International Journal of Drug Policy 64 (2019) 34–39

Contents lists available at ScienceDirect

International Journal of Drug Policy

journal homepage: www.elsevier.com/locate/drugpo

Research Paper

4-Fluoramphetamine in the Netherlands: Text-mining and sentiment analysis of internet forums

Matthijs Blankers[a,b,c,*], Daan van der Gouwe[a], Margriet van Laar[a]

[a] Trimbos Institute, The Netherlands Institute of Mental Health and Addiction, Da Costakade 45, 3521VS, Utrecht, the Netherlands
[b] Department of Research, Arkin Mental Health Care, Klaprozenweg 111, 1033NN, Amsterdam, the Netherlands
[c] Amsterdam UMC, Location AMC, Department of Psychiatry, University of Amsterdam, Meibergdreef 9, 1105AZ, Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Text mining
Sentiment analysis
4-Fluoramphetamine
New psychoactive substances
Internet-based drug forums
Trend analysis

ABSTRACT

Background: Users of new psychoactive substances including 4-fluoroamphetamine (4-FA/4-FMP) frequently share their experiences or opinions in online drug forums. We have tested the potential of computerised analysis of drug users' forum posts for monitoring and early detection of trends. Specifically, we tested whether changes in the volume of 4-FA related posts and sentiments expressed in those posts can be observed around the time 4-FA was increasingly reported by Dutch drug monitoring sources (2012–2017).

Methods: Opening posts from two popular Dutch internet-based drug discussion forums, written between January 1 st, 2012 and January 1 st, 2018 were scraped: Portions of the forum posts about 4-FA were collected. To contrast 4-FA findings against other categories of forum posts, we also collected posts on two other substances (ecstasy and cocaine) and posts not related to a specific substance. Sentiments expressed in these posts were inferred using text recognition software, and analysed for trends using linear mixed modelling.

Results: The number of 4-FA posts increased between 2012 and 2015: 76 posts in 2012, 138 in 2013, 322 in 2014, 323 in 2015, and decreased thereafter: 264 in 2016 and 135 in 2017; $X^2(5) = 271.8$, $p < .001$. Over time, a decrease in positive sentiment towards 4-FA can be observed starting in 2015, compared to the period before 2015, coinciding with more news searches and reports on adverse events related to 4-FA use. Linear mixed modelling analysis confirmed a significantly higher sentiment score in 2015 compared to 2017 for 4-FA, $B = 0.062$; $SE = 0.023$; $t(1252) = 2.70$; $p = 0.007$, but not for posts on other substances.

Conclusion: Changes in the volume and sentiments of forum posts coincided with news media exposure related to 4-FA and with trends observed by established drug monitoring sources. Hence, internet forum monitoring facilitates early discovery of trends in the popularity, prevalence and adverse events related to new psychoactive substances.

Introduction

New Psychoactive Substances (NPS), substances not controlled under the United Nations conventions on drugs (1961 and/or 1971), have become a growing global phenomenon (UNODC, 2016, 2018). Over 100 countries and territories from all regions of the world have reported the presence of one or more NPS (UNODC, 2016). In the Netherlands, we recently observed an increase in popularity of the NPS 4-fluoroamphetamine (common acronyms: 4-FA and 4-FMP) alongside a rapid and substantial increase in the number of acute emergencies and reports of life-threatening complications due to 4-FA (Wijers, van Litsenburg, Hondebrink, Niesink, & Croes, 2017). 4-FA is a releasing agent and reuptake inhibitor of dopamine, serotonin, and norepinephrine (Croes et al., 2017). The earliest appearances of 4-FA on the Dutch drug market have occurred around 2009, as an adulterant in amphetamine or ecstasy (Hondebrink, Nugteren-van Lonkhuyzen, van der Gouwe, & Brunt, 2015; Linsen et al., 2015). More recently, it became popular as a recreational drug (Wijers, van Litsenburg et al., 2017), potentially because of its relatively mild psychoactive effects including euphoria, which resemble the subjective effects of amphetamines and ecstasy (van der Gouwe & Rigter, 2016; Linsen et al., 2015). Other reported clinical effects of 4-FA include agitation, tachycardia, hypertension, and hyperthermia (Wijers, van Litsenburg et al., 2017). Although nationally representative data is lacking, a study among young people (18–35) who frequently attend dance events and festivals in the Netherlands indicates that 29.2% has ever used 4-FA and 9.9% has done so in the last year (Monshouwer, Drost, van der Pol, & van Laar, 2016). Recently, the Monitor Drug-related Incidents (MDI; see also Wijers, Croes, de Ruiter, & Valkenburg, 2017) and the Dutch

* Corresponding author at: Trimbos Institute, The Netherlands Institute of Mental Health and Addiction, Da Costakade 45, 3521VS, Utrecht, the Netherlands.
E-mail address: mblankers@trimbos.nl (M. Blankers).

https://doi.org/10.1016/j.drugpo.2018.11.016

Poisons Information Centre received several reports that 4-FA intoxication has led to severe cardiovascular and cerebrovascular complications, such as intracerebral haemorrhages (Wijers, van Litsenburg et al., 2017); of four of these patients, the presence of 4-FA in the blood of the patients was chemically confirmed.

For 4-FA, as for many other NPS (van der Gouwe, Brunt, van Laar, & van der Pol, 2017), sale over the internet was the most common method to obtain the drug in the Netherlands in 2016, together with acquisition via friends; prices are reportedly around €4-6 per 4-FA dose/pill (Croes et al., 2017). Although epidemiological data are scarce, data on the number and quality of consumer 4-FA samples submitted to the Dutch Drugs Information and Monitoring System (DIMS; see also van der Gouwe & Rigter, 2016) (Hondebrink et al., 2015) and data supplied by the MDI in the Netherlands indicate that the use of 4-FA substantially increased over the last years. In a recent report, the rising popularity of 4-FA has been documented in a timeline (van der Pol, Nijkamp, Nabben, & van Laar, 2017). In 2008, 4-FA was first found in a sample submitted to DIMS. In 2012, the first 4-FA related health incident was reported by MDI. In 2014 there was a substantial increase in the number of 4-FA samples submitted to DIMS, followed by an increase in the number of health incidents reported by the MDI in 2015. In 2016, the first 4-FA related fatality was reported, followed by health warnings issued nationwide by various addiction treatment and research institutes. A risk assessment by the Dutch National Institute for Public Health and the Environment (RIVM) has led to a ban of the production, sale, and possession of 4-FA since the 25th of May 2017 (van der Pol et al., 2017). 4-FA has been placed on 'Schedule 1', which means it now has the same legal status as for instance cocaine and ecstasy in the Netherlands (Ministry of Health, Welfare & Sport, 2017). Throughout 2017, health incidents related to 4-FA have been covered by the Dutch media.

In the Netherlands, as in most other countries, NPS users increasingly discuss their positive or negative experiences and opinions on substances over the internet, for example in online drug forums (e.g. Davey et al., 2012). In these drug forums, discussions are initiated, read and commented on by others who are interested in or have used the same or similar substances. Hence, online drug forum discussions potentially contain relevant information on drug users' experiences with various substances. Tapping in to this potential source of information on drug users' experiences and opinions regarding substances is useful to identify emerging trends in drug use (Davey et al., 2012; Lamy et al., 2017; Thanki & Frederick, 2016). It has a number of advantages over traditional, survey- or interview-based research, as it for example minimises potential interference with study subjects, and data can be collected relatively quickly. Previous social media and internet studies have for example looked into structures and trends on drug cryptomarkets (Tzanetakis, 2018), evaluated positive and negative effects of synthetic cannabinoids reported on drugs forums (Lamy et al., 2017), or did trade network analyses using crawled data from drug cryptomarkets (Norbutas, 2018). Google Trends has been used to monitor changes in popularity of certain substances including steroids (Tay Wee Teck & McCann, 2018) and Google searches for evaluation of interest in synthetic cannabinoids (Curtis et al., 2015).

A challenge for internet content analysis, or more specific forum data analysis is that the amount of data can be huge and easily drain the resources of a research team, should all forum posts be read and analysed by researchers. However, developments in natural language processing software (e.g. de Smedt & Daelemans, 2012) have made it increasingly feasible to manage and analyse large corpora of unstructured text data using computer algorithms. In this study, we have used "Pattern", a web mining package for the Python programming language (de Smedt & Daelemans, 2012). Pattern is capable of performing natural language processing, machine learning, and network analysis. It identifies sentence constituents such as nouns and verbs (de Smedt & Daelemans, 2012) and is capable of indicating whether the overall sentiment of a text fragment in general is negative or positive, on a continuous scale (sentiment analysis). Pattern has been evaluated and compared against 23 other natural language processing tools by Ribeiro, Araújo, Gonçalves, Gonçalves, and Benevenuto (2016). This evaluation is based on a benchmark study involving 18 different text datasets, containing messages posted on social networks, movie and product reviews, as well as opinions and comments in news articles. Ribeiro et al. (2016) found that over these datasets, the performance of Pattern is among the top 33% of the 24 natural language processing tools in terms of accuracy (detection of the correct sentiments (3 levels: positive, neutral or negative) and coverage (any sentiment detected). The key reason for us to work with Pattern was that it is one of the only tools capable of processing the Dutch language. We have used the possibilities of Pattern to detect trends in the popularity of 4-FA and other drugs, by performing computerised content and sentiment analysis of drug users' forum post data in the period 2012–2017. In this paper we will address the following two research questions:

1 To what extent can changes in the volume and proportion of 4-FA related posts be observed around the time this substance was increasingly seen by DIMS and MDI monitoring services?
2 Can changes in the sentiment of the posts about 4-FA be observed between 2012 and 2017, and do these changes coincide with health incidents and media coverage related to 4-FA?

Methods

Two popular Dutch language internet-based drug discussion forums were scraped: opening posts, which are the first posts of a new discussion thread on the forums were collected and stored in a database for follow-up analysis (any registered forum member can start a new discussion thread). To maximally take into account the privacy of forum users, we chose not to mention the names of the selected forums in this paper, however if necessary they can be requested from the first author of this paper. To this end, we also refrained from downloading any information which could directly or indirectly be related to individual forum members. While scraping, we adhered to the forum rules of the two forums. As no data were prospectively collected for this study and there were no participants subjected to research procedures, this study is exempted from medical ethics approval.

Internet forum scraping

To scrape the forums, we made use of the R statistical programming environment (Microsoft R Open version 3.4.0), RSelenium version 2.1.1, and the PhantomJS headless browser version 3.5.3. Using this software, we were able to interact with the web forums via scripts as other users would do via a web browser. To collect relevant posts, we used the search functions of the web forums to find posts on 4-FA. We used search terms commonly used to refer to 4-FA (4(-)fa, (4)(-)fmp, (4)(-)fluor*, flux, flava) to maximise the probability that relevant posts would be found. In order to contrast specific trends in the number of posts and sentiments towards 4-FA against more general trends, we also collected posts containing terms commonly used to refer to two other frequently discussed substances: ecstasy (search terms: 'xtc', 'ecstasy', 'mdma') and cocaine (search terms: 'cocaïne', which is the Dutch spelling, 'cocaine', 'coke'). We further collected posts not related to a specific substance, by using the search term: 'forum'. We chose the word 'forum' as our non-specific 'control' search term, as this word is used in many of the posts to refer to the drug discussion forum itself (instead of to a specific substance). We only used the opening posts (the first post in a new thread) between the years 2012–2017 on the two forums containing one or more of the search terms in our analysis, but not the responses to these opening posts. The reason for this choice is threefold: (a) we observed that responses sometimes go 'off topic'; (b) responses repeat/cite parts of previous responses or the opening post, and (c) in general, responses do not reflect an independent opinion or

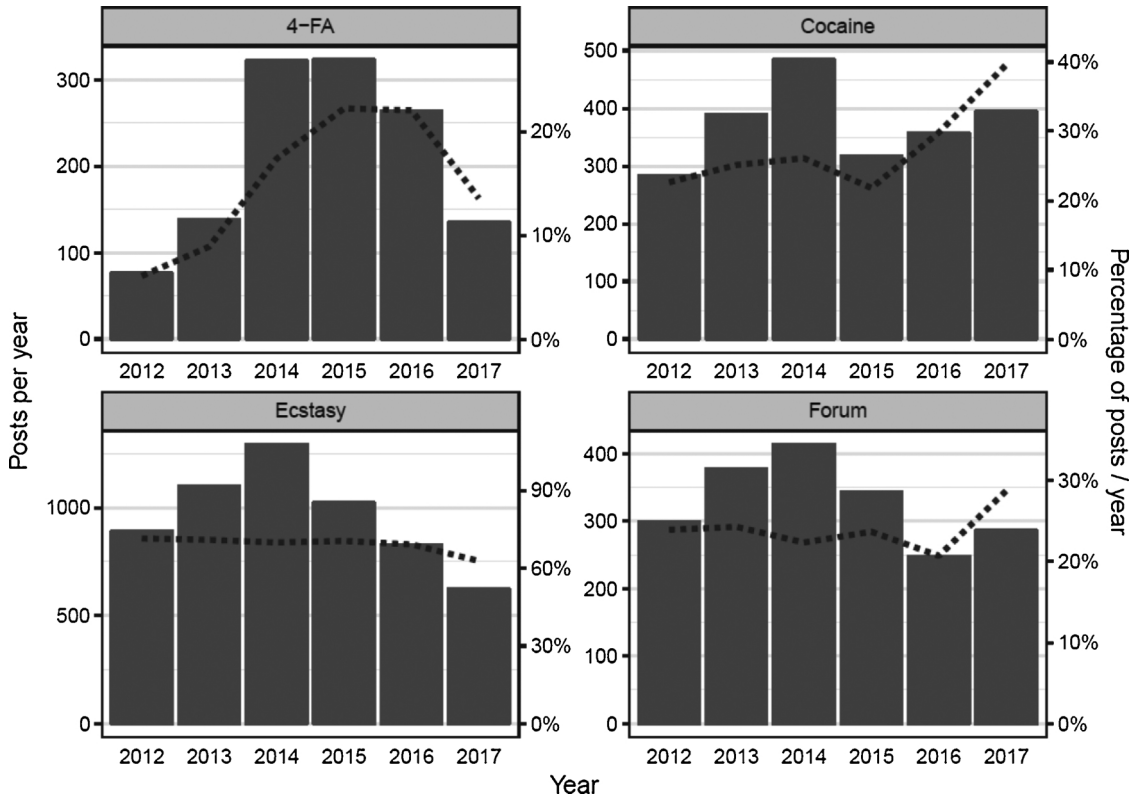

Fig. 1. Number and proportion of 4-FA, ecstasy, cocaine and neutral substance non-specific posts, years 2012–2017.
Note. This plot presents the number of posts (bars, counts on the left vertical axis) and proportion of posts (dotted line, percentages on the right vertical axis). This percentage is calculated as the number of posts on 4 FA, divided by the total number of posts on the four post categories for the same year (x100%). Data from the two drug forums are combined for this plot.

experience with the subject discussed in the posts, as responses are dependent on the information in the opening post. Within the opening posts, we only used the 200 characters before and the 200 characters after the search term occurred in the posts, to optimally capture sentiments related to the substance of interest. To give an impression, part of a paraphrased and translated example of one of the analysed 4-FA posts is: "[…] I do not have much experience with drugs, but over the last few years I have occasionally used 4-FA on dance festivals. For me this is the nicest drug. I have tried other drugs, but these made me vomit, […]".

Sentiment analysis using pattern

In a next step, relevant posts and date of posting from the two forums have been combined in one data set to gain optimal statistical power. Redundant posts were excluded from the analysis. Next, the contents of the opening posts were semantically analysed using the pattern.nlp package for the R statistical programming environment, which relies on Pattern for its text analysis (de Smedt & Daelemans, 2012). Pattern has functions for sentiment analysis, and returns a sentiment score between -1 and 1 for each text fragment. A score close to -1 indicates a very negative sentiment in the text (e.g. "This is very awful!"), whereas a score close to 1 indicates very positive sentiment (e.g. "This is very good!"). !"). A score around 0.5 indicates a moderately positive sentiment (e.g. "This is okay."). A negative nor positive post (e.g. "This is a fact") would receive a score close to 0.

News exposure data

In order to evaluate whether changes in sentiments coincide with news media coverage and exposure related to 4-FA, we extracted the relative number of news searches in the Netherlands related to 4-FA between 2012–2018 from Google Trends (2018). Google Trends does not provide information about the actual number of searches, but does provide a relative number of searches for the selected time interval standardized between 0 and 100, which indicates on which dates relatively many searches have been performed, and on which dates relatively few.

Statistical analysis

The distribution of the number of posts over the years was analysed using a Chi-squared test. The overall sentiments per search term (e.g. 4-FA, ecstasy, …) were compared among the search terms using *t*-tests. Based on the polarity score of each post in our post corpus, we analysed changes in sentiments towards 4-FA and the other substances among the forum posters between 2012 and 2017. We contrasted the polarity of 4-FA posts with the polarity of posts on ecstasy, cocaine, and our substance non-specific control posts (containing the word 'forum') using linear mixed modelling (LMM; R package lme4) to analyse relative changes in polarity of post sentiments over time. In this analysis, a random intercept for the forum where the post appeared was fitted. Time (year of the post) was entered as a fixed factor variable in the model. In addition, a fixed intercept was included in the model. Separate LMMs were fitted for each of the post categories (4-FA, ecstasy, cocaine, and forum). In addition, a set of LMMs were fitted for each of the post categories with Time as a numeric instead of a factor variable. In this analysis, both linear, quadratic and cubic time trends have been evaluated; the other specifications were the same as for the other set of LMMs.

Results

Number of forum posts

Between January 1, 2012 and January 1, 2018, a total of 8296 unique opening posts containing one or more of the search words were found. Of those, 5759 contained one or more ecstasy-related search words, 2225 posts were cocaine-related, 1968 contained the word forum, and 1258 were 4-FA related. Almost a third of the posts (30.4%) contained terms referring to more than one of the search words. Over the years, some variation in the total number of posts containing the search terms was observed. There were 1249 posts in 2012; 1551 in 2013; 1852 in 2014; 1452 in 2015; 1197 in 2016 and 995 posts in 2017. Chi-square statistics confirm that this number is not stable over time for any of the post categories, 4-FA: $X^2(5) = 271.8$, $p < .001$; Ecstasy: $X^2(5) = 281.0$, $p < .001$; Cocaine: $X^2(5) = 66.1$, $p < .001$; Forum: $X^2(5) = 58.1$, $p < .001$. The number of posts related to 4-FA shows a clear increase between 2012 and 2015, and a decrease thereafter (Fig. 1). Proportional to the total number of posts belonging to each of the post categories, there is an overall year-by-year increase observed in the number of 4-FA related posts between 2012–2015 and again a decrease thereafter. Observed trends for the other post categories are presented in Fig. 1.

Sentiment of forum posts

The mean sentiment polarity of the posts on 4-FA in 2012–2017 ($M = 0.085$ on a -1 to +1 scale; $SD = 0.23$) is not significantly different from the polarity of posts on ecstasy ($M = 0.095$; $SD = 0.22$; $t(1813) = 1.45$; $p = 0.15$), while it is marginally but not significantly higher than cocaine-related posts ($M = 0.072$; $SD = 0.21$; $t(2476) = 1.68$; $p = 0.092$). The sentiment expressed in posts containing the word 'forum' was however more positive than the sentiment expressed towards 4-FA ($M = 0.122$; $SD = 0.23$; $t(2710) = 4.52$; $p < .001$).

Over time, an increase in sentiment polarity in the period before 2015 can be observed, followed by a decrease in sentiment polarity towards 4-FA after 2015 (Fig. 2). In 2015, the sentiment score for 4-FA is at its maximum, indicating that in 2015 more positive and less negative posts containing '4-FA' have appeared, relative to the other years. In the other substances and in the substance non-specific posts category, no such pattern was observed. The positive polarity trend in 4-FA posts between 2012–2015 precedes the increase in 4-FA related acute health incidents as registered by MDI and DIMS among young people in 2013–2017 (Fig. 3). As can be further observed in Fig. 3, in the years 2012 and 2013, changes in sentiment scores are not reflected in changes in news exposure, whereas from 2014, changes in sentiment scores temporally coincide with changes in news exposure based on historic data about Google Trends news searches.

Mixed modelling analysis, with the forum source as a random intercept, confirms this observation. Based on the observed non-linear association between Year and Sentiment score (Fig. 2), we have fitted up to a 3rd order polynomial of Year in the model. For 4-FA, the association between Sentiment score (dependent) and Year as a quadratic term ($B = -0.470$; $SE = 0.224$; $t(1254) = 2.10$; $p = 0.036$) is found to be significant, indicating a non-linear association between Year and Sentiment score. The cubic term has a marginally significant association with Sentiment score ($B = -0.424$; $SE = 0.224$; $t(1253) = 1.89$; $p = 0.059$), and the linear, first order polynomial has no significant association with Sentiment score ($p = 0.81$). With Year entered in the model as a categorical variable with the year 2017 as the reference category, the difference in Sentiment score between 2015 and 2017 was found to be statistically significant ($B = 0.062$; $SE = 0.023$; $t(1252) = 2.70$; $p = 0.007$).

For cocaine, a cubic association between Year and Sentiment score was found ($B = -0.483$; $SE = 0.211$; $t(2221) = 2.28$; $p = 0.023$) in the mixed modelling polynomial regression analysis. The first ($p = 0.16$) and second ($p = 0.39$) order polynomial did not show a significant association. For ecstasy and forum, no first ($p = 0.77$ and $p = 0.27$, resp.), second ($p = 0.79$ and $p = 0.23$, resp.), nor third ($p = 0.42$ and $p = 0.25$, resp.) order polynomial association between Year and Sentiment score was found.

Discussion

In this paper we have assessed the utility of online drug user forum data for drug sentiment trend monitoring. We focused on 4-FA, a recently controlled new psychoactive substance, increasingly reported by various drug monitors in the Netherlands over the last years. We have addressed the questions to what extent changes in the volume of 4-FA related posts can be observed around the time this substance was increasingly seen by other monitoring services; and whether changes in the sentiments of the posts on 4-FA can be observed between 2012 and 2017.

With regard to the volume of posts, we found that the number of 4-FA related posts increased more than 4-fold between the years 2012–2015. This increase occurred somewhat before the increased popularity of 4-FA in the Netherlands was reported by other monitoring activities in the Netherlands (e.g. based on the health incidents covered by MDI) and before an increase in the number of Google searches related to news about 4-FA could be observed. The MDI data over 2015 and 2016 illustrates the increase of incidents, as 16% of the acute toxic effects reported by first aid stations of large-scale events in 2016 were 4-FA related, compared to less than 1% in 2012 and 2013, 2% in 2014, and 11% in 2015 (Wijers, Croes et al., 2017; Wijers, van Litsenburg et al., 2017). The increased news exposure of adverse 4-FA incidents after 2015 is reflected in a gradual decrease in sentiment scores in those years. Whether this association is due to an impact of news coverage and exposure on the sentiment as expressed in the forums cannot be evaluated using the data we have. A possible explanation could be that before using the substance, potential users visited online forums to inform themselves about 4-FA, e.g. about effects, effective dosage, and risks. An alternative explanation is that these forums facilitated access to information on new substances such as 4-FA and hence fostered the increase in 4-FA use and related accidents. The increase in sentiment presented in Fig. 3 in the second half of 2017 is difficult to interpret given the large confidence intervals in this period.

Parallel with the increase in the number of posts on 4-FA in 2012–2015, a change in the sentiment of posts related to 4-FA was observed. Although these changes were not as pronounced as the changes in the number of posts, the sentiment of the content of the posts moved somewhat towards the positive end of the polarity spectrum. This means that more positive and less negative words and phrases were used in posts on 4-FA and can be interpreted as that messages about 4-FA had a more positive sentiment towards 4-FA. After 2015, a reverse trend was observed: less new posts appeared each year, and the sentiment of the content of those posts was less positive. These trends coincided with or even preceded increases in health incidents related to 4-FA as reported by MDI, and (increased) media coverage in 2015–2017 (Google Trends, 2018) in which among other things side effects, and health incidents related to 4-FA were reported (van der Pol et al., 2017).

Strengths and limitations

This study is among the first to present both a quantitative and a qualitative evaluation of online drug users forum analysis and is as such one of the first to evaluate the practical utility of online forum data sentiment trend monitoring. Although the first findings indicate that monitoring of drug users forums may be useful for drug use trend monitoring, the results should be interpreted in the light of the limitations of this study.

First, we included data from only two forums in our analysis. Although these two forums are among the most active drug forums in

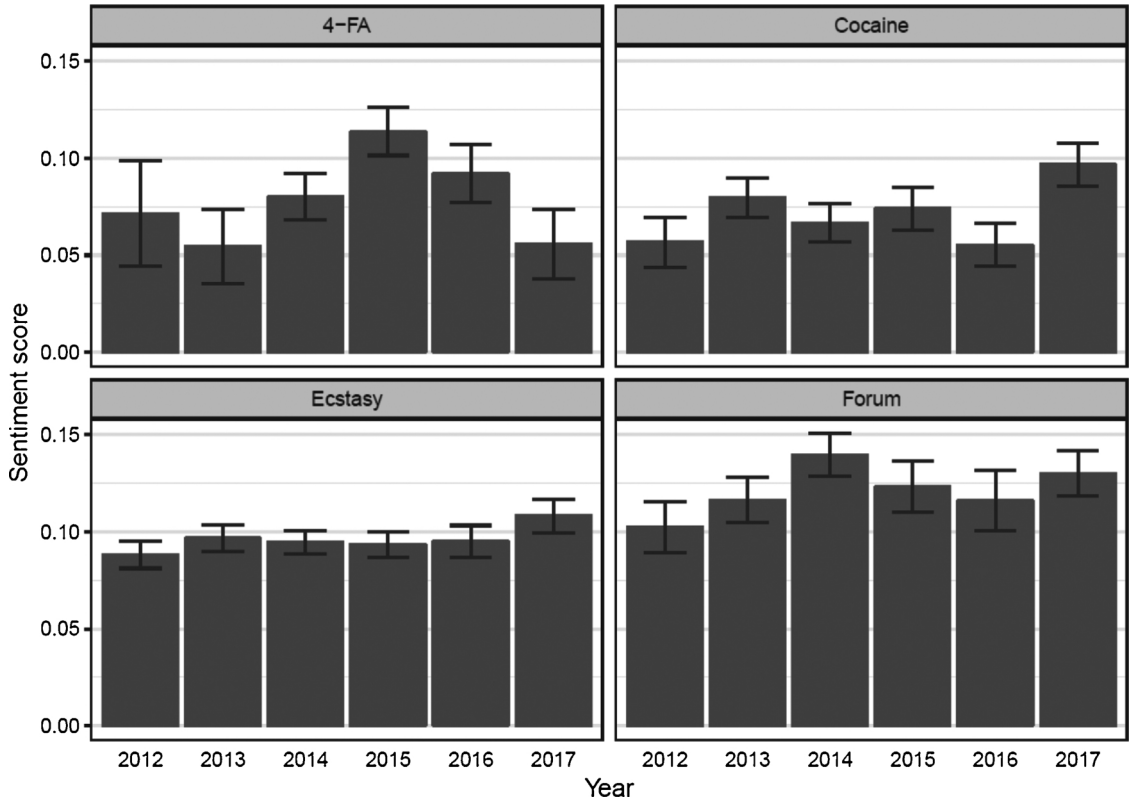

Fig. 2. Sentiment towards 4-FA, ecstasy, cocaine and substance non-specific posts, years 2012–2017.
Note. Year by year sentiment scores for 4-FA, cocaine, ecstasy and forum (substance non-specific posts). The sentiment score has a theoretical range between -1 and 1, with 0 expressing a neutral sentiment. The error bars in the plot indicate the ± 1 standard error range around the mean.

the Netherlands, it is possible that some trends in sentiments among 4-FA users may have been overlooked, e.g. due to (missing) slang words for substances in our search terms, or possible non-representativeness of drug forum users to the wider population of NPS users. Second, the results of forum sentiment analyses are merely observational in nature, and hence we cannot verify or falsify hypotheses about mechanisms of change of 4-FA sentiments and 4-FA use. Therefore we can conclude that changes in sentiments coincide with or precede changes in the number of health incidents and with news coverage of adverse events due to 4-FA use, but we cannot support nor reject the hypothesis that one of those changes has caused or was caused by any of the other changes. Thus far, we also do not know to what extent changes in number of posts or sentiments can be predictive of changes in popularity or use of a substance in the recent past or near future. In order for forum sentiment trend monitoring to have practical value for drug use monitoring purposes, additional research regarding the predictive validity of automated forum sentiment monitoring to discover emerging trends in (attitudes towards) substance use behaviour is needed. This could for example be done by including independent researcher coding of text to determine the validity of automated sentiment monitoring.

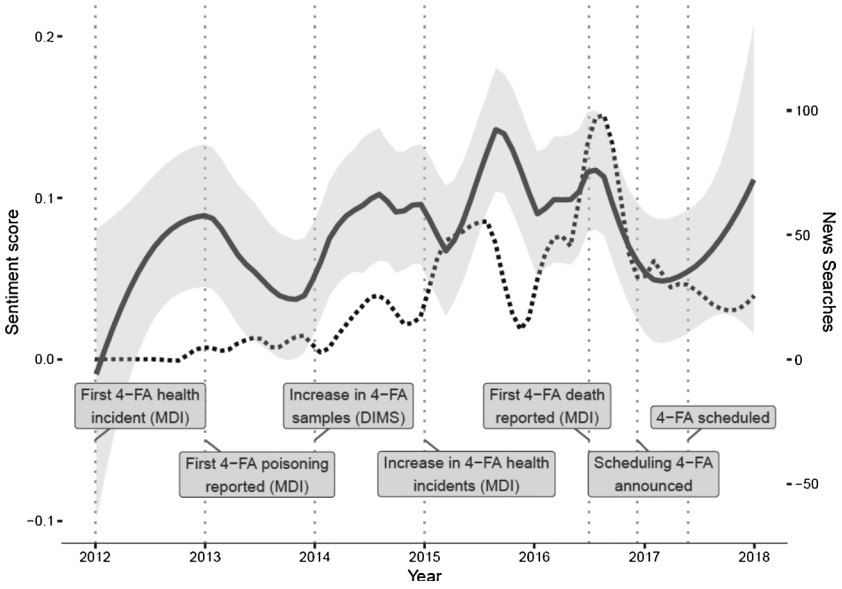

Fig. 3. Sentiment polarity trend in forum posts and timeline of news exposure and adverse events related to 4-FA in the Netherlands.
Note. Association between news coverage and exposure based on Google Trends information on the number of 4-FA news searches, adverse events related to 4-FA (presented in van der Pol et al. (2017)) and sentiment towards 4-FA expressed in forum posts. On the left vertical axis, A sentiment score of 0 indicates that on average the sentiment towards 4-FA is not positive nor negative. A sentiment score > 0 indicates an averaged positive sentiment towards 4-FA. The solid line which indicates the smoothed sentiment score and the corresponding confidence interval (in grey) and is obtained by fitting the date of the posts to the sentiments of the posts using local regression (LOESS). On the right vertical axis, standardized Google Trends data have been presented. A score close to 0 indicates few or no news searches, a score of 100 indicates the maximum number of 4-FA related news searches. The dotted line indicates the LOESS smoothed fit through the number of news searches. MDI is the Monitor Drug-related Incidents in the Netherlands; DIMS is the Dutch Drugs Information and Monitoring System.

A clear strength of this method is that a large amount of data with high ecological validity can be used to monitor trends in drug use discussions. As these data can be obtained without having a researcher interfere with the actual discussion there is no risk that the researcher unwillingly influences the discussion and, as a consequence, contaminates the data. In addition, the discussions on these drug forums are initiated by users of the forum themselves, and are not suggested by a research community. And as long as the data are available through the web fora, new emerging substances and trends and what has been discussed about these substances and trends, can easily and reliably be evaluated retrospectively using drug forum data without recall bias. These analyses could be sentiment analyses such as we have presented in this paper, but could also be more descriptive in nature (see e.g. Bilgrei (2016) and Soussan and Kjellgren (2014)) or could look into reported effects of (prescription) drug abuse (see e.g. Cameron et al., 2013; Nikfarjam, Sarker, O'Connor, Ginn, & Gonzalez, 2015)), or a combination thereof. As the forums are continually used and new posts are immediately available, it is relatively easy for a forum data researcher to work with very recent data.

Ethical aspects of drug forum monitoring

Also with regard to the research ethics discussion, drug forum monitoring is still in its infancy. In a recent systematic review (Golder, Ahmed, Norman, & Booth, 2017), 17 studies that conducted qualitative methods to collect data on attitudes on the ethical implications of research using social media were included. Expressed attitudes varied strongly. Some studies concluded that social media research was essential for the public good, while others where very concerned about ethical aspects of this type of research. The authors conclude that many conflicting issues contribute to the complexity of good ethical practice in social media research (Golder et al., 2017). Another recent review on using Twitter for health research concludes that policies that address privacy and ethical concerns in social media research are needed (Sinnenberg et al., 2017).

In order to take into account and acknowledge the ongoing discussion on social media research ethics, and to optimally respect the privacy of the forum users, in this study we have adhered to the forum guidelines which apply to the forums we used in our analyses. In addition, we have refrained from collecting any material which could link posts to individual forum users or their nick names, and we further refrained from collecting data which would allow us to link individual posts to a common source or forum user. As such, we have minimised the amount of personal data in our database. As drug forums tend to work with anonymised users (e.g. the do not provide email addresses or contact details of forum users) it was not feasible to contact individual forum members to ask for their consent to use their posts for research purposes.

Conclusion

In this paper, we have tested the feasibility of drug forum sentiment analysis. Using our approach, it was feasible to detect trends in both the volume of posts, and in the sentiment expressed in those posts towards 4-FA, a recently controlled NPS in the Netherlands. Changes in the volume and sentiments of forum posts coincided with news media exposure related to 4-FA and with trends observed by established drug monitoring sources. In future research we hope to further evaluate the predictive validity of drug sentiment trends expressed in drug user forums. This may facilitate earlier discovery of trends in the popularity, prevalence and adverse events related to scheduled drugs and NPS.

Declarations of competing interest

The authors declare that they have no competing interest to declare.

Acknowledgements

This work was supported by a grant from the Ministry of Health, Welfare and Sport (VWS) to the Trimbos Institute for drug monitoring innovation.

References

Bilgrei, O. R. (2016). From "herbal highs" to the "heroin of cannabis": Exploring the evolving discourse on synthetic cannabinoid use in a Norwegian Internet drug forum. *The International Journal of Drug Policy, 29*, 1–8.

Cameron, D., Smith, G. A., Daniulaityte, R., Sheth, A. P., Dave, D., Chen, L., ... Falck, R. (2013). PREDOSE: A semantic web platform for drug abuse epidemiology using social media. *Journal of Biomedical Informatics, 46*(6), 985–997.

Croes, E. A., de Ruiter, N., Wijers, C. H., Niesink, R., Brunt, T. M., & van Goor, M. (2017). *Fact sheet 4-FA: Update October 2017*. Utrecht: Trimbos-instituut.

Curtis, B., Alanis-Hirsch, K., Kaynak, Ö., Cacciola, J., Meyers, K., & McLellan, A. T. (2015). Using Web searches to track interest in synthetic cannabinoids (a/k/a' herbal incense'). *Drug and Alcohol Review, 34*(1), 105–108.

Davey, Z., Schifano, F., Corazza, O., Deluca, P., Web, P., & Mapping Group (2012). e-Psychonauts: Conducting research in online drug forum communities. *Journal of Mental Health, 21*(4), 386–394.

Golder, S., Ahmed, S., Norman, G., & Booth, A. (2017). Attitudes toward the ethics of research using social media: A systematic review. *Journal of Medical Internet Research, 19*(6), e195.

van der Gouwe, D., & Rigter, S. (2016). *). Annual report 2015. Drugs information and monitoring system (DIMS)*Utrecht: Trimbos-instituut.

van der Gouwe, D., Brunt, T. M., van Laar, M., & van der Pol, P. (2017). Purity, adulteration and price of drugs bought on-line versus off-line in the Netherlands. *Addiction, 112*(4), 640–648.

Google Trends (2018) Trends in 4fa, 4fmp, 4-fluoramphetamine news items (01-01-2012–31-12-2017). https://trends.google.com/trends/explore?date=2012-01-01%202017-12-31&geo=NL&gprop=news&q=4fmp.4fa.%2Fm%2F081sxj. [Accessed: 26 July 2018]. Archived: https://archive.is/2HUk7.

Hondebrink, L., Nugteren-van Lonkhuyzen, J. J., van der Gouwe, D., & Brunt, T. M. (2015). Monitoring new psychoactive substances (NPS) in the Netherlands: Data from the drug market and the poisons information centre. *Drug and Alcohol Dependence, 147*, 109–115.

Lamy, F. R., Daniulaityte, R., Nahhas, R. W., Barratt, M. J., Smith, A. G., Sheth, A., ... Carlson, R. G. (2017). Increases in synthetic cannabinoids-related harms: Results from a longitudinal web-based content analysis. *The International Journal of Drug Policy, 44*, 121–129.

Linsen, F., Koning, R. P., van Laar, M., Niesink, R. J., Koeter, M. W., & Brunt, T. M. (2015). 4-Fluoroamphetamine in the Netherlands: More than a one-night stand. *Addiction, 110*(7), 1138–1143.

Ministry of Health, Welfare and Sport (2017). *Verbod op drug 4-Fluoramfetamine gaat donderdag 25 mei in [As of May 25, the substance 4-Fluoramphetamine will be scheduled]*. 2017, Retrieved from WebCitation website:Den Haag: Rijksoverheid. http://www.webcitation.org/6y8XZjAZX.

Monshouwer, K., Drost, Y., van der Pol, P., & van Laar, M. W. (2016). *Het Grote Uitgaansonderzoek 2016*. Utrecht: Trimbos-instituut.

Nikfarjam, A., Sarker, A., O'Connor, K., Ginn, R., & Gonzalez, G. (2015). Pharmacovigilance from social media: mining adverse drug reaction mentions using sequence labeling with word embedding cluster features. *Journal of the American Medical Informatics Association, 22*(3), 671–681.

Norbutas, L. (2018). Offline constraints in online drug marketplaces: An exploratory analysis of a cryptomarket trade network. *The International Journal of Drug Policy, 56*, 92–100.

van der Pol, P., Nijkamp, L., Nabben, T., & van Laar, M. (2017). *4-Fluoramfetamine: gebruikers en gebruik in beeld [4-Fluoramphetamine: an assessment of consumers and patterns of use]*. Utrecht/Amsterdam: Trimbos institute/Bonger institute for Criminology.

Sinnenberg, L., Buttenheim, A. M., Padrez, K., Mancheno, C., Ungar, L., & Merchant, R. M. (2017). Twitter as a tool for health research: A systematic review. *American Journal of Public Health, 107*(1), e1–e8.

de Smedt, T., & Daelemans, W. (2012). Pattern for Python. *Journal of Machine Learning Research, 13*, 2031–2035.

Ribeiro, F. N., Araújo, M., Gonçalves, P., Gonçalves, M. A., & Benevenuto, F. (2016). SentiBench–A benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Science, 5*(1), 23.

Soussan, C., & Kjellgren, A. (2014). Harm reduction and knowledge exchange-a qualitative analysis of drug-related Internet discussion forums. *Harm Reduction Journal, 11*, 25.

Tay Wee Teck, J., & McCann, M. (2018). Tracking internet interest in anabolic-androgenic steroids using Google Trends. *The International Journal of Drug Policy, 51*, 52–55.

Thanki, D., & Frederick, B. (2016). *Social media and drug markets. EMCDDA. The internet and drug markets*. Luxembourg: Publications Office of the European Union.

Tzanetakis, M. (2018). Comparing cryptomarkets for drugs. A characterisation of sellers and buyers over time. *The International Journal of Drug Policy, 56*, 176–186.

UNODC (2016). *New psychoactive substances.* Retrieved from WebCitation website:Vienna: UNODC. http://www.webcitation.org/6y8XSZK0H.

UNODC (2018). *World drug report 2018*Retreived from UNODC website:. Vienna: UNODC. http://www.unodc.org/wdr2018/.

Wijers, C. H., Croes, E. A., de Ruiter, N., & Valkenburg, H. (2017). *Monitor Drugs incident: Fact sheet 2016 [Monitor drug incidents: Fact sheet 2016]*. Utrecht: Trimbos-instituut.

Wijers, C. H., van Litsenburg, R. T., Hondebrink, L., Niesink, R. J., & Croes, E. A. (2017). Acute toxic effects related to 4-fluoroamphetamine. *Lancet, 389*(10069), 600.