Received: 15 May 2023 | Accepted: 23 October 2023

DOI: 10.1111/bcpt.13955

FOCUSED REVIEW

BCPT Basic & Clinical Pharmacology & Toxicology

Epidemiological surveillance of drug safety using cumulative sequential analysis in electronic healthcare data

Mia Aakjær[1] | **Marie Louise De Bruin**[2] | **Morten Andersen**[1]

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

[2]Utrecht Institute for Pharmaceutical Sciences (UIPS), Division of Pharmacoepidemiology and Clinical Pharmacology, Utrecht University, Utrecht, the Netherlands

Correspondence
Mia Aakjær, Pharmacovigilance Research Center, Department of Drug Design and Pharmacology, University of Copenhagen, Jagtvej 160, DK-2100 Copenhagen, Denmark.
Email: mia.aakjaer@sund.ku.dk

Funding information
Novo Nordisk Foundation, Grant/Award Number: NNF15SA0018404

Abstract

Background: Methods for safety signal detection in electronic healthcare data analysing data sequentially are being developed to meet the limitations of spontaneous reporting systems.

Objectives: This study aims to provide an overview of the literature on sequential analysis of electronic healthcare data and describe the development and testing of a novel epidemiological surveillance system.

Methods: We searched Medline, Embase, PubMed, Scopus, Web of Science, and the Cochrane Library applying similar in- and exclusion criteria as those of a previous systematic review. The proposed system consisted of repeated cohort studies and was tested in an emulated prospective setting. Two signal evaluations were performed with several sensitivity analyses and a target trial emulation.

Findings: In the literature, 11 studies analysed the data sequentially of which two applied traditional epidemiological methods. Epidemiological surveillance of several exposures and outcomes can be successfully conducted with the newly proposed sequential analysis of electronic healthcare data. Signal evaluation studies confirmed the results of the system.

Conclusions: Very few studies in the literature analysed data at multiple time points, although this seems to be a prerequisite for testing the methods in a realistic setting. We demonstrated the feasibility of a sequential surveillance system using electronic healthcare data.

KEYWORDS
adverse drug reactions, pharmacoepidemiology, pharmacovigilance, postmarketing surveillance

1 | INTRODUCTION

Routine safety surveillance comprises mainly spontaneous reporting systems that consist of individual case safety reports filled by healthcare professionals or medicine users. The aim of spontaneous reporting systems is to provide early and new information on drug-associated risks. This early information is based on signal detection, "the process of looking for and/or identifying signals using data from any source",[1] in which a safety signal is "information on a new or known adverse event that may be caused by a medicine and requires further

investigation."[1] Spontaneous reporting systems are important for newly marketed drugs and in situations where medical events or drug exposures are rare. Spontaneous reporting systems have well-known limitations such as underreporting and biased reporting as the reporting is dependent on the healthcare professional or medicine user's suspicion about a relationship between a specific drug and the event and the subsequent fill of the report.

Electronic healthcare data has increased in availability, and to meet some of the limitations of spontaneous reporting, new methods utilizing electronic healthcare data are being developed for safety surveillance. Aronson et al. compared definitions of surveillance in drug safety and proposed the following definition: "a form of non-interventional public health research, consisting of a set of processes for the continued systematic collection, compilation, interrogation, analysis and interpretation of data on benefits and harms (including relevant spontaneous reports, electronic medical records and experimental data)."[2] Hence, it is distinguished from monitoring by including populations rather than individuals. Surveillance can be performed at any time during the lifecycle of a drug and is not restricted to a single or particular period of time.[2] Electronic healthcare databases can be used for screening or data mining of drug-event associations and several studies have investigated methods applicable for this purpose.[3] In a routine surveillance setting, data will be analysed repeatedly at multiple time points as they accumulate. In studies that evaluate the performance of screening methods in a realistic surveillance setting, the analysis should be performed sequentially to emulate a near real-time analysis.[4,5] To assess how early signals are detected, sequential analysis is essential.

A surveillance system conducts analyses of systematically collected data at regular intervals as it becomes available to enable early detection of signals. The performance of such a sequential system should be evaluated in a similar context, as it has been done with disproportionality analysis in spontaneous reporting.[6] The aim of this focused review is to provide an overview of the literature on sequential analysis performed in electronic healthcare data, describe the development and testing of a novel epidemiological safety surveillance system and discuss possible further improvements to our system and future perspectives of similar initiatives.

2 | SAFETY SIGNAL DETECTION IN ELECTRONIC HEALTHCARE DATA

Several methods have been proposed for safety signal detection in electronic healthcare data. These methods are described in a review by Arnaud et al. from 2017. They divided the methods into seven overall categories: disproportionality analysis, traditional pharmacoepidemiological designs, sequence symmetry analysis, sequential statistical testing, temporal association rules, supervised machine learning and tree-based scan statistics.[3] The authors highlighted that no method is superior but when considering whether the method 1) achieves the goal of signal detection, 2) is understandable to stakeholders and 3) provides guidance to help stakeholders handle the enormous amount of expected signals to be detected, the sequence symmetry analysis was proposed as the most promising method for signal detection in electronic healthcare data. The method is simple, easy to understand and, because of the self-controlled design, it is also likely to reduce the detection of false positives. However, it is less robust towards protopathic and indication bias.[3]

Recently, the list of methods for signal detection in electronic healthcare data was updated by Coste et al. They further compared the performance of the methods overall and for specific types of exposures and outcomes.[7] The authors concluded that no method was superior for all drugs and outcomes and propose that more than one method may be applied for signal detection among multiple drug-outcome pairs. Furthermore, they highlighted that only one of their included studies investigated the timeliness of signal detection.[7] Therefore, we investigated which of the papers included in the review by Coste et al. analysed their data at multiple time points as the authors did not focus on this aspect. Additionally, on 3 March 2023, we updated the systematic search conducted by Coste et al. to include papers available after 13 July 2021. Thus, we applied the same search strategy in the same databases (Medline and Embase via OVID, PubMed, Scopus, Web of Science and the Cochrane Library).[7] We also applied the same in- and exclusion criteria screening the identified records first by title and abstract followed by a full-text screening (Supplementary Figure S1). We identified 570 new papers after 13 July 2021, after removal of duplicates. After title and abstract screening, 39 papers were left and 17 were included after full-text screening. Of these, three studies were review papers (Supplementary Figure S1).

3 | SEQUENTIAL ANALYSIS OF DATA

Surveillance in healthcare databases using sequential analyses may be conducted with other purposes than signal detection, that is, screening for new drug-event associations among a broader range of drugs and conditions.

The objective may also be to assess the safety and effectiveness of a new drug compared with an existing drug using predefined outcomes,[4] or to address emerging safety concerns, estimating the relative risk of one or a few specific predefined events in a target drug comparing with a suitable reference.[8] In this review, we chose to focus only on drug surveillance in healthcare data with the purpose of signal detection.

In total, 11 original papers analysed their data at multiple time points including both papers from the systematic search conducted by Coste et al. and the updated search. Furthermore, one review paper was identified. The studies used different approaches covering the maximized sequential probability ratio test (maxSPRT),[5,9,10] the conditional sequential sampling procedure (CSSP),[11,12] shrinkage observed-to-expected (OE) ratios,[13] the Weibull Shape Parameter plotting hazard rates over time,[14] sequence symmetry analysis tested cumulatively,[15,16] a self-controlled design with incidence rates over time,[17] and finally, our own study, a new-user cohort design plotting hazard ratios and heat maps of hazard ratios over time.[18] These studies are described in more detail in the following sections and listed in Table 1. Of the 11 identified studies, three compared different methods.[10,12,16] In addition to these studies, a review compared studies using group sequential methods with either regression adjustment or weighting as confounder control.[19]

3.1 | The maximized sequential probability ratio test (maxSPRT)

In 2007, Brown et al. tested a refinement of the sequential probability ratio test (SPRT). SPRT is a sequential hypothesis test in which a signal appears if the likelihood ratio exceeds a certain threshold and the observation ends if the likelihood crosses a certain lower limit. The refinement, the maxSPRT, uses an alternative hypothesis of a relative risk greater than one instead of a single alternative hypothesis with a specific value. This means that the maxSPRT does not require a predefined specification of the level of the increased risk. With maxSPRT, a signal is generated if the log likelihood ratio reached a critical value established based on an α level, for example, 0.05. In the test of the maxSPRT, the authors used automated claims data for the prospective monitoring. They identified increased risks in four of five known drug-event associations and no signals in the two negative controls, suggesting that these findings support the potential for claims data in drug safety surveillance.[5]

Key methodological considerations of this maxSPRT were investigated in a later paper by Brown and colleagues using health plan data. Seven drug-event associations were selected to assess the performance of the method of which seven were known or suspected associations and two were negative controls. The authors suggest that specifications unique to the different drug-medical event combinations are needed and that there will not be a one-size fits all solution for signal detection using maxSPRT.[9]

In 2011, Kulldorff et al. compared the maxSPRT method with the classic SPRT using historical vaccine data. They compared the methods investigating fever and neurological symptoms following the Pediarix™ vaccine. The authors concluded that the maxSPRT works well for vaccine surveillance with good statistical power and timely signal detection.[10]

3.2 | Conditional sequential sampling procedure (CSSP)

Lingling Li proposed a practical group sequential method, the CSSP, in 2009. The CSSP method is a semi-parametric Poisson regression model that adjusts for multiple testing and was proposed to test for excess risk when little historical data is available. This proposal was motivated by issues identified in a study testing the maxSPRT method. The maxSPRT method requires a large amount of historical data to provide reliable estimates. The authors tested the method on administrative claims data using simulation in which they could evaluate the performance in many different scenarios. Moreover, the method was applied to the use of rofecoxib and the potential risk of acute myocardial infarction compared with the risk among users of diclofenac and naproxen. The authors conclude that the CSSP is especially useful if both the drug of interest and the comparator drug are new or when the comparator is prescribed to a different patient population, as this will not require historical data.[11]

The CSSP method proposed in 2009 was compared with a log-linear model with Poisson distribution (LLMP) in a study by Zhou et al. in 2018 in which they used data from administrative health claim databases. The LLMP is a parametric-based method where a generalized linear model is fitted within the available study population with the log of the expected number of events at each interim analysis. The two methods were tested on 50 combinations of drugs and outcomes of which nine were considered known associations and the remaining 41 were negative controls. The authors concluded that both methods have the potential for safety surveillance; however, the LLMP was often better and faster for identifying known associations and that it may be an alternative to or complement the CSSP method.[12]

TABLE 1 Papers analysing data at multiple time points.

Paper	Method	Concept	Conclusion	Comments
Brown et al. 2007[5]	Maximized sequential probability ratio test (MaxSPRT)	Requires information about the number of observed and expected adverse events under the assumption that the null hypothesis (of no excess risk) is correct	Prospective and periodic evaluation can provide estimates of event rates benefitting timely after marketing surveillance for selected adverse drug events	Observed and expected number of outcomes plotted over time
Brown et al. 2009[9]	MaxSPRT	The alternative hypothesis is composite with the relative risk defined as being greater than one	Unique specifications for each drug-event pair is needed. No one-size fits all approach will be applied.	
Kulldorff et al. 2011[10]	MaxSPRT compared to classical SPRT	See Brown et al. 2009[9]	Works well across a range of relative risks	Log-likelihood ratio plotted over time
Lingling Li 2009[11]	Conditional sequential sampling procedure (CSSP)	Semi-parametric Poisson regression model for the number of incident events	Flexible approach with wide application	
Zhou et al. 2018[12]	CSSP compared to a Poisson log-linear model (LLMP)	LLMP is a parametric-based method in which parameters are estimated in a generalized linear model using the maximum likelihood method	No method correctly identifies all drug-outcome pairs but LLMP provides better ability and shorter time for identification	Interim analyses were conducted
Norén et al. 2010[13]	Shrinkage observed-to-expected (OE) ratios	Shrinkage introduces a bias towards the null and obtains better variance.	Clear advantages of using the same shrinkage OE ratio for pattern discovery	Logarithm of the OE ratios over time plotted
Sauzet et al. 2013[14]	Weibull Shape Parameter (WSP)	Weibull distribution with a shape parameter indicating if the hazard is not constant	With rare events this, the method may depend on the reporting time accurately reflecting the true event onset time.	Hazard rates over time
Wahab et al. 2014[16]	Sequence symmetry analysis compared to disproportionality methods	Asymmetry in sequence of dispensing between two medicines or a drug dispensing and the occurrence of a hospital event	Administrative claims data may complement methods known from spontaneous reporting systems	Tested time to signal detection. Cumulatively tested yearly
Arnaud et al 2018[15]	Sequence symmetry analysis	Asymmetry in sequence of drug dispensing and the occurrence of a hospital event	Performed well with a promising potential for routine use	Adjusted sequence ratios plotted by quarter
Zhou et al. 2018[17]	Self-controlled case series (SCCS)	Incidence rate ratios are calculated comparing the rate of events after exposure to the rate of events in unexposed periods	Potential for early signal identification	Changes in incidence rate ratio over time can be plotted to check for earlier signal identification
Aakjær et al. 2021[18]	Cohort study	Cumulative hazard ratios repeated yearly	Demonstration of epidemiological surveillance	No assessment of timeliness

3.3 | Shrinkage observed-to-expected ratios

In a paper from 2010, Norén et al. propose a statistical shrinkage transformation that can be applied to OE ratios that reflects how an observed pattern deviates from an expected pattern.[13] The method is used for temporal pattern discovery in electronic patient records, and has been further described in subsequent papers.[20,21] OE ratios are particularly sensitive when the expected number of events is low. This can be stabilized by statistical shrinkage. The authors visualize the evolution over time by plotting separately both the logarithm of the OE ratio with shrinkage and the observed and expected number of events for several drug-event pairs, aligned according to the time of drug prescription. In this study, the analysis at multiple time points was an integral part of the signal detection method using temporal association, but OE ratios with shrinkage may also be useful for sequential follow-up as illustrated with data from the WHO Global Individual Case Safety Reports database, VigiBase.[20] The authors highlight how this approach ensures practical relevance in addition to transparency of the patterns and will protect against the generating of many false positive signals.[20]

3.4 | Weibull Shape Parameter

In a paper from 2013, Sauzet and colleagues assessed the performance of the Weibull Shape Parameter for signal detection in which hazard rates were plotted over time.[14] The Weibull Shape Parameter tool uses time-to-event data but does not require a reference group. It is based on the Weibull Shape Parameter test. Using the Weibull distribution, the null hypothesis is that the shape parameter is equal to one, and thus, a shape parameter will indicate if the hazard is not constant when the shape parameter is different from a value of one. A signal arises when a significant shape parameter is obtained, that is, the p value is less than the defined significance level. The Weibull Shape Parameter tool is then applied to the data, which is censored at specified time points and until the end of the observation period. The aim of the study was to explore the use of the Weibull Shape Parameter tool using real data and provide guidance on how to fit the tool and interpret the results. In the test of the tool, four events related to the use of bisphosphonates were selected. These were headache, musculoskeletal pain, alopecia, and carpal tunnel syndrome. The study illustrated the usability and reliability of the Weibull tool but also that more work is needed on the definition of the censoring periods.[14] The time-dependent analysis of hazard rates in this study was a part of the signal detection method itself and not performed to evaluate repeated analyses in a surveillance setting.

3.5 | Cumulative sequence symmetry analysis

Wahab and colleagues compared time with signal detection of adverse drug reactions in a spontaneous reporting system and a claims database in 2014 by comparing different methods. The test cases were rofecoxib-induced myocardial infarction and rosiglitazone-induced heart failure. They demonstrated that the four methods investigated (the sequence symmetry analysis, reporting odds ratio, proportional reporting ratio, and Bayesian techniques) detected signals within 1–3 years after the drugs entered the market. Thus, the methods detected signals earlier than results from clinical trials, which did not lead to withdrawal or a warning before 5 and 7 years after rofecoxib and rosiglitazone entered the market, respectively. The findings from this study indicate the potential for sequence symmetry analysis to complement methods known from spontaneous reporting systems to guide stakeholders in the decision-making process.[16]

Arnaud et al.[15] investigated a system with automated signal detection based on sequence symmetry analysis in longitudinal healthcare data, selecting non-insulin glucose-lowering drugs as a case study. Data analysis was performed quarterly and detected signals prioritized according to their relevance using the Longitudinal-SNIP (L-SNIP) algorithm, which is based on strength (S), novelty (N), impact (I) of the signal, and pattern of drug use (P).[15] Detected signals were compared with positive and negative controls in a reference set of drug-event associations based on a review of Summaries of Product Characteristics (SmPCs). Of 815 associations included in the screening, 241 (29.6%) were detected as signals of which 58 (24.1%) were prioritized. When comparing with the reference set, the signal detection method had a sensitivity = 47%, specificity = 80%, positive predictive value = 33% and negative predictive value = 82%. It was concluded that the system performed well with a promising potential for routine signal detection and prioritization.

3.6 | Self-controlled case series (SCCS)

In 2018, Zhou et al. evaluated the use of the self-controlled case series (SCCS) for signal detection in data from The Health Improvement Network, a UK primary care database, and Optum, a claims database from the

US. They tested the method on five outcomes in users of desvenlafaxine and escitalopram in addition to six outcomes in users of adalimumab. They considered Signals of Disproportional Recording, that is, a positive finding when the lower level of the confidence intervals (CIs) of the incidence rate ratio was greater than 1. The authors concluded that the SCCS showed promise for signal detection and may have a potential for early signal identification.[17]

3.7 | Cumulative cohort analysis

In 2021, we, the authors, proposed a safety surveillance system based on cumulative cohort studies. Several studies have demonstrated the use of sequential surveillance of adverse events in healthcare data using a traditional cohort analysis plotting hazard ratios over time.[22–25] These studies have focused on a few outcomes, but so far, no studies have applied these methods to a wide range of medical events. Therefore, we proposed an epidemiological surveillance system based on a new-user active comparator cohort design in which the data was analysed sequentially as it accumulated and tested it emulating a prospective scenario for near real-time signal detection.[18] The surveillance system that we developed consists of data management, analysis, and evaluation (Figure 1). In the data management phase, the cohorts of new users, outcomes and covariates are generated, whereas in the analysis phase, descriptive analyses and risk quantification are performed, which leads to the detection of potential signals. Signals are presented in heat maps of hazard ratios and plots of hazard ratios over time. In the evaluation phase, signals for further evaluation are selected and investigated in more detail in advanced pharmacoepidemiological studies. The system was tested on new users of antidepressants.[18] In the following sections, this surveillance system will be described in further detail in addition to two signal evaluation studies comprising the final phase of the system to illustrate all phases of the system.

3.7.1 | Signal detection: a test of the system (Phases 1a–3a)

Electronic healthcare data from the national Danish registries from 1 January 1986 to 31 December 2016 was used to test the system using information on dispensed prescriptions from the National Prescription Registry and information on hospital admissions from the National Patient Register.[26] The inclusion period started 10 years later on 1 January 1996.

We performed repeated cohort studies with one-year intervals on new users (at least one year washout period) of selective serotonin-reuptake inhibitors (SSRIs) and serotonin-norepinephrine reuptake inhibitors (SNRIs). Patients were followed using an intention-to-treat approach for a maximum of six months from the first prescription fill until censoring (outcome, emigration, death, end of follow-up or study). New users of each drug were individually compared with new users of citalopram, and new users of all others in the class except the drug of interest. We selected 51 outcomes leading to hospital admission from the European Medicines Agency (EMA) Designated Medical Event (DME) list, a list of serious medical events that are serious and often drug-related.[27] Patients were not allowed to have had the outcome of interest at any time point prior to the index date. To adjust for confounding, we specified a general list of

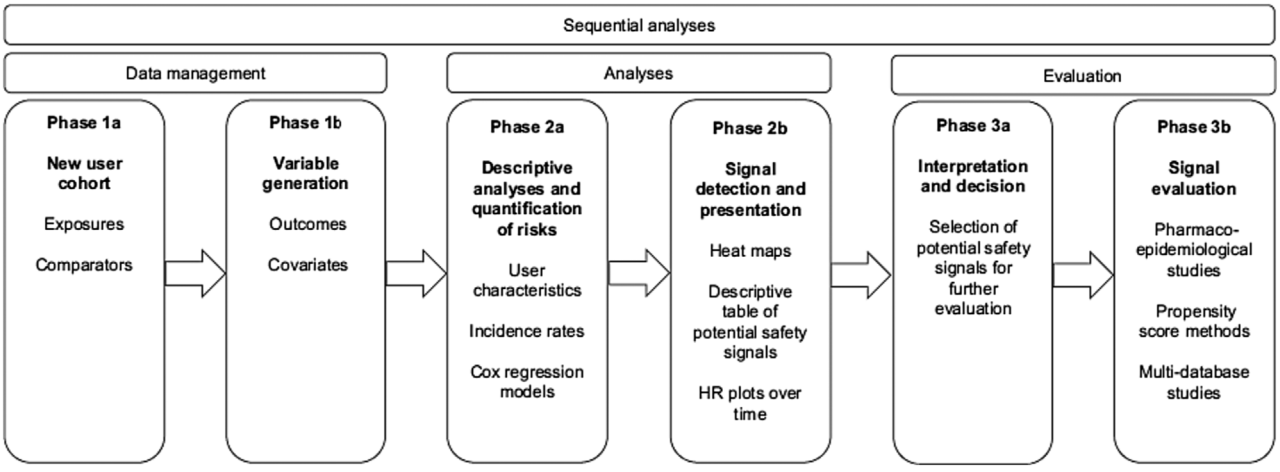

FIGURE 1 Phases of the proposed safety surveillance system. HR, hazard ratio.[18] CC BY 4.0 http://creativecommons.org/licenses/by-nc/4.0/

potential confounders that are believed to be risk factors for many medical events as many different outcomes were investigated. A crude, age-/sex-adjusted and multivariate Cox regression model was performed for each analysis providing hazard ratios with 95% CIs as epidemiological risk measures. Potential signals were defined as significant associations in at least two consecutive years, though both years include experience prior to those years. As analyses with different comparators were conducted, we defined drug-medical event combinations as unique combinations not accounting for the choice of comparator. These are, however, all parameters that can be tuned to design the most comprehensive safety surveillance system.[18]

Totally, we included 969 667 new users of SSRIs and SNRIs and identified 31 drug–medical event combinations as potential signals. Of these, more than half were considered unknown as they were not listed in the current Summary of Product Characteristics (SmPCs) and 11 were insignificant at the end of the study, meaning that they would have required further investigation in a prospective surveillance scenario. From this study, we saw how epidemiological surveillance can be conducted using annual, sequential and repeated analysis of electronic healthcare data. Our system could not identify potential signals from rare events or infrequently used drugs, which should therefore either be analysed in larger populations or identified using the spontaneous reports.[18] However, our surveillance system could be implemented in current pharmacovigilance and complement existing spontaneous reporting systems by providing preliminary risk estimates for the more common adverse events.

3.7.2 | Signal evaluation (Phase 3b)

In a preliminary version of the system with no washout period for the outcomes, an association between fluoxetine and acute pancreatitis was observed with a hazard ratio of 1.5 (95% CI 1.1–2.1). We initiated a detailed study investigating this potential association in depth.[28,29] We included 61 783 new users of fluoxetine in propensity score-matched cohorts and compared the risk of acute pancreatitis among users of fluoxetine with users of citalopram and users of other SSRIs. In this study, we did not see an increased risk of acute pancreatitis among users of fluoxetine in line with findings from a case–control study by Lancashire et al.[30] However, in a sensitivity analysis allowing patients with previous pancreatic events (no outcome washout), higher point estimates were observed in our study.[29] These higher point estimates observed in analyses without an outcome washout period may explain why an increased risk was initially observed in the preliminary[28] and not the final version of the system. This finding indicates that the potential signal identified in the preliminary version of the system might have been a false positive.[18]

In the final version of the system, a transient potential signal of cardiac arrest among new users of citalopram was detected. We wanted to investigate this potential safety signal further, as the Food and Drug Administration (FDA) and EMA previously issued safety warnings because of evidence indicating dose-dependent QT prolongation associated with citalopram[31,32] (and escitalopram).[33] Furthermore, the evidence from observational studies was conflicting. We evaluated this potential safety signal by emulating a target trial. Target trial emulation ensures that the desirable features of clinical trials such as the specification of time zero and assignment of the treatment are preserved.[34] Also, it is an efficient way to utilize the data especially if the size of the data is limited or the outcomes are rare, as the individuals are allowed to enter the study several times. Besides increasing the number of exposed patients and outcomes, selection bias is also reduced.[35–38] Overall, 257 760 person-trials corresponding to 212 309 individuals were included, meaning that 17% entered the study more than once. Overall, we found no increased risk of serious arrhythmia either among users of citalopram or escitalopram, which was in line with the majority of previous studies.[39–43] Nevertheless, when we investigated the risk before and after the safety warnings issued in 2011, lower point estimates were observed in the period after the warnings. This finding may indicate that the prescribers were actually compliant with the warnings by prescribed less citalopram and escitalopram to patients with cardiovascular risk factors.[44]

These two signal evaluation studies indicated the importance of study design and analysis choices that can be tuned in epidemiological studies. However, we believe that they also indicate the benefit and usefulness of our proposed system, since the signal detection results of the system were confirmed by the in-depth pharmacoepidemiological studies. The potential signal of fluoxetine and acute pancreatitis, which was detected in a preliminary version of the system, was not detected in the final version of the surveillance system, when recurrent outcome events were not allowed, indicating that this potential signal might have been a false positive. Similarly, the potential signal of citalopram and cardiac arrest was a transient potential signal, meaning that it was significant at some point during the study period but not at the end of the study. In the evaluation study of this potential signal, we saw a similar pattern of the risk over time with higher risk estimates before the warnings in 2011 and

lower risk estimates after, which may be explained by a shift in prescription patterns or that we investigated serious arrhythmia rather than cardiac arrest, as we believe serious arrhythmia is a more well-defined outcome. Given the very similar conclusions between the signal detection study and the evaluation of the potential signals, signal evaluation in advanced pharmacoepidemiological studies may not be needed to the same extent if the majority of the design and analysis choices are already implemented in the signal detection phase of the proposed surveillance system. We believe that signal evaluation could rather focus on validation in another database.

3.7.3 | Strengths and limitations

Major strengths of our proposed surveillance system are the use of routinely collected electronic healthcare data covering all serious medical events leading to hospitalization and the use of sequential analysis as data accumulates. The use of sequential analysis provides risk estimates during the study period, whereas risk estimates in traditional retrospective pharmacoepidemiological studies are provided for the entire study period.[18] Compared with spontaneous reporting systems, our surveillance system avoids underreporting as, for example, no clinical suspicion of an adverse event is required.[18] Furthermore, not only relative risks but also absolute risks are calculated, indicating the importance of the potential signal for the treated patient population to a greater extent, and also giving important information for a benefit–risk evaluation.

A limitation of our proposed system is the limited testing of the many tuning opportunities of design and analysis choices, for example, having citalopram as the comparator, which might not always be the most suitable.[18] Moreover, the active comparator design requires a suitable control for each exposure drug under investigation. There might, however, be situations where a suitable control is not available, which limits the generalizability of the system. In the test of the system, we focused on acute and serious events leading to hospital admission. Detecting potential signals of long-term or chronic conditions would require different analysis choices, especially regarding the exposure models, that is, time-dependent and cumulative exposure models.[18] To include less serious events, for example, events not leading to hospital admissions, other data sources such as data from general practice should be included. Moreover, in the test of our system, 35.5% of the potential signals disappeared before the end of the study indicating that they were either false positives or true risks, which were then well-managed in clinical practice. Lastly, we decided not to adjust for the multiple testing, as known from clinical trials and which have been done in some of the other previous studies analysing the data at multiple time points such as the CSSP.[11] In trials, these tests are used to decide whether the trials should be continued or not. Our purpose was to detect potential signals that should be evaluated further. The multiple testing could, however, be implemented by specifying a narrower threshold of the *p* value for a potential signal to arise. This would, however, lead to both less detection of false positive signals but it could also cause important risks to go undetected.

4 | DISCUSSION

When searching the literature on signal detection in electronic healthcare data, we found that very few studies analysed the data sequentially. The study by Wahab et al. was the only one that evaluated timeliness of signal detection,[16] and three studies compared signal detection methods.[10,12,16] In general, the authors of the included papers were positive towards the use of their methods for signal detection in electronic healthcare data. Considering the categories proposed in the review by Arnaud et al., the majority of the methods were sequential statistical testing using either maxSPRT or CSSP,[5,9–12] whereas two studies used sequence symmetry analysis[15,16] and two studies including our own analysed the data using traditional epidemiological methods with SCCS and cohort designs.[17,18] In these studies, risk estimates were available at different time points. For two of the papers in the review, however, the analysis of multiple time points was a part of the signal detection method in itself and not conducted with the purpose of assessing the performance of the method in a surveillance setting.[13,14]

Arnaud et al. highlighted three criteria for methods to be suitable for signal detection in electronic healthcare data (achievement of the goals of signal detection, understandable for stakeholders, and guidance for stakeholders to prioritize the large amount of signals).[3] Currently, no method that analyses the data sequentially has claimed to fulfill the goals of signal detection by investigating all possible combinations of drugs and medical events. Two studies investigated several drugs within a group and multiple outcomes, Arnaud et al.[15] conducting sequence symmetry analysis on noninsulin glucose-lowering drugs and our study comparing cohorts of SSRI with SNRI users.[18] In other studies, methods were in general tested on a few selected outcomes in users of specific drugs. Thus, there is still great potential for exploring the methods both in terms of drug exposures and outcomes.

The method should also be understandable to stakeholders, who are often not statisticians. Having this aspect in mind, the methods relying on simpler concepts such as the SCCS, the sequence symmetry and cumulative cohort analyses might be more appealing compared with more advanced statistical approaches such as the CSSP. Another advantage of using the traditional epidemiological methods is that they do not rely on statistical significance only, but provide a relative risk estimate, as highlighted by Arnaud et al. A relative risk estimate reflects the strength of the association and will help stakeholders prioritize the large amount of expected signals.[3] Norén et al. have proposed OE ratios for pattern discovery in observational medical data, offering a similar advantage in interpretation.[20] As OE ratios are particularly sensitive when the expected number of events is low, they propose a statistical shrinkage transformation that regularizes the measure towards the null value (one) when there are few events. This approach may also be relevant for other measures of association and in cumulative sequential analyses. However, confidentiality requirements precluded analysis of data with event counts <5 in our study.[18]

In our study, we have demonstrated with our newly proposed method, how electronic healthcare data can be applied to signal detection by developing and testing an epidemiological surveillance system. We have shown how such a surveillance system can potentially complement the existing pharmacovigilance comprising mostly of evidence from spontaneous reporting systems. We believe that safety surveillance can be conducted faster if automation of epidemiological surveillance systems is applied with limited required tuning for the individual drug-medical event combination.

4.1 | Future perspectives

The proposed surveillance system has been tested using automation incorporated into the Nordic Common Data Model developed by the Pharmacovigilance Research Center at the University of Copenhagen.[45] We used information from SmPCs to assess if the system could detect known adverse events among the DMEs. This information on prior signals and their evaluation should ideally be integrated into the system's phase 3a. Currently, information from the SmPCs is extracted manually. This step should be automatized for example using text mining methods as recently applied in Structured Products labels and regulatory approvals.[46,47] This automated extraction of information could preferably be conducted in digitalized product information, as described in a project by the EMA.[48] Also, potential signals detected by our system were not prioritized. In a system of routine use, studies of assessment evaluation processes for potential signals is needed as seen in the study by Cederholm et al.[49] If applied to all drugs and an even wider range of medical events, an automated approach for prioritizing the detected potential signals would be needed. Arnaud et al. suggested to prioritize signals according to the relevance of the signals using the Longitudinal-SNIP (L-SNIP) algorithm, which is based on strength, novelty, impact of the signal, and pattern of drug use.[15] However, the use of artificial intelligence in pharmacoepidemiological studies has increased exponentially recently and could potentially be applied to signal prioritization as well.[50]

Similarly, high-dimensional propensity scores, an extension of traditional propensity score selection that identifies a large set of proxies to be included in the propensity score model,[51] or a data-driven approach for the selection of confounders based on the identification of risk factors for the individual medical events of interest should be applied rather than confounders being selected based on the current literature or the use of a general set of potential risk factors applying to all medical events in our setting.

Assessing the system's added value and evaluation of timeliness was beyond the scope of our studies. Our reference with regard to knowledge of drug-event associations was the information in SmPCs. An alternative could be used to publish observational findings suggesting an association as positive controls as seen previously in studies by Ryan et al.[52,53] We do recommend that future analyses should incorporate analyses with, for example, a list of signals that are well-established and preferably resulted in regulatory actions. Ideally, this should also include information on signals of which it was concluded that no regulatory action was needed.[54] Timeliness could be assessed by comparing the time of detection in our surveillance system with the detection through the spontaneous reporting systems.

Lastly, once the potential signals have been generated, a future system should include analyses stratifying the analysis into relevant subgroups, for example, older adults or patients with a certain diagnosis, to identify if the risk is higher in certain patient groups. Again, these many subgroup analyses should be performed using artificial intelligence because of the huge amount of factors to investigate. These stratified analyses will help stakeholders in their subsequent work with the potential signals.

5 | CONCLUSION

Very few studies available in the current literature on signal detection in electronic healthcare data analyse the data at multiple time points, evaluate the timeliness of

signal detection and compare methods. This indicates the need to further test safety surveillance systems as we believe that this might complement existing pharmacovigilance that currently mostly comprises evidence generated through spontaneous reporting systems. We have demonstrated the feasibility of a safety surveillance system analysing the data sequentially using an epidemiological approach in electronic healthcare data. We evaluated two potential signals leading to insights into how such a system should be designed.

As surveillance is per definition done sequentially, it should also be tested sequentially. We recommend that surveillance is tested by assessing the performance of the methods, thus, whether the method is able to identify true risks without the detection of many false positive and negative signals for example by comparing the detected signals to well-known associations and finally, to compare how different methods perform. We should be able to conduct safety surveillance faster if we automate epidemiological approaches in surveillance systems that require limited tuning of design and analysis choices for each drug-event pair.

ACKNOWLEDGEMENTS

This study was partly funded by a grant from the Novo Nordisk Foundation (NNF15SA0018404) to the University of Copenhagen.

CONFLICT OF INTEREST STATEMENT

M.L.D.B. is employed by Utrecht University to conduct research under the umbrella of the Utrecht Centre for Pharmaceutical Policy and Regulation. This centre receives no direct funding or donations from private parties, including the pharma industry. Research funding from public–private partnerships (e.g., IMI, The Escher Project; http://escher.lygature.org/) is accepted under the condition that no company-specific study is conducted. The centre has received unrestricted research funding from public sources (e.g., World Health Organization, Netherlands Organization for Health Research and Development, the Dutch National Health Care Institute, EC Horizon 2020, the Dutch Medicines Evaluation Board, and the Dutch Ministry of Health). M.An. has participated in research projects funded by AstraZeneca, H. Lundbeck & Mertz, Novartis, and Pfizer, and has received fees for leading courses and teaching from Atrium, the Danish. Association of the Pharmaceutical Industry. M.A. declared no competing interests for this work.

ORCID

Mia Aakjær https://orcid.org/0000-0002-6381-6715
Marie Louise De Bruin https://orcid.org/0000-0001-9197-7068
Morten Andersen https://orcid.org/0000-0001-7029-2860

REFERENCES

1. Anonymous. *Signal management. European medicines agency*. 2018. Accessed March 10, 2021. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/signal-management
2. Aronson JK, Hauben M, Bate A. Defining "surveillance" in drug safety. *Drug Saf*. 2012;35(5):347-357. doi:10.2165/11597590-000000000-00000
3. Arnaud M, Bégaud B, Thurin N, Moore N, Pariente A, Salvo F. Methods for safety signal detection in healthcare databases: a literature review. *Expert Opin Drug Saf*. 2017;16(6):721-732. doi:10.1080/14740338.2017.1325463
4. Gagne JJ, Rassen JA, Choudhry NK, et al. Near-real-time monitoring of new drugs: an application comparing prasugrel versus clopidogrel. *Drug Saf*. 2014;37(3):151-161. doi:10.1007/s40264-014-0136-0
5. Brown JS, Kulldorff M, Chan KA, et al. Early detection of adverse drug events within population-based health networks: application of sequential testing methods. *Pharmacoepidemiol Drug Saf*. 2007;16(12):1275-1284. doi:10.1002/pds.1509
6. Bate A, Lindquist M, Edwards IR, et al. A Bayesian neural network method for adverse drug reaction signal generation. *Eur J Clin Pharmacol*. 1998;54(4):315-321. doi:10.1007/s002280050466
7. Coste A, Wong A, Bokern M, Bate A, Douglas IJ. Methods for drug safety signal detection using routinely collected observational electronic health care data: a systematic review. *Pharmacoepidemiol Drug Saf*. 2023;32(1):28-43. doi:10.1002/pds.5548
8. Dore DD, Seeger JD, Arnold CK. Use of a claims-based active drug safety surveillance system to assess the risk of acute pancreatitis with exenatide or sitagliptin compared to metformin or glyburide. *Curr Med Res Opin*. 2009;25(4):1019-1027. doi:10.1185/03007990902820519
9. Brown JS, Kulldorff M, Petronis KR, et al. Early adverse drug event signal detection within population-based health networks using sequential methods: key methodologic considerations. *Pharmacoepidemiol Drug Saf*. 2009;18(3):226-234. doi:10.1002/pds.1706
10. Kulldorff M, Davis RL, Kolczak M, Lewis E, Lieu T, Platt R. A maximized sequential probability ratio test for drug and vaccine safety surveillance. *Seq Anal*. 2011;30(1):58-78. doi:10.1080/07474946.2011.539924
11. Li L. A conditional sequential sampling procedure for drug safety surveillance. *Stat Med*. 2009;28(25):3124-3138. doi:10.1002/sim.3689
12. Zhou X, Bao W, Gaffney M, Shen R, Young S, Bate A. Assessing performance of sequential analysis methods for active drug safety surveillance using observational data. *J Biopharm Stat*. 2018;28(4):668-681. doi:10.1080/10543406.2017.1372776
13. Norén GN, Hopstadius J, Bate A, Star K, Edwards IR. Temporal pattern discovery in longitudinal electronic patient records. *Data Min Knowl Discov*. 2010;20(3):361-387. doi:10.1007/s10618-009-0152-3

14. Sauzet O, Carvajal A, Escudero A, Molokhia M, Cornelius VR. Illustration of the weibull shape parameter signal detection tool using electronic healthcare record data. *Drug Saf*. 2013;36(10):995-1006. doi:10.1007/s40264-013-0061-7
15. Arnaud M, Bégaud B, Thiessard F, et al. An automated system combining safety signal detection and prioritization from healthcare databases: a pilot study. *Drug Saf*. 2018;41(4):377-387. doi:10.1007/s40264-017-0618-y
16. Wahab A, Pratt NL, Kalisch LM, Roughead EE. Comparing time to adverse drug reaction signals in a spontaneous reporting database and a claims database: a case study of rofecoxib-induced myocardial infarction and rosiglitazone-induced heart failure signals in Australia. *Drug Saf*. 2014;37(1):53-64. doi:10.1007/s40264-013-0124-9
17. Zhou X, Douglas IJ, Shen R, Bate A. Signal detection for recently approved products: adapting and evaluating self-controlled case series method using a US claims and UK electronic medical records database. *Drug Saf*. 2018;41(5):523-536. doi:10.1007/s40264-017-0626-y
18. Aakjær M, De Bruin ML, Kulahci M, Andersen M. Surveillance of antidepressant safety (SADS): active signal detection of serious medical events following SSRI and SNRI initiation using big healthcare data. *Drug Saf*. 2021;8(11):1215-1230. doi:10.1007/s40264-021-01110-x
19. Nelson JC, Ulloa-Pérez E, Bobb JF, Maro JC. Leveraging the entire cohort in drug safety monitoring: part 1 methods for sequential surveillance that use regression adjustment or weighting to control confounding in a multisite, rare event, distributed data setting. *J Clin Epidemiol*. 2019;112:77-86. doi:10.1016/j.jclinepi.2019.04.012
20. Norén GN, Hopstadius J, Bate A. Shrinkage observed-to-expected ratios for robust and transparent large-scale pattern discovery. *Stat Methods Med Res*. 2013;22(1):57-69. doi:10.1177/0962280211403604
21. Norén GN, Bergvall T, Ryan PB, Juhlin K, Schuemie MJ, Madigan D. Empirical performance of the calibrated self-controlled cohort analysis within temporal pattern discovery: lessons for developing a risk identification and analysis system. *Drug Saf*. 2013;36(Suppl 1):S107-S121. doi:10.1007/s40264-013-0095-x
22. Mayer F, Kirchmayer U, Coletta P, 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). doi:10.1161/JAHA.117.008034
23. Schneeweiss S, Gopalakrishnan C, Bartels DB, et al. Sequential monitoring of the comparative effectiveness and safety of dabigatran in routine care. *Circ Cardiovasc Qual Outcomes*. 2019;12(2):e005173. doi:10.1161/CIRCOUTCOMES.118.005173
24. Toh S, Reichman ME, Graham DJ, 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. doi:10.2337/dc17-0476
25. Patorno E, Gopalakrishnan C, Brodovicz KG, 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):1824-1836. doi:10.1111/dom.13735
26. 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):12-16. doi:10.1177/1403494811399956
27. European Medicines Agency. *Signal management. European medicines agency*. 2018. Accessed September 23, 2020. https://www.ema.europa.eu/en/human-regulatory/post-authorisation/pharmacovigilance/signal-management
28. ICPE. All access conference abstracts. *Pharmacoepidemiol Drug Saf*. 2020;29(S3):3-634. doi:10.1002/pds.5114
29. Aakjær M, Kristiansen SB, Pape K, et al. Investigation of the potential association between the use of fluoxetine and occurrence of acute pancreatitis: a Danish register-based cohort study. *Int J Epidemiol*. 2022;dyac071. doi:10.1093/ije/dyac071
30. Lancashire RJ, Cheng K, Langman MJS. Discrepancies between population-based data and adverse reaction reports in assessing drugs as causes of acute pancreatitis. *Aliment Pharmacol Ther*. 2003;17(7):887-893. doi:10.1046/j.1365-2036.2003.01485.x
31. European Medicines Agency. *Pharmacovigilance working party (PhVWP)*. October 2011 plenary meeting. 2011.
32. Center for Drug Evaluation and Research. *FDA drug safety communication: revised recommendations for Celexa (citalopram hydrobromide) related to a potential risk of abnormal heart rhythms with high doses*. FDA; 2019. Accessed January 27, 2022. https://www.fda.gov/drugs/drug-safety-and-availability/fda-drug-safety-communication-revised-recommendations-celexa-citalopram-hydrobromide-related
33. European Medicines Agency. *Pharmacovigilance working party (PhVWP). November 2011 plenary meeting*. 2011.
34. Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. *Jama*. 2022;328(24):2446-2447. doi:10.1001/jama.2022.21383
35. Schmidt M, Sørensen HT, Pedersen L. Diclofenac use and cardiovascular risks: series of nationwide cohort studies. *BMJ*. 2018;362:k3426. doi:10.1136/bmj.k3426
36. Rossides M, Kullberg S, Di Giuseppe D, et al. Infection risk in sarcoidosis patients treated with methotrexate compared to azathioprine: a retrospective "target trial" emulated with Swedish real-world data. *Respirol Carlton Vic*. 2021;26(5):452-460. doi:10.1111/resp.14001
37. Caniglia EC, Rojas-Saunero LP, Hilal S, et al. Emulating a target trial of statin use and risk of dementia using cohort data. *Neurology*. 2020;95(10):e1322-e1332. doi:10.1212/WNL.0000000000010433
38. Dickerman BA, Gerlovin H, Madenci AL, et al. Comparative effectiveness of BNT162b2 and mRNA-1273 vaccines in U.S. veterans. *N Engl J Med*. 2022;386(2):105-115. doi:10.1056/NEJMoa2115463
39. Ray WA, Chung CP, Murray KT, Hall K, Stein CM. High-dose citalopram and escitalopram and the risk of out-of-hospital death. *J Clin Psychiatry*. 2017;78(2):190-195. doi:10.4088/JCP.15m10324
40. Coupland C, Hill T, Morriss R, Moore M, Arthur A, Hippisley-Cox J. Antidepressant use and risk of cardiovascular outcomes in people aged 20 to 64: cohort study using primary care database. *BMJ*. 2016;352:i1350. doi:10.1136/bmj.i1350

41. Zivin K, Pfeiffer PN, Bohnert ASB, et al. Evaluation of the FDA warning against prescribing citalopram at doses exceeding 40 mg. *Am J Psychiatry*. 2013;170(6):642-650. doi:10.1176/appi.ajp.2013.12030408
42. Leonard CE, Bilker WB, Newcomb C, Kimmel SE, Hennessy S. Antidepressants and the risk of sudden cardiac death and ventricular arrhythmia. *Pharmacoepidemiol Drug Saf*. 2011;20(9):903-913. doi:10.1002/pds.2181
43. Lin YT, Lu TS, Hansen RA, Wang CC. Selective serotonin reuptake inhibitor use and risk of arrhythmia: a Nationwide, population-based cohort study. *Clin Ther*. 2019;41(6):1128-1138.e8. doi:10.1016/j.clinthera.2019.04.023
44. Aakjær M, Werther SK, De Bruin ML, Andersen M. Serious arrhythmia in initiators of citalopram, escitalopram, and other selective serotonin reuptake inhibitors: a population-based cohort study in older adults. *Clin Transl Sci*. 2022;15(9):2105-2115. doi:10.1111/cts.13319
45. Sharon R, Lange T, Aakjær M, Brøgger Kristiansen S, Baltzer Houlind M, Andersen M. Incidence of hospital contacts with acute kidney injury after initiation of second-generation antipsychotics in older adults: a Danish population-based cohort study. *Eur J Clin Pharmacol*. 2022;78(8):1341-1349. doi:10.1007/s00228-022-03339-6
46. Demner-Fushman D, Shooshan S, Rodriguez L, et al. A dataset of 200 structured product labels annotated for adverse drug reactions. *Sci Data*. 2018;5(1):180001. doi:10.1038/sdata.2018.1
47. Polak TB, van Rosmalen J, Uyl-de Groot CA. Expanded access as a source of real-world data: an overview of FDA and EMA approvals. *Br J Clin Pharmacol*. 2020;86(9):1819-1826. doi:10.1111/bcp.14284
48. EMA. *Electronic product information for human medicines in the European Union – key principles*. European Medicines Agency; 2019. Accessed May 12, 2023. https://www.ema.europa.eu/en/electronic-product-information-human-medicines-european-union-key-principles
49. Cederholm S, Hill G, Asiimwe A, et al. Structured assessment for prospective identification of safety signals in electronic medical records: evaluation in the health improvement network. *Drug Saf*. 2015;38(1):87-100. doi:10.1007/s40264-014-0251-y
50. Sessa M, Khan AR, Liang D, Andersen M, Kulahci M. Artificial intelligence in Pharmacoepidemiology: a systematic review. Part 1—overview of knowledge discovery techniques in artificial intelligence. *Front Pharmacol*. 2020;11:1028. doi:10.3389/fphar.2020.01028
51. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology*. 2009;20(4):512-522. doi:10.1097/EDE.0b013e3181a663cc
52. Ryan PB, Madigan D, Stang PE, Overhage JM, Racoosin JA, Hartzema AG. Empirical assessment of methods for risk identification in healthcare data: results from the experiments of the observational medical outcomes partnership. *Stat Med*. 2012;31(30):4401-4415. doi:10.1002/sim.5620
53. Ryan PB, Stang PE, Overhage JM, et al. A comparison of the empirical performance of methods for a risk identification system. *Drug Saf*. 2013;36(Suppl 1):S143-S158. doi:10.1007/s40264-013-0108-9
54. Potts J, Genov G, Segec A, Raine J, Straus S, Arlett P. Improving the safety of medicines in the European Union: from signals to action. *Clin Pharmacol Ther*. 2020;107(3):521-529. doi:10.1002/cpt.1678

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Aakjær M, De Bruin ML, Andersen M. Epidemiological surveillance of drug safety using cumulative sequential analysis in electronic healthcare data. *Basic Clin Pharmacol Toxicol*. 2024;134(1):129-140. doi:10.1111/bcpt.13955