ORIGINAL RESEARCH ARTICLE

Surveillance of Antidepressant Safety (SADS): Active Signal Detection of Serious Medical Events Following SSRI and SNRI Initiation Using Big Healthcare Data

Mia Aakjær[1] · Marie Louise De Bruin[2,3] · Murat Kulahci[4,5] · Morten Andersen[1]

Abstract

Introduction The current process for generating evidence in pharmacovigilance has several limitations, which often lead to delays in the evaluation of drug-associated risks.

Objectives In this study, we proposed and tested a near real-time epidemiological surveillance system using sequential, cumulative analyses focusing on the detection and preliminary risk quantification of potential safety signals following initiation of selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs).

Methods We emulated an active surveillance system in an historical setting by conducting repeated annual cohort studies using nationwide Danish healthcare data (1996–2016). Outcomes were selected from the European Medicines Agency's Designated Medical Event list, summaries of product characteristics, and the literature. We followed patients for a maximum of 6 months from treatment initiation to the event of interest or censoring. We performed Cox regression analyses adjusted for standard sets of covariates. Potential safety signals were visualized using heat maps and cumulative hazard ratio (HR) plots over time.

Results In the total study population, 969,667 new users were included and followed for 461,506 person-years. We detected potential safety signals with incidence rates as low as 0.9 per 10,000 person-years. Having eight different exposure drugs and 51 medical events, we identified 31 unique combinations of potential safety signals with a positive association to the event of interest in the exposed group. We proposed that these signals were designated for further evaluation once they appeared in a prospective setting. In total, 21 (67.7%) of these were not present in the current summaries of product characteristics.

Conclusion The study demonstrated the feasibility of performing epidemiological surveillance using sequential, cumulative analyses. Larger populations are needed to evaluate rare events and infrequently used antidepressants.

✉ Mia Aakjær
  mia.aakjaer@sund.ku.dk

1 Department of Drug Design and Pharmacology, Pharmacovigilance Research Center, University of Copenhagen, Copenhagen, Denmark

2 Department of Pharmacy, Copenhagen Centre for Regulatory Science (CORS), University of Copenhagen, Copenhagen, Denmark

3 Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Utrecht, The Netherlands

4 Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark

5 Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden

1 Introduction

Early detection of safety problems is essential for drug safety. However, existing systems have considerable weaknesses. Among drugs launched after 1960 and later withdrawn, the median interval between the launch year and the year of the first report of an adverse drug reaction leading to drug withdrawal was 4 years (interquartile range [IQR] 1–10) [1]. Delays in detecting potential adverse drug reactions in these systems need to be reduced to increase patient safety.

The established evaluation of the safety of a drug starts before it is launched, where the human safety profile is evaluated in clinical trials. However, these have selected patient populations of limited size, are often short, and differ from real-world practice [2]. Duijnhoven et al. [3] investigated the total number of individuals in clinical trials before drug

Key Points

We proposed an active safety surveillance system and demonstrated the feasibility of using epidemiological methods for signal detection in a cumulative, sequential approach using historical data on antidepressants for a 20-year period.

Our system would complement the existing surveillance mainly based on spontaneous reporting by detecting potential safety signals as soon as they appeared in healthcare registries and allowing preliminary quantification of the risk of serious medical events.

Larger populations are needed to detect signals for rare events and infrequently used antidepressants.

approval in Europe from 2000 to 2010. They found that the median number of patients studied before a new medicine entered the market was 1708 (IQR 968–3195). This number is considered insufficient to study many drug safety problems [3]. Thus, rare adverse drug reactions cannot necessarily be detected in clinical trials, and nor can adverse effects occurring in patients with risk factors, as these are often excluded from trials.

Post-marketing surveillance is necessary to address these limitations, and current surveillance principally relies on spontaneous reporting systems (SRS) to detect new safety signals. However, these systems are entirely dependent on the identification of the adverse event (suspicion) and the subsequent submission of individual case safety reports by clinicians, other healthcare professionals, or medicine users. It is known that adverse events in SRS are highly underreported and that the systems suffer from biased reporting and duplication of reports [2].

Finally, SRS cannot quantify risk using epidemiological measures and do not allow a direct comparison of the risk of events between drug users and non-users or users of other drugs. When a safety concern from SRS emerges, potential safety signals may be further evaluated in pharmacoepidemiological studies, where risks can be quantified. According to a study from 2014 describing signals during the first 18 months of the European Medicines Agency (EMA) Pharmacovigilance Risk Assessment Committee (PRAC), 3% of signals that required further investigation were requested to be evaluated in a pharmacoepidemiological study [4]. However, pharmacoepidemiological studies may be time consuming, which could lead to unnecessary delays in evaluating drug-associated risks.

To complement the current post-marketing surveillance and improve signal detection timeliness, active surveillance

using epidemiological methods has been suggested as a solution [5, 6]. Active systems prospectively collect data on exposure and outcomes from healthcare databases and accumulate information, enabling signal detection closer to real time [7]. Active surveillance may be performed by conducting sequential analyses, an approach already implemented in clinical trials and in vaccine safety surveillance [8, 9] and tested in real-world data [7, 10–16]. A recent study using US claims data demonstrated the advantages of routinely collected healthcare data to monitor drug safety [17]. Therefore, we aimed to develop and test a near real-time epidemiological surveillance system using sequential, cumulative analysis and evaluate its performance focusing on detection and preliminary risk quantification of potential safety signals. We tested such a system in new users of antidepressants. Selective serotonin reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs) are among the most prescribed drugs in numerous countries, and Denmark has one of the highest consumptions [18–20]. These drugs are also known to be associated with many adverse reactions [21, 22].

2 Methods

2.1 The Active Surveillance System

Our active surveillance system is based on annual sequential analyses. The evidence-generating process is divided into three parts, as depicted in Fig. 1. In phase 1 (data management), cohorts of new users are created and variables on outcomes and covariates generated. The cohorts comprise patients with the drug exposure(s) of interest and selected comparators. In phase 2a (analyses), descriptive analyses and risk quantifications are performed. These include new user characteristics, outcome incidence rates (IRs), and the estimation of crude and adjusted hazard ratios (HRs) using Cox regression models. We adjusted for age and sex and a standard set of covariates representing relevant comorbidities and co-medications (Table S1 in the electronic supplementary material [ESM]).

In phase 2b, signal detection is performed using output from the Cox regression models, applying prespecified criteria and presenting them in heat maps and plots of cumulative HRs over time. We defined a preliminary association as a statistically significant difference in relative risk expressed as an HR between exposure and comparator. From the preliminary associations, potential safety signals are identified as those significant in at least 2 consecutive years. In phase 3, the final phase (evaluation), potential safety signals with a positive association to the event of interest in the exposed group, which are unique exposure–comparator combinations across the primary and secondary analyses, are first selected
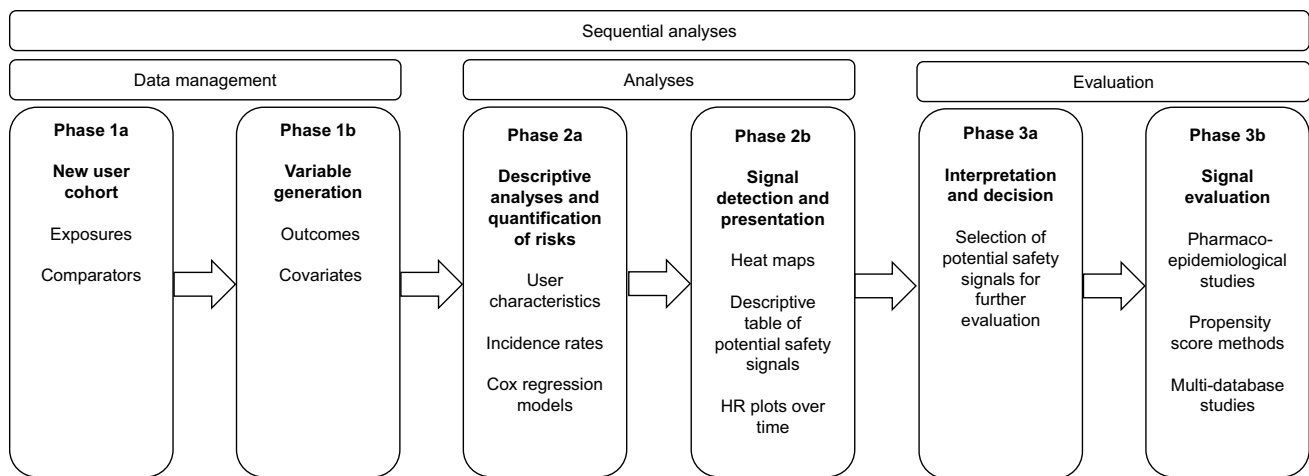
Fig. 1 Illustration of the phases of the proposed active surveillance system. *HR* hazard ratio

(3a) and then further evaluated in more advanced pharmacoepidemiological studies (3b). Also, potential signals are compared with events already listed in summaries of product characteristics (SmPCs). For this study, we chose to evaluate detected signals, compare them with information in current Danish SmPCs, and present our findings from phase 2a until and including phase 3a.

2.2 System Inputs and Test

To test our surveillance system, we used historical data on drug exposure from 1 January 1996 to 31 December 2016 but conducted the repeated annual analyses emulating a prospective setting. Thus, the analysis outputs reflect the signals identified and the risks quantified at the end of each analysis year. It is possible to adapt and tune the system by varying specific parameters (Table S2 in the ESM). These include the data source, study design and participants, study outcomes, covariates, frequency of the repeated testing, and number and type of sensitivity analyses. In the following, we present our choices of these parameters to detect potential safety signals of serious medical events following the initiation of SSRIs and SNRIs. Further details on additional parameters are provided in Table S2 in the ESM.

2.2.1 Data Sources

We used Danish healthcare register data to identify exposures, outcomes, and covariates. The unique personal identification number assigned to all residents at birth or immigration enables the linkage of information from different registers [23]. To identify drug exposure and

co-medications, we used data on all prescription drugs dispensed in community pharmacies from the Danish Prescription Registry [24]. Drugs were identified using the anatomical therapeutic chemical (ATC) classification code [25]. For information on outcomes and comorbidities, we used the Danish National Patient Register, which contains data on somatic and psychiatric inpatient and outpatient hospitalizations [26]. For outcomes, we included only primary diagnoses and inpatient hospitalizations. To obtain information on psychiatric comorbidities until the start of 1995, we used the Danish Psychiatric Central Research Register [27]. We used the *International Classification of Diseases, Eighth Edition* (ICD-8) (until 1 January 1994) and ICD-10 (onwards) for the definition of diagnoses (outcomes and covariates). Finally, we used the Cause of Death Register and the Central Person Registry for data on deaths and migration, respectively [28, 29].

2.2.2 Study Design and Participants

In the surveillance, we included new users of all SSRIs marketed in Denmark (ATC codes under N06AB) and SNRIs (N06AX16 and N06AX21), who were aged at least 18 years. To define new users, i.e., no previous use of SSRIs and SNRIs, we applied a washout period of at least 1 year starting on 1 January 1996. Patients were thus only able to enter the study once. They were followed for a maximum of 6 months from treatment initiation to the event of interest or censoring using an intention-to-treat approach. We excluded patients who had experienced the medical event of interest within 5 years before the first SSRI or SNRI dispensing (index date). We used individual drugs as exposures. In the

primary analysis, citalopram was chosen as the comparator drug since this is the most commonly used antidepressant in Denmark [20]. We compared each drug with all other SSRIs and SNRIs in the secondary analysis, allowing an HR estimate for citalopram. Furthermore, as part of the secondary analysis, we performed additional analyses for citalopram and escitalopram, excluding both of these in the comparator group.

2.2.3 Outcomes

To obtain a broad spectrum of serious medical events, we selected outcomes from the EMA designated medical event (DME) list, which is a list of medical events that are serious and often drug related. This list can be used to study any type of drug (EMA/557113/2016) [30]. We complemented this list and customized it to our exposure by adding serious medical events known from the literature and SmPCs assumed to be caused by antidepressants (Table S3 in the ESM). We excluded some of the medical events from the EMA DME list because of a lack of seriousness or diagnostic code since we could only capture events leading to hospital admission. For some events, it was necessary to use less specific diagnoses as proxies. For example, the event torsades des pointes (a life-threatening cardiac arrhythmia) is not represented with a separate ICD-10 code, so it cannot be detected in the data directly. Instead, we used the diagnoses arrhythmia, ventricular arrhythmia, and ventricular fibrillation as proxies for the event. Events were divided into system organ classes following the Medical Dictionary of Regulatory Activities structure. We considered two broad groups of outcomes, psychiatric and somatic (Table S3 in the ESM). A flowchart with further details is provided in Fig. S1 in the ESM.

2.2.4 Covariates

The analyses were adjusted for selected comorbidities and co-medications (Table S1 in the ESM). We included covariates representing major chronic diseases based on Charlson comorbidity definitions and others with potential association with the exposure and outcome [31, 32]. We created two overlapping sets of covariates used for adjustment in the analyses on somatic and psychiatric outcomes, respectively (Table S3 in the ESM). All analyses were adjusted for Charlson comorbidities and alcohol-related disorders. The analyses of somatic outcomes were additionally adjusted for hypertension, dyslipidemia, and obesity. Analyses of psychiatric outcomes were also adjusted for schizophrenia, bipolar disorders, anxiety, personality disorders, depression, abuse of other substances, and the use of antipsychotics, lithium, and anxiolytics (Table S1 in the ESM).

2.3 Statistical Analysis

Cox regression analyses were performed to estimate crude and adjusted HRs and 95% confidence intervals (CIs). The output from these regressions was visually presented in heat maps and plots of cumulative HRs with 95% CIs over time. We created heat maps on effect as the $\log_{10}$ of the HR ($\log_{10}(\mathrm{HR})$) ranging from -0.6 to 0.6. These correspond to HRs of 0.25 and 4. Other relevant values of the $\log_{10}(\mathrm{HR})$ are -0.3, 0.0, and 0.3, corresponding to HRs of approximately 0.5, 1.0, and 2. Colors reflect the direction and size of the association. Darker red indicates a positive association, and darker blue indicates negative associations. Heat maps on significance used $-\log_{10}$ of the p value ($-\log_{10}(P)$) in the range from -3 to 3, both corresponding to p values of 0.001, but with positive associations represented by positive values and negative associations by negative values. Other relevant $-\log_{10}(P)$ values are -2 and 2, both corresponding to a p value of 0.01 and -1.3, and 1.3 corresponding to a p value of 0.05, which is considered to be statistically significant in all analyses. Heat maps were presented, including all three regression models: crude (m1), age and sex-adjusted (m2), and fully adjusted (m3). Sequential analyses were performed annually. We estimated HRs (95% CIs) for 51 outcomes at 20 time points using six and nine exposure–comparator sets in the primary and secondary analysis. We chose not to adjust for multiple testing over time.

We used SAS (version 9.4) software (SAS Institute, Inc., Cary, NC, USA) for data management and analyses. Data management was done using a Nordic Common Data Model and analytics framework developed by the Pharmacovigilance Research Center, University of Copenhagen.

2.4 Ethical Considerations and Permissions

The study was approved by the Danish Data Protection Agency (0421-0022/18-7000). In Denmark, ethical approval is not required for register-based research [23].

3 Results

3.1 Phase 2a: Descriptive Analyses and Quantification of Risks

3.1.1 User Characteristics

In the total study population used for testing the surveillance system, 969,667 new users of SSRIs and SNRIs were included and followed for a total of 461,506 person-years

(Table 1). The overall mean age on the index date was 52.7 years (with a standard deviation of 20.7), and the majority of users were females (60.1%). The mean age was highest for citalopram users and lowest for paroxetine users (56.0 vs. 46.3 years). Citalopram, escitalopram, and fluvoxamine users tended to be somewhat older than users of the other antidepressants. The majority of drugs were launched before the start of the study, except for escitalopram and duloxetine, which were launched in 2002 and 2004, respectively.

Figure 2 shows the distribution of drug exposure in the annual new user cohorts. The most commonly used drug was citalopram, followed by sertraline and escitalopram. Fluvoxamine was the most seldom used. The overall number of new users decreased from 64,253 in 1997 to 25,287 in 2016 (Fig. 2).

3.1.2 Incidence Rates

Table 2 shows the number and IRs (95% CI) of the selected medical events for the cumulative data collected during the entire study period. Given the limited number of fluvoxamine users, the drug was excluded from the remaining result presentations, although it was included in the comparator group in the secondary analysis.

Fractures, ischemic stroke, arrhythmia, heart failure, and myocardial infarction were the most common medical events, occurring with IRs of 186 (181–190), 77.0 (74.4–79.6), 48.1 (46.1–50.2), 40.9 (39.1–42.8), and 37.1 (35.4–39.0) per 10,000 person-years, respectively (Table 2). Some of the medical events were not included in Table 2 or IRs were not estimated because the small numbers were incompatible with Statistics Denmark's confidentiality requirements (fewer than five events). Aplasia pure red cell, progressive multifocal leukoencephalopathy, autoimmune hepatitis, hepatic infarction, and toxic epidermal necrolysis were too rare to present total IRs. Hemolytic anemia, auto-immune hemolytic anemia (AIHA), thrombotic thrombocytopenic purpura (TTP), dermatitis exfoliative, erythema multiforme, Stevens–Johnson syndrome, and myopathy were too rare for IRs to be presented for all individual drugs and so were not included in Table 2. The total number of events for AIHA, TTP, dermatitis exfoliative, Stevens–Johnson syndrome, and myopathy ranged from seven to ten, corresponding to a total IR of 0.2 per 10,000 person-years. The total number of events of erythema multiforme, hypomania, and hemolytic anemia were 19, 22, and 31, and their IRs were 0.4 (0.2–0.6), 0.5 (0.3–0.7), and 0.7 (0.5–1.0) per 10,000 person-years, respectively. The majority of these events were observed in citalopram users (data not shown).

3.1.3 Cox Regression Models

We had 306 and 459 possible drug–medical event combinations in the primary and secondary analyses, respectively. These drug–medical event combinations were evaluated at each of the 20-year time points, equivalent to 6120 and 9180 analyses; all were analyzed in the three models (m1 crude, m2 age and sex, and m3 fully adjusted).

Table 1 Characteristics of the total cohort of Danish users of selective serotonin reuptake inhibitor and serotonin-norepinephrine reuptake inhibitor (1997–2016)

Characteristics	All (N=969,667)	Citalopram (N=520,000)	Escitalopram[a] (N=99,558)	Sertraline (N=171,289)	Venlafaxine (N=51,531)	Duloxetine[b] (N=13,571)	Paroxetine (N=59,967)	Fluoxetine (N=52,893)	Fluvoxamine (N=858)
%	100	53.6	10.3	17.7	5.3	1.4	6.2	5.5	0.1
Female sex	582,581 (60.1)	312,162 (60.0)	58,796 (59.1)	103,597 (60.5)	29,519 (57.3)	8143 (60.0)	37,127 (61.9)	32,721 (61.9)	516 (60.1)
Age, years	52.7 ± 20.7	56.0 ± 21.2	53.0 ± 20.8	48.5 ± 20.2	46.9 ± 17.5	50.4 ± 17.5	46.3 ± 17.2	47.4 ± 18.1	52.8 ± 18.5
18–39	302,094 (31.2)	139,115 (26.8)	30,457 (30.6)	65,492 (38.2)	19,357 (37.6)	3851 (28.4)	23,758 (39.6)	19,833 (37.5)	231 (26.9)
40–54	230,532 (23.8)	110,346 (21.2)	23,656 (23.8)	42,806 (25.0)	15,692 (30.5)	4353 (32.1)	18,120 (30.2)	15,296 (28.9)	263 (30.7)
55–64	127,734 (13.2)	67,407 (13.0)	13,724 (13.8)	21,194 (12.4)	7567 (14.7)	2266 (16.7)	8075 (13.5)	7382 (14.0)	119 (13.9)
65–74	116,209 (12.0)	69,621 (13.4)	11,615 (11.7)	17,547 (10.2)	4719 (9.2)	1756 (12.9)	5434 (9.1)	5412 (10.2)	105 (12.2)
>75	193,098 (19.9)	133,511 (25.7)	20,106 (20.2)	24,250 (14.2)	4196 (8.1)	1345 (9.9)	4580 (7.6)	4970 (9.4)	140 (16.3)
Person time, total, years	461,506	246,416	47,274	81,313	24,887	6325	29,179	25,690	419
First patient included	Jan 1997	Jan 1997	Aug 2002	Jan 1997	Jan 1997	Sep 2004	Jan 1997	Jan 1997	Jan 1997
Last patient included	Dec 2016	Dec 2016	Dec 2016	Dec 2016	Dec 2016	Dec 2016	Jul 2015	Dec 2016	Jul 2015

Data are presented as *n* (%) or mean ± standard deviation unless otherwise indicated

[a]Start of follow-up 2002

[b]Start of follow-up 2004

Fig. 2 Distribution of incident
selective serotonin reuptake
inhibitor and serotonin-norep-
inephrine reuptake inhibitor
users during the study period
from 1997 to 2016
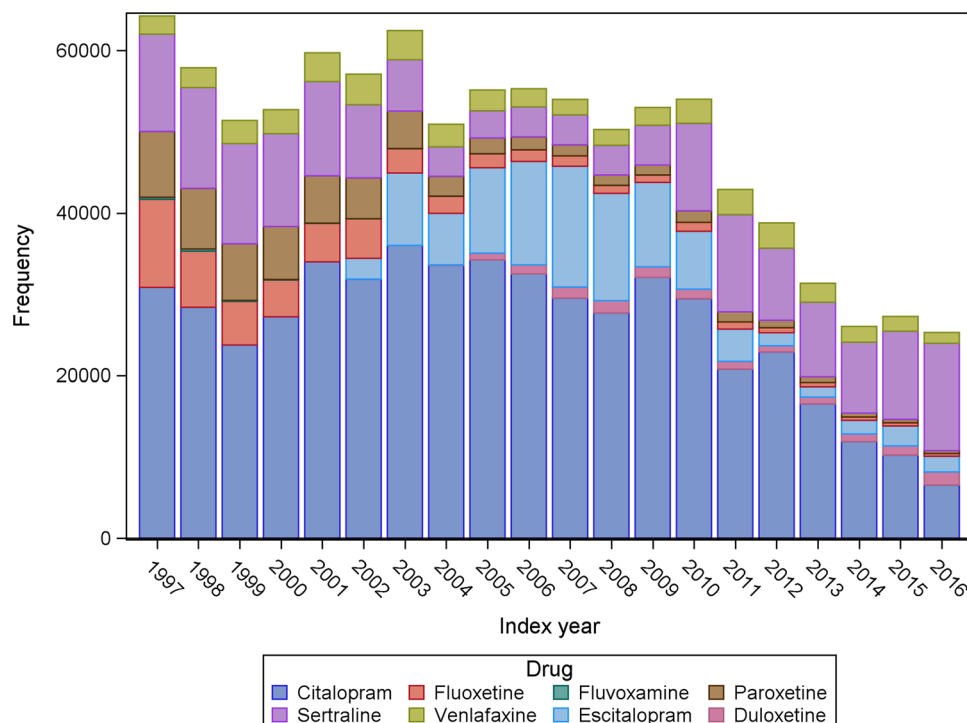
3.2 Phase 2b: Signal Detection and Presentation

3.2.1 Heat Maps

We created heat maps of all preliminary associations (sta-
tistically significant difference in HR) at each sequence (20-
year time point) to illustrate the changes over time. As an
example, Fig. 3 shows the patterns in 2006 and 2016. Heat
maps 3a and 3b are colored by effect as the $\log_{10}$ of the
HR, where a darker red indicates a more positive association
and darker blue indicates a more negative association. Heat
maps 3c and 3d are colored by significance as the $-\log_{10}$
of the p value, where darker red indicates a higher level of
significance (lower p value) for positive associations, and
darker blue indicates a higher level of significance for nega-
tive associations. Comparing heat maps colored by effect
with heat maps colored by significance, we see that only
some of the potential safety signals (preliminary associa-
tions significant in at least 2 consecutive years) that appear
in the effect heat maps are statistically significant. We can
also see that many potential safety signals disappear after
adjustment (m2 age and sex adjusted, m3 fully adjusted),
e.g., sertraline–myocardial infarction and venlafaxine–myo-
cardial infarction (Fig. 3d). From these figures, we observe
that more signals appear as data accumulate (e.g., fluox-
etine–epilepsy/seizures) and that some signals even disap-
pear (e.g., sertraline–hepatotoxicity).

3.2.2 Descriptive Signal Table

In the fully adjusted models of the primary and secondary
analyses, we detected 51 potential safety signals (prelimi-
nary associations significant in at least 2 consecutive years)
out of 765 exposure–event combinations that were unique
exposure–event combinations across the analyses. Of these,
31 were positive associations (Table 3). Examples of plots of
HRs (95% CI) over time from the primary analysis are pro-
vided for the potential safety signal venlafaxine–arrhythmia
(Fig. 4a) and for the transient signal sertraline–heart failure
(Fig. 4b). Potential safety signals were detected with IRs as
low as 0.9 (0.4–1.9) per 10,000 person-years at the first year
of detection (sertraline–hepatic necrosis).

In the primary analysis using citalopram as comparator,
we identified 46 preliminary associations, i.e., significant
in at least 1 year during the study period after full adjust-
ment. Of these, 31 were potential safety signals (significant
in at least 2 consecutive years). The mean number of years
since the start of surveillance to the first year as a significant
potential safety signal was 7.7 (IQR 5.0–10.0) years.

In the secondary analysis using all other SSRIs as com-
parator, we found 33 potential safety signals (preliminary
associations significant in at least 2 consecutive years). Of
these, 30 were also identified in the primary analysis. Six
potential safety signals with a positive association were
identified for citalopram: cardiac arrest, epilepsy and sei-
zures, fractures, hepatic failure, ischemic stroke, and renal
failure. Apparent negative associations when comparing the

Table 2 Number, incidence rates per 10,000 person-years, and 95% confidence intervals for each drug–medical event combination (phase 2a) for the whole study period 1997–2016, where the number of events was greater than or equal to five

Medical event	Total	Citalopram	Escitalopram	Sertraline	Venlafaxine	Duloxetine	Paroxetine	Fluoxetine
Blood and lymphatic system disorders								
Aplastic anemia[a]	45; 1.0 (0.7–1.3)	33; 1.3 (0.9–1.9)	5; 1.1 (0.3–2.5)	NA	NA	NA	NA	NA
ITP[a]	14; 0.3 (0.2–0.5)	7; 0.3 (0.1–0.6)	5; 1.1 (0.3–2.5)	NA	NA	NA	NA	NA
Thrombocyto-penia	88; 1.9 (1.5–2.4)	56; 2.3 (1.7–3.0)	12; 2.5 (1.3–4.4)	9; 1.1 (0.5–2.1)	5; 2.0 (0.7–4.7)	NA	NA	NA
Agranulocytosis[a]	33; 0.7 (0.5–1.0)	16; 0.6 (0.4–1.1)	6; 1.3 (0.5–2.8)	NA	NA	NA	NA	NA
Immune system disorders								
Anaphylaxis[a]	47; 1.0 (0.7–1.4)	25; 1.0 (0.7–1.5)	6; 1.3 (0.5–2.8)	NA	NA	NA	NA	NA
Psychiatric disorders								
Mania	123; 2.7 (2.2–3.2)	64; 2.6 (2.0–3.3)	10; 2.1 (1.0–3.9)	21; 2.6 (1.6–3.9)	7; 2.8 (1.1–5.8)	NA	12; 4.1 (2.1–7.2)	NA
Hallucinations	31; 0.7 (0.5–1.0)	19; 0.8 (0.5–1.2)	NA	5; 0.6 (0.2–1.4)	NA	NA	NA	NA
Delusions[a]	457; 10.0 (9.1–10.9)	216; 8.8 (7.7–10.1)	53; 11.3 (8.4–14.7)	78; 9.6 (7.6–12.0)	27; 10.9 (7.2–15.9)	8; 12.7 (5.5–25.0)	38; 13.1 (9.3–18.0)	37; 14.4 (10.2–19.9)
Suicide	841; 18.2 (17.0–19.5)	440; 17.9 (16.2–19.6)	93; 19.7 (15.9–24.1)	145; 17.8 (15.0–21.0)	66; 26.5 (20.5–33.7)	13; 20.6 (10.9–35.1)	46; 15.8 (11.5–21.0)	38; 14.8 (10.5–20.3)
Self-harm	320; 7.0 (6.2–7.8)	155; 6.3 (5.4–7.4)	37; 7.9 (5.5–10.8)	49; 6.1 (4.5–8.0)	31; 12.5 (8.5–17.8)	NA	NA	24; 9.3 (6.0–13.9)
Nervous system disorders								
Epilepsy and seizures	1278; 28.1 (26.6–29.7)	883; 36.4 (34.0–38.8)	116; 24.8 (20.5–29.8)	153; 19.0 (16.1–22.3)	39; 15.8 (11.3–21.6)	9; 14.4 (6.6–27.2)	41; 14.2 (10.2–19.3)	37; 14.5 (10.2–20.0)
Hemorrhagic stroke	663; 14.5 (13.4–15.6)	410; 16.8 (15.2–18.5)	88; 18.8 (15.1–23.1)	88; 10.9 (8.7–13.4)	29; 11.7 (7.8–16.8)	7; 11.1 (4.5–22.9)	20; 6.9 (4.2–10.6)	21; 8.2 (5.1–12.5)
Ischemic stroke	3340; 77.0 (74.4–79.6)	2189; 96.7 (92.7–108.5)	334; 74.8 (67.0–83.3)	442; 56.2 (51.1–61.7)	118; 48.5 (40.2–58.1)	30; 48.8 (32.9–69.6)	104; 36.3 (29.7–44.0)	123; 49.0 (40.7–58.4)
Polyneuropathy	79; 1.7 (1.4–2.1)	39; 1.6 (1.1–2.2)	9; 1.9 (0.9–3.6)	13; 1.6 (0.9–2.7)	NA	NA	9; 3.1 (1.4–5.9)	5; 2.0 (0.6–4.6)
Cardiac disorders								
Arrhythmia	2111; 48.1 (46.1–50.2)	1317; 56.7 (53.7–59.9)	226; 50.5 (44.2–57.6)	279; 35.6 (31.6–40.1)	103; 42.6 (34.8–51.7)	29; 47.9 (32.1–68.7)	81; 28.5 (22.6–35.4)	76; 30.2 (23.8–37.8)
Ventricular arrhythmia	193; 4.2 (3.6–4.8)	114; 4.6 (3.8–5.6)	23; 4.9 (3.1–7.3)	30; 3.7 (2.5–5.3)	7; 2.8 (1.1–5.8)	NA	NA	10; 3.9 (1.9–7.2)
Ventricular fibrillation[a]	46; 1.0 (0.7–1.3)	30; 1.2 (0.8–1.7)	5; 1.1 (0.3–2.5)	6; 0.7 (0.3–1.6)	NA	NA	NA	NA
Myocardial infarction	1679; 37.1 (35.4–39.0)	1109; 46.1 (43.4–48.9)	174; 37.6 (32.2–43.6)	216; 27.0 (23.5–30.8)	56; 22.8 (17.2–29.6)	16; 25.7 (14.7–41.7)	62; 21.5 (16.5–27.6)	46; 18.1 (13.3–24.2)
Cardiomyopathies	141; 3.1 (2.6–3.6)	82; 3.3 (2.7–4.1)	23; 4.9 (3.1–7.3)	14; 1.7 (0.9–2.9)	7; 2.8 (1.1–5.8)	NA	11; 3.8 (1.9–6.8)	NA
Cardiac arrest[a]	247; 5.4 (4.7–6.1)	159; 6.5 (5.5–7.5)	30; 6.4 (4.3–9.1)	38; 4.7 (3.3–6.4)	7; 2.8 (1.1–5.8)	NA	NA	5; 1.9 (0.6–4.5)

Table 2 (continued)

Medical event	Total	Citalopram	Escitalopram	Sertraline	Venlafaxine	Duloxetine	Paroxetine	Fluoxetine
Heart failure	1840; 40.9 (39.1–42.8)	1215; 50.9 (48.1–53.9)	185; 40.2 (34.6–46.4)	256; 32.0 (28.2–36.2)	55; 22.3 (16.8–29.1)	11; 17.7 (8.8–31.6)	54; 18.7 (14.0–24.4)	64; 25.2 (19.4–32.2)
Vascular disorders								
Pulmonary hypertension[a]	26; 0.6 (0.4–0.8)	14; 0.6 (0.3–1.0)	NA	7; 0.9 (0.3–1.8)	NA	NA	NA	NA
Respiratory, thoracic, and mediastinal disorders								
Pulmonary fibrosis[a]	93; 2.0 (1.6–2.5)	57; 2.3 (1.8–3.0)	10; 2.1 (1.0–3.9)	11; 1.4 (0.7–2.4)	8; 3.2 (1.4–6.3)	NA	NA	NA
Gastrointestinal disorders								
Intestinal perforation[a]	34; 0.7 (0.5–1.0)	24; 1.0 (0.6–1.4)	6; 1.3 (0.5–2.8)	NA	NA	NA	NA	NA
Pancreatitis acute[a]	315; 6.9 (6.1–7.7)	175; 7.1 (6.1–8.3)	40; 8.5 (6.1–11.6)	45; 5.6 (4.0–7.4)	15; 6.0 (3.4–10.0)	6; 9.5 (3.5–20.7)	19; 6.5 (3.9–10.2)	15; 5.9 (3.3–9.7)
Pancreatitis acute, sensitivity	300; 6.5 (5.8–7.3)	164; 6.7 (5.7–7.8)	39; 8.3 (5.9–11.3)	43; 5.3 (3.8–7.1)	14; 5.6 (3.1–9.5)	6; 9.5 (3.5–20.7)	19; 6.5 (3.9–10.2)	15; 5.9 (3.3–9.7)
Autoimmune pancreatitis[a]	72; 1.6 (1.2–2.0)	39; 1.6 (1.1–2.2)	11; 2.3 (1.2–4.2)	11; 1.4 (0.7–2.4)	NA	NA	NA	NA
Pancreatitis[a], all	352; 7.7 (6.9–8.5)	198; 8.1 (7.0–9.3)	45; 9.6 (7.0–12.8)	49; 6.0 (4.5–8.0)	17; 6.9 (4.0–11.0)	8; 12.7 (5.5–25.0)	20; 6.9 (4.2–10.6)	15; 5.9 (3.3–9.7)
Hepatobiliary disorders								
Hepatotoxicity	364; 7.9 (7.1–8.8)	205; 8.3 (7.2–9.6)	46; 9.8 (7.1–13.0)	57; 7.0 (5.3–9.1)	18; 7.3 (4.3–11.5)	9; 14.3 (6.5–27.1)	16; 5.5 (3.1–8.9)	13; 5.1 (2.7–8.7)
Toxic liver disease[a]	105; 2.3 (1.9–2.8)	54; 2.2 (1.6–2.9)	8; 1.7 (0.7–3.3)	22; 2.7 (1.7–4.1)	9; 3.6 (1.7–6.9)	NA	NA	5; 1.9 (0.6–4.5)
Hepatic necrosis[a]	23; 0.5 (0.3–0.7)	10; 0.4 (0.2–0.7)	NA	8; 1.0 (0.4–1.9)	NA	NA	NA	NA
Hepatic failure[a]	211; 4.6 (4.0–5.2)	134; 5.4 (4.6–6.4)	32; 6.8 (4.6–9.6)	21; 2.6 (1.6–4.0)	NA	NA	9; 3.1 (1.4–5.9)	6; 2.3 (0.9–5.1)
Acute hepatic failure[a]	63; 1.4 (1.1–1.7)	38; 1.5 (1.1–2.1)	11; 2.3 (1.2–4.2)	6; 0.7 (0.3–1.6)	NA	NA	NA	NA
Skin and subcutaneous tissue disorders								
Angio-neurotic edema[a]	67; 1.5 (1.1–1.8)	38; 1.5 (1.1–2.1)	8; 1.7 (0.7–3.3)	11; 1.4 (0.7–2.4)	NA	NA	NA	NA
Musculoskeletal and connective tissue disorders								
Fractures	7553; 186 (181–190)	5012; 233 (227–240)	748; 179 (166–192)	939; 129 (121–138)	265; 119 (105–134)	53; 93.9 (70.4–123)	261; 98.7 (87.1–111)	275; 119 (105–134)
Renal and urinary disorders								
Acute kidney injury[a]	337; 7.3 (6.6–8.1)	219; 8.9 (7.8–10.2)	49; 10.4 (7.7–13.7)	38; 4.7 (3.3–6.4)	9; 3.6 (1.7–6.9)	NA	10; 3.4 (1.6–6.3)	NA
Renal failure[a]	813; 17.8 (16.6–19.0)	542; 22.2 (20.4–24.2)	108; 23.1 (18.9–27.9)	92; 11.4 (9.2–13.9)	20; 8.1 (4.9–12.5)	7; 11.2 (4.5–23.0)	17; 5.8 (3.4–9.3)	27; 10.5 (6.9–15.3)
Hyponatremia	1257; 27.4 (25.9–29.0)	791; 32.3 (30.1–34.6)	143; 30.4 (25.6–35.8)	182; 22.5 (19.3–26.0)	51; 20.6 (15.3–27.0)	20; 31.8 (19.4–49.1)	44; 15.1 (11.0–20.3)	26; 10.1 (6.6–14.9)

Data are presented as *N*; incidence rate (95% confidence interval). Number of events less than five was marked with NA due to confidentiality requirements from Statistics Denmark

ITP immune thrombocytopenic purpura, *NA* not available

[a]Events from the European Medicines Agency's Designated Medical Event list
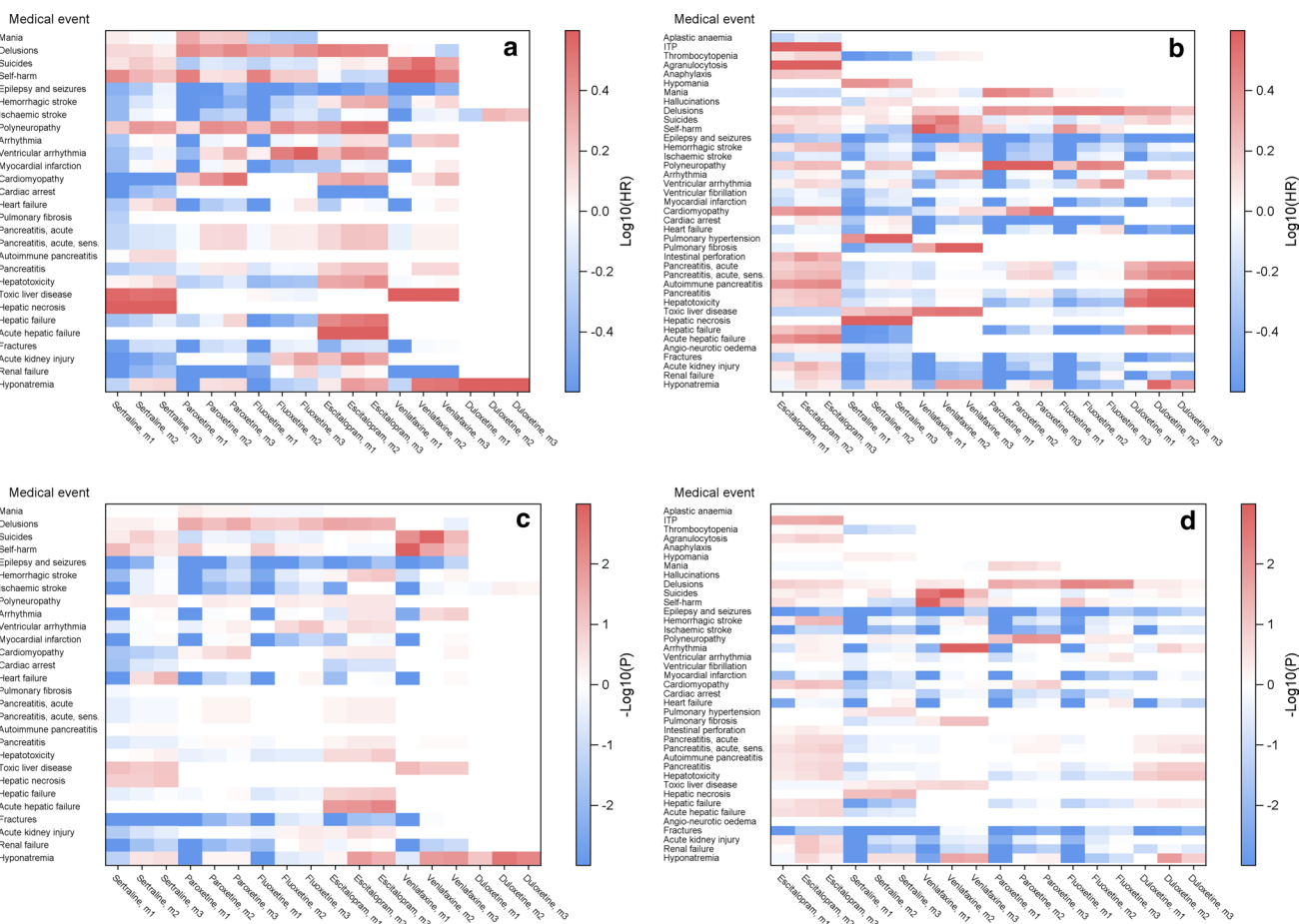
Fig. 3 Illustration of preliminary associations of the primary analysis in all three models: crude (m1), age and sex (m2), and fully adjusted (m3) in (**a**, **c**) 2006 and (**b**, **d**) 2016. Heat maps show events occurring at least five times. Figures (**a**) and (**b**) are colored by $\log_{10}$ of the hazard ratio ($\log_{10}$(HR)) ranging from -0.6 to 0.6 that correspond to HR of 0.25 and 4. Other relevant values of the $\log_{10}$(HR) are -0.3, 0.0, and 0.3, corresponding to HRs of 0.5, 1.0, and 2. Darker red indicates a more positive association, and darker blue indicates more negative associations. Figures (**c**) and (**d**) are colored by signifi-

cance as $\log_{10}$ of the *p* value ($-\log(P)$) that ranges from -3 to 3, both corresponding to *p* values of 0.001, but for negative associations and positive associations, respectively. Other relevant values are -2 and 2, corresponding to *p* values of 0.01 and -1.3 and 1.3 corresponding to *p* values of 0.05. **a** Colored by $\log_{10}$ of the HR in 2006, **b** Colored by $\log_{10}$ of the HR in 2016, **c** colored by significance as $\log_{10}$ of the *p* value in 2006, **d** colored by significance as $\log_{10}$ of the *p* value in 2016. *HR* hazard ratio, *ITP* immune thrombocytopenic purpura, *sens* sensitivity

exposure of interest with citalopram in the primary analysis reappeared as a positive association in citalopram when comparing with all other SSRIs: epilepsy and seizures, fractures, and ischemic stroke.

In the additional analysis investigating citalopram and escitalopram with a comparator group excluding both, a new potential safety signal with a positive association appeared for citalopram (hepatic failure). For escitalopram, two new signals with positive associations appeared (acute kidney injury and intestinal perforation).

3.3 Phase 3a: Interpretation and Decision

We detected 31 unique exposure–comparator combinations of potential safety signals (preliminary associations

significant in at least 2 consecutive years) that showed positive associations to the event of interest in the exposed group (Table 3). Of the 31 unique combinations, 20 were significant at the end of the study in at least one of the exposure–comparator combinations. This means that 11 potential safety signals with different exposure–comparator combinations with positive associations disappeared during the study period (e.g., duloxetine–hyponatremia, fluoxetine–ventricular arrhythmia, and citalopram–renal failure). Since this study was historical, these 11 potential safety signals would have required further investigation in a prospective setting. Of the 31 unique combinations, our system identified 21 (67.7%) new potential safety signals that are not listed in the current Danish SmPCs (Table S4 in the ESM). These were all medical events from the EMA DME list or added

Table 3 Potential safety signals and the year in the 20 sequential analyses in which they appeared the first and last time, including the hazard ratio and the 95% confidence interval at the last appearance

Antidepressants (years of surveillance) and serious medical event	First appearance of potential signal (years)			Last appearance of potential signal (years)			HR (95% CI), last year as a potential signal		
	P	S	A	P	S	A	P	S	A
Duloxetine (13)									
Fractures	5	5		13*	13*		0.68 (0.52–0.89)	0.72 (0.55–0.94)	
Hyponatremia[b]	4	4		8	8		**1.86 (1.07–3.24)**	**1.88 (1.08–3.25)**	
Suicides[b]	5	5		5	5		**2.06 (1.06–4.00)**	**2.05 (1.06–3.97)**	
Escitalopram (15)									
Acute hepatic failure[a]	5	5	5	6	6*	5	**2.37 (1.00–5.63)**	**2.41 (1.05–5.53)**	**2.25 (1.00–5.04)**
Acute kidney injury[a]			7			15*			**1.49 (1.02–2.16)**
Cardiomyopathy[c]		10	10		15*	14		**1.61 (1.02–2.51)**	**1.75 (1.02–3.01)**
Delusions[a]	4	4		5	5		**1.59 (1.05–2.39)**	**1.49 (1.01–2.21)**	
Epilepsy/seizures	4	8		15*	11		0.78 (0.64–0.95)	0.81 (0.66–0.99)	
Heart failure			7			13			0.83 (0.70–1.00)
Hepatic failure[a]		10	5		15*	15*		**1.52 (1.04–2.21)**	**2.13 (1.34–3.38)**
Fractures	8			15*			0.91 (0.85–0.99)		
Hemorrhagic stroke[c]	13	6	6	15*	15*	9	**1.28 (1.02–1.62)**	**1.30 (1.04–1.62)**	**1.34 (1.01–1.77)**
Hyponatremia[b]	6			6			**1.38 (1.07–1.78)**		
ITP[a]	11	11	11	15*	15*	15*	**4.04 (1.27–12.8)**	**4.97 (1.66–14.9)**	**8.08 (1.55–42.2)**
Intestinal perforation[a]			15			15*			**4.27 (1.16–15.7)**
Ischemic stroke	7	8		8	8		0.86 (0.76–0.99)	0.88 (0.77–1.00)	
Renal failure[a]		14	9		15*	15*		**1.23 (1.00–1.51)**	**1.42 (1.11–1.82)**
Fluoxetine (20)									
Delusions[a]	4	4		20*	20*		**1.61 (1.13–2.30)**	**1.52 (1.09–2.14)**	
Epilepsy/seizures	5	6		20*	8		0.69 (0.49–0.96)	0.69 (0.48–0.99)	
Myocardial infarction	8	8		8	8		0.70 (0.50–0.98)	0.72 (0.52–1.00)	
Ventricular arrhythmia[a]	3	3		8	8		**2.11 (1.04–4.29)**	**2.14 (1.08–4.23)**	
Paroxetine (20)									
Cardiomyopathy[c]		2			14			**1.93 (1.03–3.64)**	
Delusions[a]	9	11		20*	15		**1.42 (1.00–2.02)**	**1.42 (1.01–1.99)**	
Epilepsy/seizures	6			20*			0.71 (0.52–0.98)		
Fractures	10	18		20*	19		0.84 (0.74–0.95)	0.88 (0.77–1.00)	
Heart failure	7	7		7	7		0.65 (0.46–0.94)	0.67 (0.47–0.96)	
Ischemic stroke	5			20*			0.78 (0.64–0.96)		
Polyneuropathy[c]	15	15		20*	20*		**2.74 (1.30–5.79)**	**2.63 (1.30–5.33)**	
Renal failure	9	9		20*	13		0.61 (0.38–1.00)	0.50 (0.28–0.92)	
Sertraline (20)									
Cardiomyopathy		10			18			0.50 (0.26–0.95)	
Epilepsy/seizures	19			20*			0.78 (0.66–0.93)		
Fractures	5	5		20*	20*		0.84 (0.78–0.90)	0.86 (0.81–0.93)	
Heart failure[c]	11	11		14	14		**1.20 (1.03–1.40)**	**1.22 (1.05–1.42)**	
Hepatotoxicity[b]	2	2		4	3		**1.90 (1.04–3.45)**	**1.79 (1.02–3.13)**	
Hepatic necrosis[a]		19			20*			**2.66 (1.12–6.30)**	
Renal failure	6	6		15	15		0.80 (0.64–1.00)	0.75 (0.57–0.99)	
Toxic liver disease[a]		2			3			**1.73 (1.01–2.97)**	
Venlafaxine (20)									
Arrhythmia[a]	13	13		20*	20*		**1.43 (1.17–1.75)**	**1.41 (1.16–1.73)**	
Epilepsy/seizures	4			9			0.70 (0.51–0.97)		
Hyponatremia[b]	4	4		20*	20*		**1.40 (1.05–1.87)**	**1.38 (1.04–1.83)**	

Table 3 (continued)

Antidepressants (years of surveillance) and serious medical event	First appearance of potential signal (years)			Last appearance of potential signal (years)			HR (95% CI), last year as a potential signal		
	P	S	A	P	S	A	P	S	A
Pulmonary fibrosis[a]		18			20*			**2.30 (1.11–4.79)**	
Suicides[b]	18	9		20*	20*		**1.31 (1.00–1.71)**	**1.33 (1.03–1.71)**	
Citalopram (20)									
Cardiac arrest[a]		10	10		12	10		**1.46 (1.02–2.09)**	**1.73 (1.11–2.71)**
Delusions		10			12			0.79 (0.63–0.98)	
Epilepsy/seizures[b]		5	5		20*	20*		**1.33 (1.18–1.51)**	**1.35 (1.18–1.56)**
Fractures[b]		5	5		20*	20*		**1.13 (1.08–1.19)**	**1.16 (1.09–1.22)**
Hepatic failure[a]		20				20*			**1.49 (1.06–2.11)**
Hyponatremia		6	6		18	7		0.88 (0.78–0.99)	0.84 (0.72–0.98)
Ischemic stroke[c]		16	20		20*	20*		**1.11 (1.03–1.19)**	**1.11 (1.02–1.20)**
Renal failure[a]		6	6		8	20		**1.35 (1.05–1.75)**	**1.27 (1.06–1.52)**
Self-harm		4	4		4	4		0.49 (0.28–0.87)	0.65 (0.42–0.99)

Bold medical events and HRs (95% CI) indicate signals selected for further assessment. Blank fields illustrate analysis not relevant or that no potential safety signal was detected

A additional secondary analysis for citalopram and escitalopram excluding both in the comparator group, *CI* confidence interval, *DME* designated medical event, *EMA* European Medicines Agency, *HR* hazard ratio, *ITP* immune thrombocytopenic purpura, *S* secondary comparing each drug with all others excluding drug of interest, *SmPC* summary of product characteristics, *SNRI* serotonin-norepinephrine reuptake inhibitor, *SSRI* selective serotonin reuptake inhibitor, *P* primary comparing each drug with citalopram

*Significant signal ($p < 0.05$) the last year of surveillance

[a]From the EMA DME list

[b]Found in SmPCs for some SSRIs or SNRIs

[c]Medical events missing from the EMA DME list and the SmPCs added because of importance

because of importance (Table 3). Ten potential safety signals concerned events listed in the current SmPCs, of which two were also present in the EMA DME list (ventricular arrhythmia and arrhythmia). Finally, 82 of our events were present in the SmPCs but not detected as signals by our system (Table S4 in the ESM). Further details on each medical event and their appearance on the EMA DME list and the SmPCs are provided in Table S5 in the ESM.

4 Discussion

4.1 The Active Surveillance System

We developed and tested an active surveillance system based on annual sequential analyses using Danish healthcare data to detect potential safety signals (preliminary associations significant in at least 2 consecutive years) closer to real time. We emulated a prospective scenario following new users of SSRIs and SNRIs using 20 years of historical data. Among potential safety signals, IRs were as low as 0.9 (0.4–1.9) per 10,000 person-years at the first year of detection. The system identified 51 potential safety signals, counted as unique exposure–medical event combinations, across the analyses.

Of these, 31 showed positive associations to the event of interest in the exposed group, and 20 of them were present at the end of the study. We propose that all signals indicating positive associations should have been further assessed when they appeared in a prospective setting.

In 2007, Brown et al. [6] evaluated the utility of healthcare data for near real-time surveillance and found that prospective sequential evaluation of routinely collected data can provide estimates of drug–medical event rates to support routine and timely surveillance. They investigated five known drug–event associations and two pairs of negative controls and used a statistical test developed for sequential data, the maximized sequential probability ratio test. Our study differed from their study by including a broad spectrum of serious medical events often associated with drug exposure, of which most were included from the EMA DME list. Furthermore, we used a definition of a potential safety signal inspired by traditional pharmacoepidemiological risk measures, i.e., a significantly increased HR in 2 consecutive years based on multivariable Cox regression analyses. Recent studies have used epidemiological approaches to investigate how active systems can be applied to healthcare data and found them useful as complements to the existing surveillance systems [13, 16, 17]. Similar to our study,

Fig. 4 Plot of hazard ratios and
confidence intervals for venla-
faxine compared with citalo-
pram and the association with
arrhythmia (**a**) and sertraline
compared with citalopram and
the association with heart fail-
ure (**b**). *CI* confidence interval,
HR hazard ratio
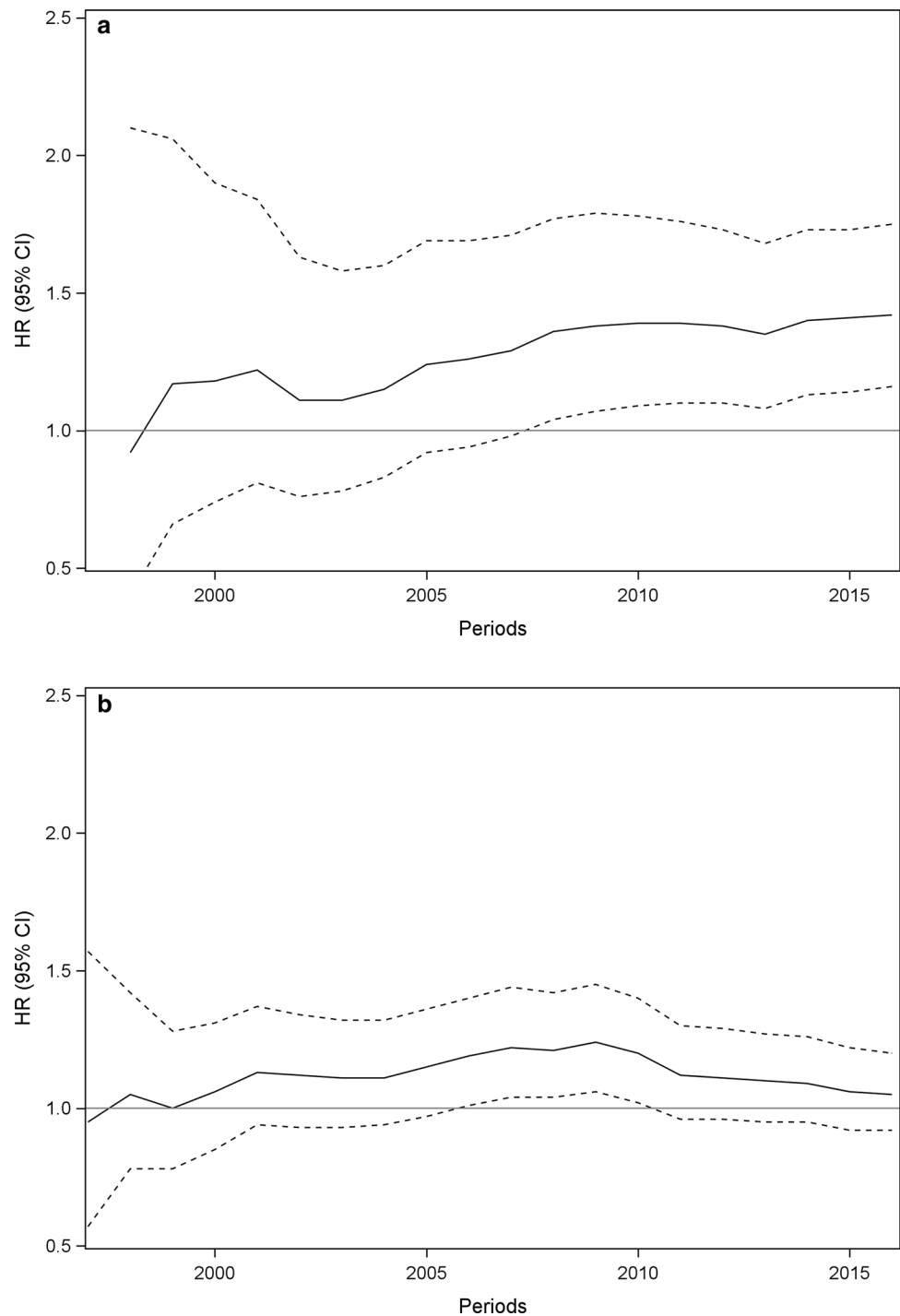
these three studies also evaluated safety by estimating HRs.
Our study differed by focusing on detecting potential safety
signals among many different medical events rather than
evaluating one or a few specified outcomes. We also sought
to extend signal detection by applying pharmacoepidemio-
logical methods, usually known from signal evaluation, to
provide a preliminary risk quantification as a threshold for
signal detection.

4.2 Strengths and Limitations of the Active Surveillance System

Our proposed system's primary strength compared with
existing surveillance systems is that it uses routinely col-
lected electronic secondary healthcare data to assess the
associations with serious medical events. It does so by ana-
lyzing data as they accumulate. Theoretically, this enables

the system to obtain a preliminary risk estimate more quickly than conventional pharmacoepidemiological studies. The comprehensive data cover all serious medical events leading to hospital admission. Thus, our system requires no additional data collection but only human resources to initiate and maintain the surveillance [13]. Underreporting is avoided because the system does not require evaluation of individual patients, clinical suspicion of specific exposure–event relationships, willingness to report, or extra effort to collect information and submit individual case safety reports. When including all patients in Denmark treated with SSRIs and SNRIs, the system had sufficient statistical precision to identify rare events with IRs as low as 0.2 (0.1–0.3) per 10,000 person-years for the entire drug group but detected potential safety signals with an IR above 0.9 (0.4–1.9) per 10,000 person-years. Our system allows the detection of potential safety signals as data accumulate and provides epidemiological measures of absolute and relative risk (in the form of IRs and HRs) useful in risk/benefit evaluations. Furthermore, comparisons are adjusted for confounding using multivariable models in contrast to the subjective evaluation of alternative causes of the event used in case reports.

Traditional pharmacoepidemiological studies will usually be initiated when a safety concern has been identified. After this, different time-consuming steps will follow before the results are ready for publication: the study will have to be planned and designed, data and permissions must be applied for, the applications must be handled and data retrieved before access can be provided, and, finally, the analysis can be conducted. This may lead to an unnecessary delay between the detection of the signal and the final evaluation. Currently, there is approximately a 1-year lag time in the data availability at Statistics Denmark and 1–2 months at the Danish Health Data Authority [33] in addition to the handling of the data application at the authorities. We suggest that an active surveillance system should have general access to healthcare data and analyze data as they are continuously entered in the registries. This will provide an updated preliminary epidemiological risk estimate at the chosen time intervals. If so, the delay will be reduced to the delay between data generation and analysis. However, for the rare events included in the EMA DME list, it will still take years to detect a signal in our system.

Our system also has limitations. A general limitation is its inability to detect non-serious medical events since these do not appear in our registries. Furthermore, 11 (35.5%) of the potential safety signals that were identified disappeared before the end of the study. Most of them were significant in only a few years (e.g., sertraline–heart failure; Fig. 4b). This means that the signals were either false positives or true risks that were consequently well-managed, e.g., by contraindicating the drugs in high-risk patients. We identified

21 new potential safety signals for events that were not listed in the current Danish SmPCs. Also, 82 drug–medical event combinations listed in the SmPCs were not detected by our system (Table S4 in the ESM). However, some of the events were detected as signals in closely related medical terms, e.g., cardiac arrest and ventricular arrhythmia, which were both listed in the SmPCs, but the latter was not detected by the system. Furthermore, we used an active comparator design in our system and therefore estimated individual associations of the event of interest between two similar drugs rather than class effects. Finally, it could also indicate that the undetected medical events listed in the SmPCs have limited evidence and might not actually be true risks. Hence, our system should not be considered as a replacement of the existing system but rather a complement, especially since most of the detected potential safety signals were not listed in the SmPCs. However, assessing our system's direct added value over existing surveillance systems was beyond the scope of this project. We recommend that future analyses incorporate comparisons with a list of gold standard signals (e.g., those resulting in regulatory action) and assessment of false-positive, false-negative signals and timing of the potential safety signal.

We tested the system using a comprehensive dataset, allowing confounder adjustment through information on each patient's long-term history of comorbidities and co-medications. A general set of confounders relevant to all the included medical events is difficult to establish. Therefore, we decided to divide the confounders into two standard adjustment sets, for somatic and psychiatric events, respectively. Furthermore, some of our events were too rare to include all covariates in the multivariable analysis. This stresses the need for more advanced confounder adjustment methods such as propensity score methods, as done in other recent studies [13, 16, 17].

Moreover, though we adjusted for numerous potential confounders, residual confounding is possible. Also, national guidelines and unmeasured baseline risks may influence the choice of drug, leading to confounding by indication. This is especially relevant to consider in events such as suicides and self-harm but also in the remaining psychiatric events investigated in this study (delusions, hallucinations, mania, and hypomania) that can be associated with the depression itself rather than the drug use.

4.3 Strengths and Limitations of Design and Analysis Choices

We opted to use an intention-to-treat approach of 6 months, which may have caused exposure misclassification. This could be reduced by shortening the follow-up, as suggested in Table S2 in the ESM. Using a time-dependent exposure approach might have enabled the detection of more signals.

Our system was tuned to focus on medical events occurring shortly after drug initiation, and we did not consider long-term medical events such as cancers that would require years of follow-up and time-dependent exposure models, including cumulative exposure. However, some of the very acute events that we investigated, such as anaphylaxis, may have been better investigated using cumulative incidence proportions and as binary outcomes using logistic regression (Table S2 in the ESM). To avoid misclassification of outcomes, we limited our diagnosis to inpatient admissions and main diagnoses. Furthermore, we noticed that some potential safety signals either appeared in their narrow or broad definitions, such as fluoxetine–ventricular arrhythmia and venlafaxine–arrhythmia, respectively (Table 3). This could indicate that using too broad or unspecific definitions may dilute the effect of a signal. However, using a very narrow definition could also reduce the power to detect a signal.

We chose to define a potential safety signal based on the relative risk between the exposure and comparator expressed as HRs statistically significant for at least 2 consecutive years and to use two sets of comparators. In the primary analysis, we compared each drug with citalopram, whereas the secondary analysis compared each drug with all others in the groups. In general, for many of the exposure–event combinations, the change in comparator changed neither the estimate nor the time of detection (e.g., all potential safety signals that appeared for duloxetine). However, in some cases, the choice of comparator affected especially the time of detection (e.g., escitalopram–hemorrhagic stroke). A more general limitation in the choice of comparator appears in drugs that are unique in their class where no comparator is available. In this situation, we recommend a comparison with related drug classes or a sample of the general population. Furthermore, we did not adjust for multiple testing since our system was developed to provide a preliminary risk estimate closer to real time, similar to what would have been obtained by a pharmacoepidemiological study rather than assessing whether the accumulated data were sufficient to make a definitive decision, i.e., terminate the study as in clinical trials [34]. Finally, our surveillance was only tested in a single country. We acknowledge that clinical practice may vary between countries but believe that similar systems can be developed in other countries or regions with databases similar to Denmark's.

4.4 Future Perspectives

The system could be extended to assess all drugs newly introduced to the market and all serious medical events that lead to hospital admission. Instead of prespecifying the medical events of interest, newer methods within data mining known from other recent studies, e.g., tree-based scan statistics, could also be applied, selecting a broad range of outcomes, for example, diseases represented by all the third- or fourth-level ICD-10 codes [35, 36]. Additionally, the implementation of data from general practice could enable the system to detect less serious medical events that do not necessarily lead to hospital admission.

Many design and analysis choices must be made when conducting pharmacoepidemiological studies [37]. We outlined several fine-tuning possibilities (Table S2 in the ESM), which should all be further assessed to design the most comprehensive active surveillance system for Danish healthcare data. We propose that automated sensitivity analyses focusing on outcome definitions and follow-up are performed to check the robustness of the findings. Sensitivity analyses of outcome definitions could check whether the definitions are too nonspecific, especially since some of our definitions, such as arrhythmia, are very broad. In contrast, sensitivity analysis of the follow-up will address exposure misclassification. Finally, augmenting the system with confounder adjustment based on propensity scores should be evaluated.

5 Conclusions

Our study demonstrates the feasibility of performing epidemiological surveillance of medical events in Danish big healthcare data using sequential, cumulative analyses. Our active surveillance system could be a useful supplement to the existing passive system. However, several fine-tuning possibilities should be further assessed. The system is especially suitable for the detection of serious medical events among users of commonly prescribed drugs. However, larger populations are needed to evaluate rare events and infrequently used drugs.

Declarations

References

1. Onakpoya IJ, Heneghan CJ, Aronson JK. Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. BMC Med [Internet]. 2016;14. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4740994/.

2. Gibbons RD, Amatya AK, Brown CH, Hur K, Marcus SM, Bhaumik DK, et al. Post-approval drug safety surveillance. Annu Rev Public Health. 2010;31:419–37.

3. Duijnhoven R, Straus S, Raine J, Boer A, Hoes A, Bruin M. Number of patients studied prior to approval of new medicines: a database analysis. PLoS Med. 2013;10:e1001407.

4. Pacurariu AC, Coloma PM, van Haren A, Genov G, Sturkenboom MCJM, Straus SMJM. A description of signals during the first 18 months of the EMA pharmacovigilance risk assessment committee. Drug Saf. 2014;37(12):1059–66.

5. Platt R, Madre L, Reynolds R, Tilson H. Active drug safety surveillance: a tool to improve public health. Pharmacoepidemiol Drug Saf. 2008;17(12):1175–82.

6. Brown JS, Kulldorff M, Chan KA, Davis RL, Graham D, Pettus PT, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. Pharmacoepidemiol Drug Saf. 2007;16(12):1275–84.

7. Gagne JJ, Rassen JA, Choudhry NK, Bohn RL, Patrick AR, Sridhar G, et al. Near-real-time monitoring of new drugs: an application comparing prasugrel versus clopidogrel. Drug Saf. 2014;37(3):151–61.

8. Lieu TA, Kulldorff M, Davis RL, Lewis EM, Weintraub E, Yih K, et al. Real-time vaccine safety surveillance for the early detection of adverse events. Med Care. 2007;45(10 Supl 2):S89-95.

9. Greene SK, Kulldorff M, Lewis EM, Li R, Yin R, Weintraub ES, et al. Near real-time surveillance for influenza vaccine safety: proof-of-concept in the Vaccine Safety Datalink Project. Am J Epidemiol. 2010;171(2):177–88.

10. Chrischilles EA, Gagne JJ, Fireman B, Nelson J, Toh S, Shoaibi A, et al. Prospective surveillance pilot of rivaroxaban safety within the US Food and Drug Administration Sentinel System. Pharmacoepidemiol Drug Saf. 2018;27(3):263–71.

11. Gagne JJ, Glynn RJ, Rassen JA, Walker AM, Daniel GW, Sridhar G, et al. Active safety monitoring of newly marketed medications in a distributed data network: application of a semi-automated monitoring system. Clin Pharmacol Ther. 2012;92(1):80–6.

12. Gagne JJ, Wang SV, Rassen JA, Schneeweiss S. A modular, prospective, semi-automated drug safety monitoring system for use in a distributed data environment. Pharmacoepidemiol Drug Saf. 2014;23(6):619–27.

13. Mayer F, Kirchmayer U, Coletta P, Agabiti N, Belleudi V, Cappai G, et al. Safety and effectiveness of direct oral anticoagulants versus vitamin K antagonists: pilot implementation of a Near-Real-Time Monitoring Program in Italy. J Am Heart Assoc. 2018;7(6):e008034.

14. Rodríguez-Martín S, Martín-Merino E, Lerma V, Rodríguez-Miguel A, González O, González-Herrada C, et al. Active surveillance of severe cutaneous adverse reactions: a case-population approach using a registry and a health care database. Pharmacoepidemiol Drug Saf. 2018;27(9):1042–50.

15. Schneeweiss S, Gopalakrishnan C, Bartels DB, Franklin JM, Zint K, Kulldorff M, et al. Sequential monitoring of the comparative effectiveness and safety of dabigatran in routine care. Circ Cardiovasc Qual Outcomes. 2019;12(2):e005173.

16. Toh S, Reichman ME, Graham DJ, Hampp C, Zhang R, Butler MG, et al. Prospective postmarketing surveillance of acute myocardial infarction in new users of saxagliptin: a population-based study. Diabetes Care. 2018;41(1):39–48.

17. Patorno E, Gopalakrishnan C, Brodovicz KG, Meyers A, Bartels DB, Liu J, et al. Cardiovascular safety of linagliptin compared with other oral glucose-lowering agents in patients with type 2 diabetes: a sequential monitoring programme in routine care. Diabetes Obes Metab. 2019;21(8):1824–36.

18. Kantor ED, Rehm CD, Haas JS, Chan AT, Giovannucci EL. Trends in prescription drug use among adults in the United States from 1999–2012. JAMA. 2015;314(17):1818–31.

19. Rotermann M, Sanmartin C, Hennessy D, Arthur M. Prescription medication use by Canadians aged 6 to 79. Health Rep. 2014;25(82):9.

20. Forns J, Pottegård A, Reinders T, Poblador-Plou B, Morros R, Brandt L, et al. Antidepressant use in Denmark, Germany, Spain, and Sweden between 2009 and 2014: incidence and comorbidities of antidepressant initiators. J Affect Disord. 2019;15(249):242–52.

21. Gartlehner G, Thieda P, Hansen RA, Gaynes BN, Deveaugh-Geiss A, Krebs EE, et al. Comparative risk for harms of

second-generation antidepressants: a systematic review and meta-analysis. Drug Saf. 2008;31(10):851–65.

22. Carvalho AF, Sharma MS, Brunoni AR, Vieta E, Fava GA. The safety, tolerability and risks associated with the use of newer generation antidepressant drugs: a critical review of the literature. Psychother Psychosom. 2016;85(5):270–88.

23. Thygesen LC, Daasnes C, Thaulow I, Brønnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving. Scand J Public Health. 2011;39(7 Suppl):12–6.

24. Kildemoes HW, Sørensen HT, Hallas J. The Danish National Prescription Registry. Scand J Public Health. 2011;39(7 Suppl):38–41.

25. WHO Collaborating Centre for Drug Statistics Methodology. ATC/DDD Index [Internet]. WHO Collaborating Centre for Drug Statistics Methodology. https://www.whocc.no/atc_ddd_index/. Accessed 22 Sept 2020.

26. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. Scand J Public Health. 2011;39(7 Suppl):30–3.

27. Mors O, Perto GP, Mortensen PB. The Danish Psychiatric Central Research Register. Scand J Public Health. 2011;39(7 Suppl):54–7.

28. Helweg-Larsen K. The Danish Register of Causes of Death. Scand J Public Health. 2011;39(7 Suppl):26–9.

29. Pedersen CB. The Danish Civil Registration System. Scand J Public Health. 2011;39(7 Suppl):22–5.

30. European Medicines Agency. Signal management [Internet]. European Medicines Agency. 2018. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/signal-management. Accessed 23 Sept 2020.

31. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40(5):373–83.

32. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi J-C, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43(11):1130–9.

33. Bliddal M, Broe A, Pottegård A, Olsen J, Langhoff-Roos J. The Danish Medical Birth Register. Eur J Epidemiol. 2018;33(1):27–36.

34. Martin D, Gagne JJ, Gruber S, Izem R, Nelson JC, Nguyen MD, et al. Sequential surveillance for drug safety in a regulatory environment. Pharmacoepidemiol Drug Saf. 2018;27(7):707–12.

35. Harpaz R, DuMouchel W, Shah NH, Madigan D, Ryan P, Friedman C. Novel data-mining methodologies for adverse drug event discovery and analysis. Clin Pharmacol Ther. 2012;91(6):1010–21.

36. Kulldorff M, Dashevsky I, Avery TR, Chan AK, Davis RL, Graham D, et al. Drug safety data mining with a tree-based scan statistic. Pharmacoepidemiol Drug Saf. 2013;22(5):517–23.

37. Schneeweiss S, Rassen JA, Brown JS, Rothman KJ, Happe L, Arlett P, et al. Graphical depiction of longitudinal study designs in Health Care Databases. Ann Intern Med. 2019;170(6):398–406.