Combining big data search analytics and the FDA Adverse Event Reporting System database to detect a potential safety signal of mirtazapine abuse

Dimitrios Spachos, Spyridon Siafis, Panagiotis Bamidis, Dimitrios Kouvelas and Georgios Papazisis
Aristotle University of Thessaloniki, Greece

Abstract
This study sought to detect a potential safety signal of mirtazapine abuse by combining two different sources of surveillance, specifically Google Analytics (Google, Inc., Mountain View, CA, USA) and the FDA Adverse Event Reporting System database. Data from the first quarter of 2004 to the second quarter of 2017 were collected and analysed. The search interest over time, the frequencies of abuse-related terms in the search analytics domain, and the odds ratio of abuse events in FDA Adverse Event Reporting System were determined. Correlations between the two aforementioned domains using quarterly data from the timeline series were also assessed. Our results suggest a positive correlation between abuse-related searches in the Google domain and abuse-related events in FDA Adverse Event Reporting System database. These results indicate that these methods can be used in combination with each other as a pharmacovigilance supplementary tool to detect drug safety signals.

Keywords
big data, FAERS database, Google search analytics, mirtazapine abuse, safety signal

Introduction
Post-marketing drug safety surveillance mainly relies on data retrieved from spontaneous report systems and observational databases. However, due to some important limitations of these data...
sources, interest in exploring the use of new methods of surveillance to detect potential drug safety signals is growing. Prescription drug abuse is an emerging field in pharmacovigilance characterised already as an epidemic crisis in the United States, with Europe approaching a similar conclusion. The most commonly abused drugs include opioids, stimulants, benzodiazepines, and to a lesser extent, substances such as the gabapentinoids (pregabalin and gabapentin). Antidepressants are considered to have a low abuse potential and their usage has been reported mostly in patients with comorbid substance use disorder seeking euphoric and/or stimulant-like effects. Mirtazapine is a noradrenergic and specific serotoninergic antidepressant (NaSSA) marketed since 1996. PsychonautWiki, a community-driven encyclopaedia covering psychoactive substances, has classified mirtazapine as a recreational substance whose hallucinogenic, sedative, and delirium-like effects make the users rapidly develop a tolerance to but without showing habit-forming actions. Studies conducted among heroin users or inpatients treated with mirtazapine for methamphetamine withdrawal similarly report these properties. Recently, published studies have further validated this classification: mirtazapine was the most prevalent nonprescribed psychoactive drug found in the post-mortem toxicological analysis in drug poisoning deaths in Finland and was also detected in urine samples from patients with suspected illicit drug ingestion in the United Kingdom.

Considering the abovementioned data and the World Health Organization definition that ‘a signal is essentially a hypothesis of a risk with a medicine with data and arguments that support it, derived from data from one or more of many possible sources’, the aim of our study was to detect a potential safety signal of mirtazapine abuse by combining two different methods of surveillance. The first was a recently introduced technique in the pharmacovigilance domain that is based on finding patterns in big data search analytics results. During the H1N1 pandemic, Google (Mountain View, CA, USA) developed a model that could track the spread of the influenza-like illness throughout the 2008–2009 influenza season in the United States. This model was based on the idea that people in areas infected by the flu virus were conducting online searches for related information before consulting a physician. As a result, Google detected the spread of the disease in near-real-time, while the Centers for Disease Control and Prevention predicted it with a 2-week lag—a long delay for a pandemic disease. Our second method was the disproportionality analysis of the US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS) using OpenVigil2.1-MedDRA, a freely accessible pharmacovigilance data extraction, mining, and analysis tool of the FAERS database. Disproportionality analysis is a data mining algorithm of pharmacovigilance that is able to detect the disproportionality of a specific group of reports among the total gathered reports. Both methods were used to detect mirtazapine’s abuse potential in comparison with two controls, clonazepam and escitalopram, and, further, to investigate the correlation between search analytics and the FAERS domain. To our knowledge, this is the first study conducted using these two methodological approaches to detect a previously nonreported signal (i.e. mirtazapine’s abuse potential).

**Methods**

**Data sources**

The aim of employing both methods was to spot evidence of mirtazapine’s abuse liability by collecting and analysing data from the United States between 2004 and the second quarter of 2017 (2017Q2). Clonazepam, a benzodiazepine with a well-known abuse potential, served as the positive control, while escitalopram, considered an selective serotonin reuptake inhibitor (SSRI) with a low abuse potential, served as the negative control. Google Analytics was used to measure the frequency of abuse-related searches conducted in the Google search engine linked to each drug.
Disproportionality analysis was designed to detect the disproportionality of abuse-related events in the FAERS database linked to each drug. Data extraction from both domains required terms assigned to the examined drugs and terms related to drug abuse.

**Google Analytics.** The Google search engine receives billions of queries per day and some large datasets are publicly accessible, providing valuable data. Users search Google for information related to drugs using both generic and brand names, combined together in most cases with some additional terms. Both the generic name and a common brand name of each drug were considered separately for data extraction, with the resulting data subsequently combined (Table 1). For each term, Google offered us two options to search: the first employed search semantics (from Google’s internal search machine) and recognised the term as a prescription drug, while the second performed the search for the drug as a general search term. The prescription drug category generally provides accurate results, but, depending on the popularity of the term, it is not always possible to secure a large enough search data volume. As a result, the extraction of data was done using both the ‘prescription drug’ category and the ‘general search terms’ category.

Abuse-related search terms were utilised to trigger abuse-related search queries in Google. Eight terms similar to the MedDRA abuse-related terms were used in Google Analytics, as follows: ‘abuse’, ‘overdose’, ‘misuse’, ‘withdrawal’, ‘addiction’, ‘dependence’, ‘tolerance’, and ‘high’ (Table 1). Search terms with low results volumes, duplicate searches, and searches with special characters (i.e. queries with apostrophes and other special characters) were excluded. Among the billions of search queries, only those containing combinations of the selected drugs’ names and abuse-related terms were identified. However, not all queries that contained a combination of the above terms were associated with abuse. For example, the search query ‘can you get high off clonazepam’ is related to abuse, while the query ‘clonazepam and high blood pressure’ is not. In consideration of this, the isolation of the queries’ data was performed manually.

**FAERS.** FAERS is a spontaneous adverse events database that relies on individual safety reports collected by the pharmaceutical industry, health professionals, and patients hailing predominantly from the United States. OpenVigil2.1-MedDRA, a freely available pharmacovigilance tool,
operates only on cleaned FAERS data – that is, results where duplicates or reports with missing data are removed.\textsuperscript{13} It requires only the generic drug name for data extraction, as drug entries are normalised to the format of generic name (US adopted name) based on external databases (e.g. DrugBank and Drugs@FDA).\textsuperscript{13} The MedDRA terminology is incorporated into OpenVigil2.1-MedDRA for hierarchical ontological mapping of adverse events.\textsuperscript{17} High-level MedDRA terms and groups of events were selected for use in the disproportionality analysis.\textsuperscript{16} The narrow scope of the Standardised MedDRA Query (SMQ) ‘drug abuse and dependence’ was used, which is more specific than the broader scope SMQ.\textsuperscript{18} Complementary events related to drug abuse were also examined, including drug withdrawal, overdose, drug tolerance, and euphoria-related terms (Table 1).\textsuperscript{19,20} In addition to the aggregated data for the period of 2004–2017Q2, the quarterly number of abuse-related events for each drug was extracted.

**Statistical analysis**

In order to provide quantitative metrics linked to abuse liability, the following indicators were used: search interest over time, measured by the search popularity score (SPS); search frequencies of abuse-related terms in the Google Analytics domain; and the reporting odds ratio (ROR) for abuse-related adverse events in the FAERS domain.

**Search interest over time.** For each search term, Google returns the top related queries, suggesting that users looking for the term have also searched for these queries. The most popular search queries are sorted by a metric, the SPS, with a value that varies from 0 to 100 points. The SPS represents the total number of searches for each term divided by the total number of related searches of the geography and time range in order to compare relative popularity between topics. The most-searched-for query takes a score of 100 points, while a score of 50 points is assigned to queries that are searched half as often. A score of 0 points is assigned to those queries searched for at a rate of less than 1 per cent of that of the activity focused on the top query. In other words, terms with a popularity of less than 1 per cent in comparison with the peak have a score of 0 points.\textsuperscript{15}

The monthly SPS of overall abuse-related terms was also obtained for each drug. Thus, timelines for each drug were created, which represent the average search interest over time for abuse-related terms between the start of 2004 and 2017Q2. Search analytics of clonazepam, mirtazapine, and escitalopram were initially evaluated with regard to the normality assumption using a formal test (Kolmogorov–Smirnov test), histograms, boxplots, and normal Q–Q plots. Kurtosis and skewness parameters were also considered. Normality assumption findings were considered in order to calculate the appropriate descriptive statistics and choose the proper statistical test for comparing the search analytics between drugs. As search analytics were not approximately normally distributed, medians and interquartile ranges (IQRs) were calculated as measures of centrality and variability, respectively, and between-drug comparisons were performed using the Kruskal–Wallis $H$ test. In the case of statistical significance with the Kruskal–Wallis $H$ test, we performed post hoc tests using the Mann–Whitney $U$ test for between-group comparisons.

**Search frequencies of search queries linked to abuse-related terms.** The exact search frequencies for abuse-related terms for each drug were gathered and analysed according to the following method. Using Google’s AdWords Keyword Planner, additional sets of search queries for each abuse-related term were created, containing not only exact matches but also phrases that users had typed in the search engine. For example, for the terms ‘clonazepam’ and ‘addiction’, we received a set of six related search queries linked with this abuse-related term (Table 2). The search volume of the query represents the average monthly searches made for each query. By summing the search
volumes for abuse-related queries, as seen in Table 2, and dividing by the search volume of the queries related to clonazepam, we can obtain the search frequency for the pairing of this drug with a certain abuse-related search term. In the above example, the drug is clonazepam and its average search volume per month was 246,000. We repeated this procedure for all pairings. Our analyses ultimately contained more than 200 different search queries and their respective volume sizes, including abuse-related search terms.

The frequentist method, ROR, was used to calculate signals in the FAERS database. ROR is a disproportionality measure that estimates the frequency of a selected adverse event in association with the tested drug in comparison with all other drugs in the database. A higher ROR result signifies a more frequent reporting of the selected combination of the event and drug in question and indicates stronger signals. Based on previous studies, safety signals were detected when the number of drug–event reports was greater than three and the lower boundary of the ROR was greater than one.16,21 RORs and 95 per cent confidence intervals (CIs) were calculated in OpenVigil2.1-MedDRA, which was validated by performing the above statistics.13

### Correlation between FAERS and search analytics domains.

Search analytics (FAERS and Google Analytics) were computed for the three drugs considered in this research. Our purpose was to investigate whether a significant cross-correlation could be found between time series. The most straightforward measure of similarity between time series of the same length \((n)\) is the cross-correlation \(r_{XY}(\tau)\) as a function of the time (lag \(\tau\) of a 3-month period)

\[
r_{XY}(\tau) = \frac{c_{XY}(\tau)}{s_X s_Y}, \quad c_{XY}(\tau) = \frac{1}{n} \sum_{t=1}^{n-\tau} (x_t y_{t+\tau} - \bar{x} \bar{y})
\]

In equation (1), \(c_{XY}(\tau)\) denotes the cross-covariance at a given lag \(\tau\), and \(s_X\) and \(s_Y\) are the sample standard deviation and mean of \(X\), respectively (represented similarly for \(Y\)). In more detail, for a zero lag, equation (1) corresponds to the Pearson correlation coefficient of \(X\) and \(Y\), \(r_{XY}(0) = s_{XY} / s_X s_Y = r(X, Y)\). Although uncorrelated time series have a nonsignificant cross-correlation varying around zero, we thought it prudent to perform a significance test for cross-correlation in order to infer whether a genuine correlation between time series existed or not. Thus, we explored

Table 2. Example of search queries for clonazepam and an abuse-related search term (addiction) with their search volume.

| Search query                                      | Search volume | Search volume (%) |
|---------------------------------------------------|---------------|-------------------|
| 1 Is clonazepam addictive?                        | 260           | 0.106             |
| 2 Clonazepam addiction                            | 260           | 0.106             |
| 3 Addiction to clonazepam                         | 70            | 0.028             |
| 4 Clonazepam addiction withdrawal                 | 50            | 0.020             |
| 5 Clonazepam addictive                            | 40            | 0.016             |
| 6 Addicted to clonazepam                          | 10            | 0.004             |
| Summary                                           | 690           | 0.280             |

Additional sets of search queries for each abuse-related term were created, containing not only the exact matches but also phrases that users type in the search engine. For the terms ‘clonazepam’ and ‘addiction’, a set of six related search queries linked with the abuse-related term was retrieved. The search volume of the query is a relative number that represents the average monthly searches for each query, comparing to the search volume all drug searches. In this example, the average search volume per month for clonazepam was 246,000.
the cross-correlation of time series in the context of each drug, taking advantage of the basic scheme of a randomisation significance test for cross-correlation.\textsuperscript{22,23} In more detail, we initially plotted the time series and then computed their cross-correlation for a lag value equal to 5 (including zero lag, which corresponds to the Pearson correlation coefficient). The null hypothesis was tested using the methodology presented by Pandria and Kugiumtzis.\textsuperscript{23}

Subsequently, a resampling approach (random permutation) was used in time series and surrogates of the time series were produced ($M=1000$). Following this approach, we formed the null hypothesis that we sought to test [$H_0: \rho(X, Y)=0$] by calculating it in randomised pairs of time series that preserve the marginal distribution of the original time series but were, by construction, uncorrelated. More precisely, random permutation destroyed any time-matching between time series and, as such, they were uncorrelated. The level of statistical significance was set at 5 per cent.

**Results**

**Google Analytics**

Analysis for the cumulative period revealed that overall abuse-related terms had an average SPS of 56.0 for clonazepam, 24.4 for mirtazapine, and 11.4 for escitalopram. Again, larger values of SPS indicate more queries were made. Considering that Google receives more than 3.5 billion queries per day,\textsuperscript{21} even low values of SPS suggest that millions of search queries were posed about a topic.

Similar patterns were also found in the search interest over time (Figure 1(a)). The descriptive analyses showed that the median values of search analytics over time were 65.50, IQR=[34.17, 76.08] for clonazepam; 20.67, IQR=[14.50, 38.25] for mirtazapine; and 8.67, IQR=[5.92, 17.83] for escitalopram. Based on the Kruskal–Wallis $H$ test, it was found that search analytics between the three drugs were significantly different ($\chi^2(2) = 100.558; p<0.001$). By performing a post hoc test, we concluded that the search analytics results for clonazepam were significantly higher than those of mirtazapine ($U=386.00; p<0.001$) and escitalopram ($U=53.00; p<0.001$), respectively. Similarly, those for mirtazapine were significantly greater than those of escitalopram ($U=540.50; p<0.001$).

The frequencies of search queries linked to abuse-related terms are displayed in Figure 2 and Table 3. Queries linked to ‘high’ seem to be those that were most commonly posed for each drug, along with those related to ‘overdose’ and ‘withdrawal’. The frequency of search queries linked to ‘high’ was 5.78 per cent for clonazepam, 5.04 per cent for mirtazapine, and 2.33 per cent for escitalopram. Similar patterns appeared to be followed by the other search queries. Queries containing the terms ‘misuse’, ‘dependence’, and ‘tolerance’ were not as frequently used.

**FAERS disproportionality analysis**

After data cleaning including duplicate deletion, the study population consisted of 4,704,663 reports. Of these, 15,652 involved mirtazapine, 36,026 involved clonazepam, and 26,473 involved escitalopram. The percentage of co-occurring drug abuse/dependence with mirtazapine (5.62%) was comparable to that associated with clonazepam (5.95%) but greater than that associated with escitalopram (3.18%) (Table 4). Safety signals for events related to drug abuse could be detected for mirtazapine and both control drugs, though they differed in magnitude from one another. Drug abuse/dependence, withdrawal, and tolerance seem to be associated more with clonazepam and mirtazapine and to a lesser degree with escitalopram, respectively. A stronger signal for overdose was generated by mirtazapine, in contrast to clonazepam and escitalopram, while euphoria-related
terms appeared to be associated more with clonazepam and escitalopram and to a lesser degree with mirtazapine (Table 4).

Correlation between FAERS and search analytics domains

The cross-correlation of time series (Figure 1(a) and (b)) was calculated for a lag value equal to 5 as presented in Figure 3. A lag refers to a difference in time between two observations. As displayed in Figure 3, the search analytics of clonazepam were found to be uncorrelated for each of the lags, while the search analytics of escitalopram were uncorrelated for each lag, as the cross-correlation of time series was between the approximate upper and lower confidence bounds.

Figure 1. (a) Search interest over time for abuse-related terms in search analytics domain and (b) number of abuse-related events submitted in FAERS. In the search analytics domain, for the period 2004Q1–2017Q2, the search interest over time for overall abuse-related terms is represented in timelines for each drug and is expressed as quarterly relative search volume for overall abused-linked terms. During the same period, the quarterly number of drug-abuse-related events in FAERS for each drug is presented. The narrow scope of the SMQ ‘Drug abuse and dependence’ was used to identify abuse-related events.
Moreover, search analytics for most of the lags were uncorrelated, whereas the opposite was true for the lags equal to −3 and −5. From Figure 3, we can additionally export the zero-lag cross-correlation of time series in each drug. Thus, the zero-lag cross-correlation of search analytics outcomes were 0.37, 0.38, and 0.05 for clonazepam, mirtazapine, and escitalopram, respectively. Our hypothesis regarding the resampling method used was confirmed, as the cross-correlation between the randomised pairs of time series was between the approximate lower and upper bounds of zero correlation, respectively (Figure 4). Based on the results of significance testing for cross-correlation, search statistics for clonazepam ($p=0.0007$) and mirtazapine ($p=0.004$) were found to be significantly correlated between the search analytics and FAERS domains, while the corresponding data for escitalopram did not show such results ($p=0.34$) (Figure 5).

**Discussion**

To our knowledge, this is the first study to investigate the abuse potential of mirtazapine using two different pharmacovigilance methods: disproportionality analysis in the FAERS and Google Analytics. A positive control (benzodiazepine clonazepam with a well-known abuse profile) and a negative control (SSRI escitalopram with a previously unreported abuse potential) were used for comparison purposes.

Due to their volume, velocity, and variety, Google Analytics data are considered to be big data and, while the recognition of the potential of big data analytics is rising, its use in pharmacovigilance is
still emerging. A recently published study of the French Addictovigilance Network combined Google Trends with the analysis of the global database of individual case safety reports (VigiBase). In our study, Google's big data analysis revealed that a significant number of US residents searched for terms related to the abuse liability of mirtazapine. In the time frame of our study, these are hundreds of millions of queries being posed worldwide, providing a potential signal of drug abuse. The identified signal of mirtazapine abuse was in agreement with the signals for the positive and negative control drugs (i.e. clonazepam and escitalopram, respectively). The generic pattern was

\[
\text{signal indicator (clonazepam)} \geq \text{signal indicator (mirtazapine)} \geq \text{signal indicator (escitalopram)},
\]

where 'indicator' refers to the SPS, search interest over time, and search frequency of abuse-related terms. A further finding of our study was that an analysis over time revealed the same patterns regarding the search interest, suggesting that this behaviour is not a temporary one; rather, it seems to have been consistent over the past 13 years.

Importantly, the signal patterns in the search analytics domain were similar to those found in the FAERS database domain, in which signals of drug abuse-related events have been detected for clonazepam, mirtazapine, and, surprisingly, escitalopram. Mirtazapine’s signals were in between clonazepam’s and escitalopram’s, except for in the context of ‘overdose’ and ‘euphoria’. However, this exception seems not to be critical since both an ALERTT Working Group (ACTTION) systematic review and the FDA have suggested that euphoria- and overdose-related terms are non-specific and should not be interpreted alone in the assessment of abuse potential. Escitalopram showed the lowest abuse-related reporting rates in comparison with the other two drugs, both in the search analytics domain and the FAERS database; however, the rates were indicative of a signal of abuse.

### Table 4. Number of reports and ROR (95% CI) related to drug-abuse per drug.

| Event and drug        | Reports | Total drug (%) | ROR (95% CI) |
|-----------------------|---------|----------------|--------------|
| **Drug abuse/dependence** |         |                |              |
| Clonazepam            | 2144    | 5.95           | 4.27 (4.08–4.46) |
| Mirtazapine           | 880     | 5.62           | 3.96 (3.70–4.24) |
| Escitalopram          | 842     | 3.18           | 2.18 (2.03–2.33) |
| **Drug withdrawal**   |         |                |              |
| Clonazepam            | 851     | 2.36           | 3.99 (3.72–4.27) |
| Mirtazapine           | 268     | 1.71           | 2.83 (2.5–3.19)  |
| Escitalopram          | 398     | 1.5            | 2.48 (2.25–2.74) |
| **Overdose**          | 29,002  |                |              |
| Clonazepam            | 1825    | 5.07           | 4.52 (4.31–4.74) |
| Mirtazapine           | 1012    | 6.47           | 5.8 (5.44–6.18)  |
| Escitalopram          | 1069    | 4.04           | 3.52 (3.31–3.75) |
| **Drug tolerance**    |         |                |              |
| Clonazepam            | 79      | 0.22           | 7.9 (6.3–9.92)  |
| Mirtazapine           | 12      | 0.08           | 2.63 (1.49–4.65) |
| Escitalopram          | 10      | 0.04           | 1.29 (0.69–2.41) |
| **Euphoria-related terms** | 196,087 |                |              |
| Clonazepam            | 3395    | 9.42           | 2.42 (2.33–2.50) |
| Mirtazapine           | 1130    | 7.22           | 1.79 (1.69–1.91) |
| Escitalopram          | 2222    | 8.39           | 2.12 (2.03–2.21) |

ROR: reporting odds ratio; CI: confidence interval; FAERS: FDA Adverse Event Reporting System.

Each drug has been compared with all other drugs in the FAERS database. The study population consisted of 4,704,663 reports. Of them, 15,652 included mirtazapine, 36,026 clonazepam, and 26,473 escitalopram.
Figure 3. Cross-correlation of time series for (a) clonazepam, (b) mirtazapine, and (c) escitalopram, respectively. The two lines (cyan colour) indicate the approximate upper and lower confidence bounds assuming the input series are completely uncorrelated.
Figure 4. Cross-correlation of randomised pairs of time series for (a) clonazepam, (b) mirtazapine, and (c) escitalopram, respectively. Element vectors that are displayed as cyan lines are the approximate upper and lower confidence bounds of zero correlation.
Figure 5. Results of randomisation significance test for cross-correlation of search analytics in (a) donepezil, (b) mirtazapine, and (c) escitalopram, respectively.
To confirm escitalopram’s signal of drug abuse-related events, further clinical assessment of the initially submitted reports is needed. In addition, factors typically not included in a disproportionality analysis such as concomitant medications, co-suspected drugs, and comorbidity should be evaluated for standardised case causality assessment.

To date, limited evidence supports the understanding of the abuse properties of mirtazapine. Opioid users on antidepressant treatment seem to be aware of the possible use of mirtazapine for achieving intoxicating effects. According to this, patients being for methamphetamine withdrawal using mirtazapine reported significantly greater (yet still limited) feelings of a ‘high’ or ‘rush’ in comparison with modafinil-treated patients. In addition, a recent toxicological screening of urine samples from patients with suspected illicit drug ingestion showed that almost half of the patients used multiple drugs, with mirtazapine detected in 11.25 per cent of them. Mirtazapine seems to increase the dopaminergic efflux selectively in the cortex rather than in the nucleus accumbens, differentiating it from common drugs of abuse. Pregabalin and gabapentin may also fail to increase the dopaminergic efflux in this region, yet the evidence of their abuse potential is growing.

Mirtazapine has a unique pharmacodynamic profile, with $H_1$, 5-HT2A, 5-HT2C, 5-HT3, and alpha2 adrenergic antagonistic effects; lower affinities to muscarinic, alpha1, and dopaminergic receptors; and negligible effects on monoamine reuptake. Adrenergic alpha2, $H_4$, and muscarinic receptor antagonism could promote abuse-related behaviours. Despite mirtazapine’s low affinity to muscarinic receptors, muscarinic receptor antagonism could be achieved by high doses of mirtazapine, which could contribute to the suggested abuse properties.

Finally, there was a positive correlation between search interest in Google and incidents reported to the FAERS database, which presented a monotonic correlation for mirtazapine and clonazepam but not for escitalopram. At its core, a statistically significant correlation quantifies the relationship between two data domains. When one of the data values changes in one domain, the other is highly likely to change as well in response in the other domain. Thus, changes in abuse-related searches on Google for mirtazapine or clonazepam are expected to be accompanied by analogous changes of abuse-related events in FAERS. This outcome does not indicate causality but, rather, a similarity in the behaviours of two data domains. Interestingly, in opposition to clonazepam and mirtazapine, no significant correlation between the two domains was noticed for escitalopram. Concerns about the abuse liability of SSRI antidepressants might explain the significant number of Internet searches performed by users and, thus, the existence of a Google signal that is not reflected in a spontaneous report to the FDA given that SSRIIs are rarely misused and seem to have a low abuse liability. On the contrary, abuse-related searches for mirtazapine or clonazepam might conceal a motivation to abuse, since searches and abuse-related events are correlated.

This study has some methodological considerations and limitations that should be highlighted. First, while disproportionality analysis may be a suitable tool to quantitate signals of abuse of known and novel psychoactive substances, it has a limited capacity to differentiate the type of abuse, for example, recreational use or pharmaceutical neuroenhancement. In addition, confounding factors such as comorbidity and concomitant drugs cannot be assessed properly. Finally, the causal relationship between drugs and the adverse event (abuse) cannot be verified without a clinical causality assessment. Regarding search analytics, Google only reports large datasets. In our study, terms like dependence, tolerance, and misuse did not provide substantial numbers for the analysis of the search frequency in contrast to other terms such as high or overdose. For this reason, the evaluation of the correlation between the two domains was conducted with overall abuse-related terms in Google Analytics but only using the more precise narrow scope of the SMQ ‘drug abuse and dependence’ category in FAERS. In addition, the algorithms and their updates utilised by Google to analyse data are not publicly available. Finally, data before 2004 were excluded since both Google and OpenVigi2.1-MedDRA operate on data from 2004.
Conclusion

In summary, our results achieved using two different methods strengthen the evidence towards a potential signal of mirtazapine abuse, confirming the findings from post-marketing reports. Further signal validation requires clinical assessment in order to confirm causality and is a step in the signal management process performed by the regulatory authorities. However, based on the positive correlation found between search analytics and incidents reported to the FAERS in this study and in line with other recent studies that combined web sources with more conventional pharmacovigilance data, we suggest that similar methods could be invoked as supplementary pharmacovigilance tools to detect drug safety signals.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Spyridon Siafis https://orcid.org/0000-0001-8264-2039
Panagiotis Bamidis https://orcid.org/0000-0002-9936-5805
Georgios Papazisis https://orcid.org/0000-0003-1641-9095

References

1. Hernandez SH and Nelson LS. Prescription drug abuse: insight into the epidemic. Clin Pharmacol Ther 2010; 88(3): 307–317.
2. Papazisis G and Tzachanis D. Pregabalin’s abuse potential: a mini review focusing on the pharmacological profile. Int J Clin Pharmacol Ther 2014; 52(8): 709–716.
3. Evans EA and Sullivan MA. Abuse and misuse of antidepressants. Subst Abuse Rehabil 2014; 5: 107–120.
4. Haukka J, Kriikku P, Mariottini C, et al. Non-medical use of psychoactive prescription drugs is associated with fatal poisoning. Addiction 2018; 113(3): 464–472.
5. Davis R and Wilde MI. Mirtazapine: a review of its pharmacology and therapeutic potential in the management of major depression. CNS Drugs 1996; 5(5): 389–402.
6. PsychonautWiki. Mirtazapine, https://psychonautwiki.org/wiki/Mirtazapine (accessed 20 October 2017).
7. Cornwall CS, Umeh K and Manshani N. Heroin users’ experiences of depression: a qualitative study. Fam Pract 2012; 29(5): 586–592.
8. McGregor C, Srisurapanont M, Mitchell A, et al. Symptoms and sleep patterns during inpatient treatment of methamphetamine withdrawal: a comparison of mirtazapine and modafinil with treatment as usual. J Subst Abuse Treat 2008; 35(3): 334–342.
9. Lowe DJ, Torrance HJ, Ireland AJ, et al. SODAS: surveillance of drugs of abuse study. Eur J Emerg Med 2017; 24: 126–129.
10. WHO-UMC. What is a signal? https://www.who-umc.org/research-scientific-development/signal-detection/what-is-a-signal/ (2017, accessed 30 November 2017).
11. Abbott R. Big data and pharmacovigilance: using health information exchanges to revolutionize drug safety. Iowa L Rev 2013; 99: 225.
12. Ginsberg J, Mohebbi MH, Patel RS, et al. Detecting influenza epidemics using search engine query data. Nature 2009; 457(7232): 1012–1014.
13. Bohm R, von Hehn L, Herdegen T, et al. OpenVigil FDA – inspection of U.S. *PLoS ONE* 2016; 11(6): e0157753.
14. Bossard JB, Ponte C, Dupouy J, et al. Disproportionality analysis for the assessment of abuse and dependence potential of pregabalin in the French pharmacovigilance database. *Clin Drug Investig* 2016; 36(9): 735–742.
15. supportgooglecom. The Homepage Explained – Trends Help, https://support.google.com/trends/answer/6248105?hl=en&ref_topic=6248052
16. Sakaeda T, Tamon A, Kadoryama K, et al. Data mining of the public version of the FDA adverse event reporting system. *Int J Med Sci* 2013; 10: 796–803.
17. Brown EG, Wood L and Wood S. The medical dictionary for regulatory activities (MedDRA). *Drug Saf* 1999; 20: 109–117.
18. Gahr M, Connemann BJ, Schonfeldt-Lecuona C, et al. Sensitivity of quantitative signal detection in regards to pharmacological neuroenhancement. *Int J Mol Sci* 2017; 18(1): E101.
19. Smith SM, Dart RC, Katz NP, et al. Classification and definition of misuse, abuse, and related events in clinical trials: ACTTION systematic review and recommendations. *Pain* 2013; 154(11): 2287–2296.
20. Food Drug Administration H. Assessment of abuse potential of drugs-guidance for industry. Docket no. FDA-2010-D-0026, Center for Drug Evaluation and Research, 2017, https://www.fda.gov/node/382407
21. Van Puijenbroek EP, Bate A, Leufkens HG, et al. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions. *Pharmacoepidemiol Drug Saf* 2002; 11(1): 3–10.
22. Kugiumtzis D. Surrogate data test on time series. In: Soofi AS and Cao L (eds) *Modelling and forecasting financial data: techniques of nonlinear dynamics*. Boston, MA: Springer, 2002, pp. 267–282.
23. Pandria N and Kugiumtzis D. Testing the correlation of time series using dynamic time warping. In: 27th Panhellenic Conference of Statistics, 2014.
24. Kruse CS, Goswamy R, Raval Y, et al. Challenges and opportunities of big data in health care: a systematic review. *JMIR Med Inform* 2016; 4(4): e38.
25. Ponte C, Pi C, Palmaro A, et al. Early signal of diverted use of tropicamide eye drops in France. *Br J Clin Pharmacol* 2017; 83(8): 1791–1800.
26. Millan MJ, Gobert A, Rivet JM, et al. Mirtazapine enhances frontocortical dopaminergic and corticolimbic adrenergic, but not serotoninergic, transmission by blockade of alpha2-adrenergic and serotonin2C receptors: a comparison with citalopram. *Eur J Neurosci* 2000; 12: 1079–1095.
27. Volkow ND, Wang G-J, Fowler JS, et al. Addiction: beyond dopamine reward circuitry. *Proc Natl Acad Sci U S A* 2011; 108(37): 15037–15042.
28. Bonnet U and Scherbaum N. How addictive are gabapentin and pregabalin? A systematic review. *Eur Neuropsychopharmacol* 2017; 27(12): 1185–1215.
29. Caplan JP, Epstein LA, Quinn DK, et al. Neuropsychiatric effects of prescription drug abuse. *Neuropsychol Rev* 2007; 17(3): 363–380.
30. Sofuoglu M and Sewell RA. Norepinephrine and stimulant addiction. *Addict Biol* 2009; 14(2): 119–129.
31. Chiappini S and Schifano F. A decade of gabapentinoid misuse: an analysis of the European medicines agency’s ‘suspected adverse drug reactions’ database. *CNS Drugs* 2016; 30(7): 647–654.
32. Rodriguez EM, Staffa JA and Graham DJ. The role of databases in drug postmarketing surveillance. *Pharmacoepidemiol Drug Saf* 2001; 10(5): 407–410.
33. Nuti SV, Wayda B, Ranasinghe I, et al. The use of Google Trends in health care research: a systematic review. *PLoS ONE* 2014; 9(10): e109583.