analytical chemistry

pubs.acs.org/ac Article

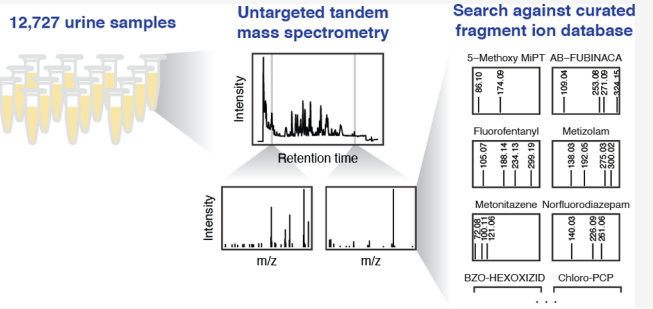
Identification of Emerging Novel Psychoactive Substances by Retrospective Analysis of Population-Scale Mass Spectrometry Data Sets

Michael A. Skinnider,* Sandrine A. M. Mérette, Daniel Pasin, Jason Rogalski, Leonard J. Foster, Frank Scheuermeyer, and Aaron M. Shapiro

Cite This: https://doi.org/10.1021/acs.analchem.3c03451 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Over the last two decades, hundreds of new psychoactive substances (NPSs), also known as "designer drugs", have emerged on the illicit drug market. The toxic and potentially fatal effects of these compounds oblige laboratories around the world to screen for NPS in seized materials and biological samples, commonly using high-resolution mass spectrometry. However, unambiguous identification of a NPS by mass spectrometry requires comparison to data from analytical reference materials, acquired on the same instrument. The sheer number of NPSs that are available on the illicit market, and the pace at which new compounds are introduced, means that forensic laboratories must make difficult decisions about which reference materials to acquire. Here, we asked whether retrospective suspect screening of population-scale mass spectrometry data could provide a data-driven platform to prioritize emerging NPSs for assay development. We curated a suspect database of precursor and diagnostic fragment ion masses for 83 emerging NPSs and used this database to retrospectively screen mass spectrometry data from 12,727 urine drug screens from one Canadian province. We developed integrative computational strategies to prioritize the most reliable identifications and tracked the frequency of these identifications over a 3 year study period between August 2019 and August 2022. The resulting data were used to guide the acquisition of new reference materials, which were in turn used to validate a subset of the retrospective identifications. Last, we took advantage of matching clinical reports for all 12,727 samples to systematically benchmark the accuracy of our retrospective data analysis approach. Our work opens up new avenues to enable the rapid detection of emerging illicit drugs through large-scale reanalysis of mass spectrometry data.



INTRODUCTION

Over the last two decades, the illicit drug market has been reshaped by the proliferation of novel psychoactive substances (NPS), also known as "designer drugs" or "legal highs".[1,2] These compounds are often structurally similar to the existing drugs of abuse but with slight modifications to their chemical structures that allow them to circumvent drug control laws.[3] They are commonly derived from the scientific or patent literature, synthesized by clandestine chemists, and packaged for sale as innocuous products such as "bath salts" or "air fresheners", with the disclaimer that they are "not for human consumption".[4] Because the vast majority of NPSs have never been evaluated in preclinical studies, their safety profiles are rarely well-characterized, and a number of them have been associated with serious toxidromes or fatalities.[5,6] As a result, the emergence of NPS has become a major public health concern, particularly given the proliferation of NPSs that carry a high potential for overdose and death, such as synthetic opioids or designer benzodiazepines.[1,5]

The toxic and potentially fatal effects of NPSs oblige clinical, forensic, and law enforcement laboratories around the world to screen for these compounds in seized materials and biological samples. High-resolution mass spectrometry (HR-MS) has emerged as one of the primary methods used to detect and identify NPS.[7] Unlike immunochemical approaches, HR-MS can provide component-resolved drug profiles, and unlike nuclear magnetic resonance (NMR) spectroscopy, HR-MS can be applied to detect low-abundance compounds within complex biological mixtures like blood or urine.[8,9] Despite these advantages, several factors make screening for NPS by HR-MS a challenging task. Unambiguous identification of an

Received: August 2, 2023
Revised: October 13, 2023
Accepted: October 23, 2023

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

NPS by mass spectrometry requires data from analytical reference materials, acquired contemporaneously on the same instrument. However, the costs of these reference materials can be prohibitive, and even when they are not, the sheer volume and diversity of these compounds make it impractical to acquire reference materials for every NPS.[10] Moreover, because these compounds are often introduced to circumvent legislation, the introduction of new legislation to regulate NPS can prompt the emergence of a new generation of compounds, requiring an entirely new suite of reference materials and pitting analytical laboratories against clandestine chemists in a cat-and-mouse game.[11]

With so many new potentially toxic substances emerging on the illicit market, analytical laboratories must make difficult decisions about which NPS to prioritize for the development of new assays. Generally, these decisions require individual scientists to consider and integrate multiple sources of information.[12,13] At the British Columbia Provincial Toxicology Centre (Vancouver, BC, Canada), for example, decisions about which reference materials to acquire are based on a compound's reported national and international prevalence, its toxicity and forensic value, and clinician input. This information must be culled from a number of different sources, including the scientific literature, monographs, communications from public health agencies, and forensic databases, and then synthesized by experts in a subjective manner.

In an ideal scenario, decisions about which reference materials to acquire would be made in a more data-driven manner. One potential strategy toward this end is to reanalyze mass spectrometry data from clinical or forensic samples in order to identify NPSs that may be present in these samples but which are not detected by existing assays. In the absence of purchased reference materials, these NPSs could be presumptively identified by comparison to published mass spectrometry data acquired by other laboratories, such as that compiled in databases like MassBank,[14] Thermo mzCloud, or HighResNPS.[10,15] This is an example of a suspect screening approach, whereby compounds from a suspect list are tentatively identified by searching for the accurate masses of their precursor and/or fragment ions, although comparison to reference standards is ultimately required for unambiguous identification.[16] The concept of retrospective data analysis has previously been explored,[13,17−23] but key questions remain: for example, how best to leverage mass spectrometric data acquired on different instruments; how to refine tentative identifications to prioritize the most confidently detected NPS; and how to estimate the sensitivity and specificity of retrospective approaches.

Here, we asked whether we could retrospectively leverage large volumes of existing mass spectrometry data in order to prioritize emerging NPSs for assay development. We curated a database of precursor and diagnostic fragment ion masses for 83 emerging NPSs, for which reference materials were unavailable or had not been acquired. We then used this database to retrospectively search more than 12,000 mass spectrometry runs from urine drug screens performed at the British Columbia Provincial Toxicology Centre. We developed integrative computational strategies to prioritize the most reliable identifications and tracked the frequency of these identifications over a 3 year study period. We used these data to guide the acquisition of new reference materials, which were in turn used to validate a subset of retrospective identifications. Our data suggest that large-scale retrospective analysis of data from clinical or forensic samples can contribute to data-driven decision-making in analytical laboratories.

METHODS

Clinical Samples. A retrospective analysis was performed of all urine drug screens performed at the Provincial Toxicology Centre at the British Columbia Centre for Disease Control between August 1, 2019 and August 31, 2022. Samples originated from hospitals and clinics throughout the province of British Columbia. Urine samples were received in sterile urine containers or vacutainers with no preservatives. All 12,727 consecutive samples collected over the study period were included in the study. There were no exclusion criteria. Samples were assigned anonymized identifiers for all analyses described here. The study was approved by the UBC Clinical Research Ethics Board (#H22−02722).

Mass Spectrometry. Urine samples were analyzed by liquid chromatography−high-resolution mass spectrometry. Full scan with targeted data-dependent MS^2 (full MS/dd-MS^2) was performed in positive electrospray ionization mode with an inclusion list containing over 200 drugs. Complete details are provided in the Supplementary Information.

Reference Materials. Reference standards (bromazolam, *para*-fluorofentanyl, eutylone, deschloroetizolam, and furanyl UF-17) were purchased from Cayman Chemical and prepared as described in the Supplementary Information.

Suspect Database. A database of emerging NPSs that have been identified in the seized drugs in Canada or by organizations that focus on NPS detection (e.g., CFSRE and UNODC), but that are not currently reported to clinicians in British Columbia, was manually curated by two of the authors (S.A.M.M. and A.M.S.) as part of the standard operating procedures of the British Columbia Provincial Toxicology Centre. NPS structures and accompanying analytical data (i.e., fragment ions and retention time, when available) were obtained from a variety of sources, including HighResNPS,[10,15] monographs published by the Center for Forensic Science Research & Education's NPS Discovery early warning system, the Thermo mzCloud mass spectral database, and individual publications. In total, the database comprised 83 compounds, each of which was associated with a precursor mass and between one and nine fragment ions (Figure S1). The suspect database is provided in Table S1.

Retention Time Prediction. A previously developed machine learning model[24] was adapted here to predict the retention times of unseen compounds on our LC gradient. The training data set consisted of 4,846 retention times from 24 laboratories including 153 retention times from the LC-HR-MS method. Retention times for 21 compounds were held out as a test set for model validation (Figure S2). The trained model was then used to predict the retention times of the tentatively identified compounds. Additional details are provided in the Supplementary Information.

Data Analysis. Complete details are provided in the Supplementary Information. In brief, mass spectrometry files were converted to mzML format using msconvert,[25] and peak detection was performed in xcms,[26] using the "centWave" algorithm.[27] MS/MS spectra for each peak were then searched against the suspect database of emerging NPSs. Our initial filter required the detection of the precursor ion (M+H adduct) and at least two fragment ions, with tolerances of 10 and 20 ppm, respectively. Tentative NPS identifications were

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

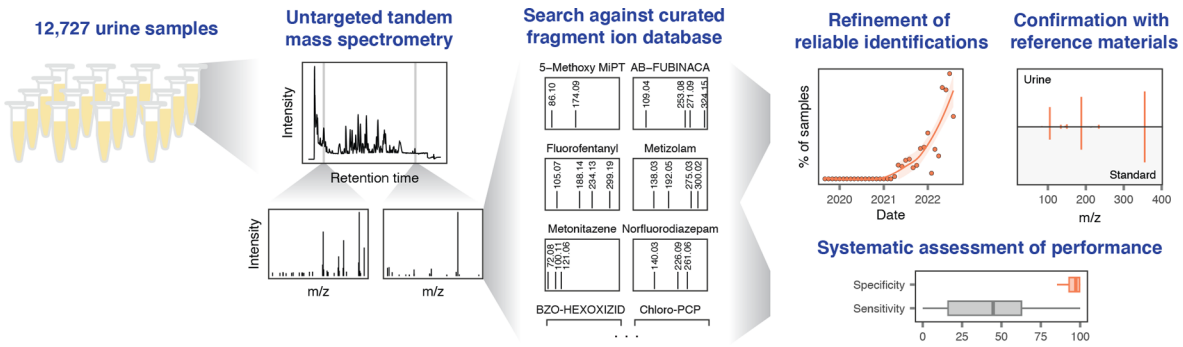


Figure 1. Schematic overview of the analyses described in this study. Full-scan mass spectrometry data from 12,727 urine samples were retrospectively searched against the curated precursor and fragment ions from a database of emerging NPSs. Tentative identifications were manually inspected and refined using multiple sources of chromatographic and spectral data. A subset of identifications were validated by the acquisition of reference spectra for five compounds. The performance of the approach was characterized more systematically by comparing clinical reports with tentative identifications based on the HighResNPS database.

then manually evaluated and refined using several additional lines of evidence, including the number and proportion of fragment ions detected, the mass error of the precursor ion, the distribution of retention times, their correspondence with retention time predictions, and the similarity between MS/MS spectra tentatively associated with a given NPS. Spectral similarity was calculated with the normalized dot-product, as implemented in the "Spectra" R package.[28] All of the codes generated in this study are available via GitHub at https://github.com/skinnider/NPS-screening.

Terminology. Throughout the paper, we refer to matches made on the basis of the accurate mass of a precursor ion and the presence of at least two diagnostic fragment ions as "tentatively identified." We use this terminology in keeping with the standard practice in the field[16,29] and notwithstanding the fact that the deliberately permissive threshold we selected to make these tentative identifications will result in a number of false-positive identifications. We refer to the identifications for which further supportive evidence was obtained by the acquisition of new reference materials as being "validated" but underscore that these identifications still do not meet the evidentiary threshold that would be required for reporting, which would require contemporaneous analysis of a re-extracted sample along with the reference material in question. Only adjudicated identifications from historical reports are referred to as "positives" or "confirmed."

RESULTS

Retrospective Data Analysis Prioritizes NPS for Assay Development. Definitive identification of a given NPS in biological samples requires comparison to authenticated reference materials analyzed contemporaneously on the same instrument. However, the rapid emergence and disappearance of new compounds on the illicit market mean that analytical laboratories must make difficult and often subjective decisions about which reference materials to acquire. We explored the possibility of using published mass spectrometry data to retrospectively analyze clinical samples in order to prioritize the detection of NPSs that may be present in these samples but not captured by existing assays. An overview of our approach is presented in Figure 1. We retrospectively analyzed a total of 12,727 urine drug screens performed by the British Columbia Provincial Toxicology Centre between August 1, 2019 and September 1, 2022, each of which was profiled by full-scan data-dependent tandem mass spectrometry. We searched each data set against a database of diagnostic fragment ions for 83 suspect NPSs, which we curated from forensic monographs, mass spectrometric databases, and the scientific literature. Tentative NPS identifications were made when two or more fragment ions in a urine MS/MS spectrum matched to a database entry, with the understanding that this liberal criterion served as a starting point that would require further refinement to eliminate false-positives caused by isobaric or coisolated molecules.

A total of 31 NPSs were tentatively identified by the presence of the precursor ion and two or more fragment ions (Figure 2). The majority of these identifications were supported by just two matching fragment ions, but 19% of the tentative NPS identifications were supported by three fragments, 2.7% by four fragments, and 3.6% by five or more fragments. When expressed instead as a proportion of the fragment ions in the suspect database that were matched in MS/MS spectra from urine samples, these numbers translated into a range of 22.2−100% of fragment ions matched. Among the tentatively identified NPSs, 17.6% matched at least 50% of the fragment ions within the suspect database.

The frequencies at which the 31 NPSs were tentatively identified varied markedly, ranging from a single sample (nine compounds) to 3,156 samples (eutylone). Ten NPSs were tentatively identified in 20 or more samples, including eutylone, fluorofentanyl, α-PiHP, 8-aminoclonazolam, bromazolam, N-ethylpentedrone, naphthyl-U-47700, metizolam, furanyl UF-17, and N-ethyl-U-47700. The frequency with which these NPSs were tentatively identified suggests that they might represent the most logical candidates for the acquisition of reference standards.

Refinement of Tentative NPS Identifications by Mass Spectral and Chromatographic Data. Our initial suspect screening criterion of two or more matching fragment ions was a deliberately permissive threshold that we selected to maximize the sensitivity of our approach, and we recognized the potential for this permissive threshold to lead to false-positive identifications. We therefore integrated multiple additional sources of mass spectrometric and chromatographic information to better assess the reliability of these tentative identifications. These sources of data included (i) the maximum number of diagnostic fragment ions overlapping with another compound in the HighResNPS database,[10,15] reflecting potential misidentifications; (ii) the mass error of the precursor ion, in ppm; (iii) the distribution of retention times

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

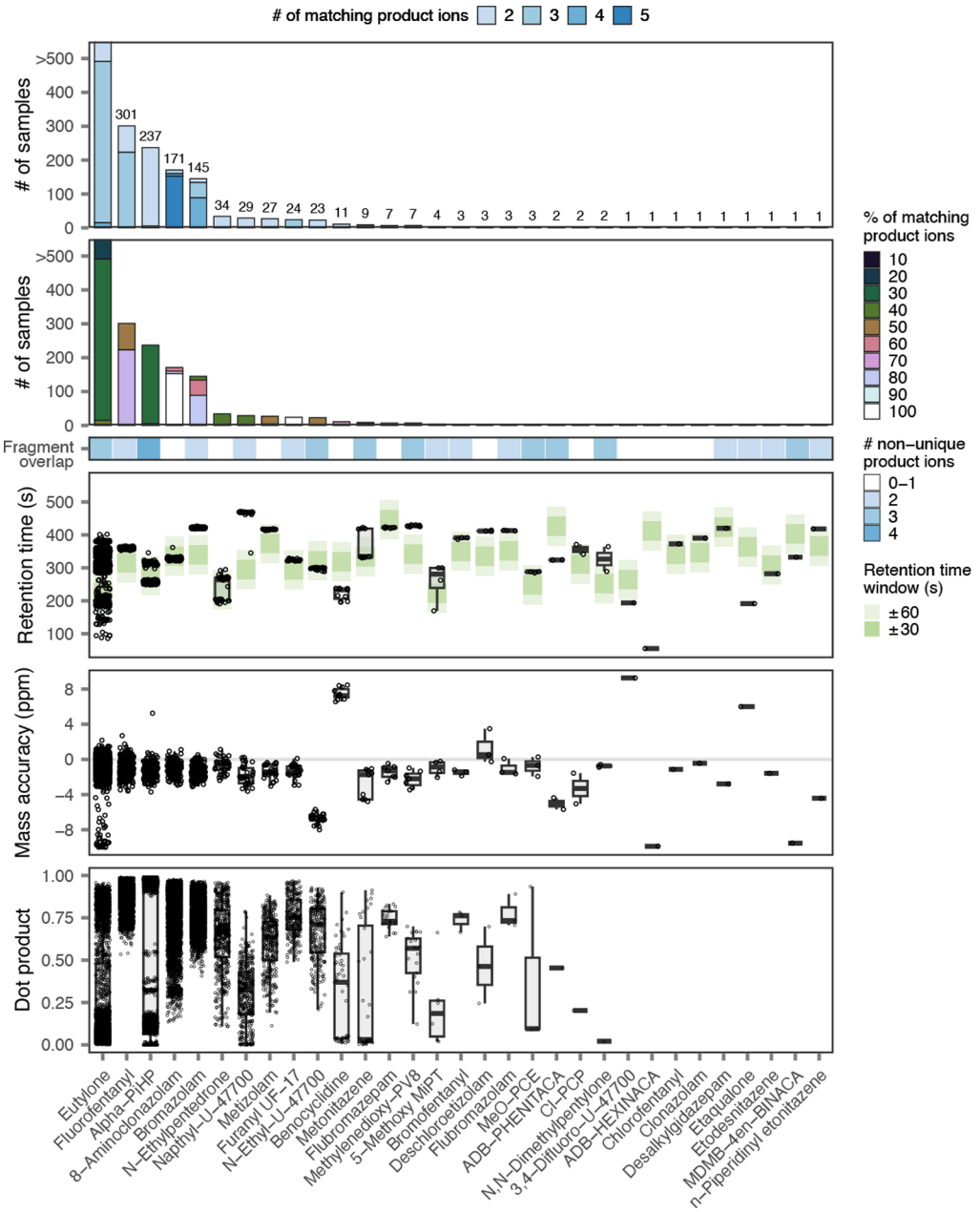


Figure 2. Overview of 31 emerging NPSs tentatively identified in urine samples by the presence of two or more diagnostic fragment ions. First row: number of database fragment ions identified in each urine sample. Inset text shows the total number of samples in which each NPS was tentatively identified. Second row: proportion of database fragment ions identified in each urine sample. Third row: maximum number of database fragment ions overlapping with the fragment ions for another compound in the HighResNPS database. Fourth row: retention times at which emerging NPSs were tentatively identified (light and dark green boxes show retention times predicted from chemical structures with ±60 and 30 s windows, respectively). Fifth row: mass error (in parts per million) for the precursor ions of tentatively identified NPSs. The sixth row shows the spectral similarity between MS/MS spectra tentatively identified as emerging NPSs, as quantified by the normalized dot-product.

at which each NPS was tentatively identified; (iv) the correspondence between the observed retention time and that predicted based on the chemical structures of each NPS by a machine-learning model;[24] and (v) the spectral similarity of tentatively identified MS/MS spectra to one another, as quantified by the normalized dot-product. Our expectation was that NPSs that were reliably identified would be characterized by a mass error close to zero, elute within a single chromatographic window that was approximately aligned with predicted retention times, share few fragments with other NPSs, and exhibit a high degree of spectral similarity to one another. Conversely, unreliable identifications would be characterized by larger or heterogeneous mass errors, elute in multiple chromatographic windows or at retention times falling well outside of the predicted windows, share multiple fragments with other NPSs, or comprise multiple clusters of mutually dissimilar mass spectra.

Inspection of these data suggested that several tentative NPS identifications did indeed reflect false-positives (Figures 2 and 3). Several NPSs that were tentatively identified, including α-PiHP, N-ethylpentedrone, and metonitazene, exhibited two distinct clusters of retention times, suggesting that our permissive initial criterion had identified a mixture of at least two compounds. Notably, both eutylone and α-PiHP share a

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

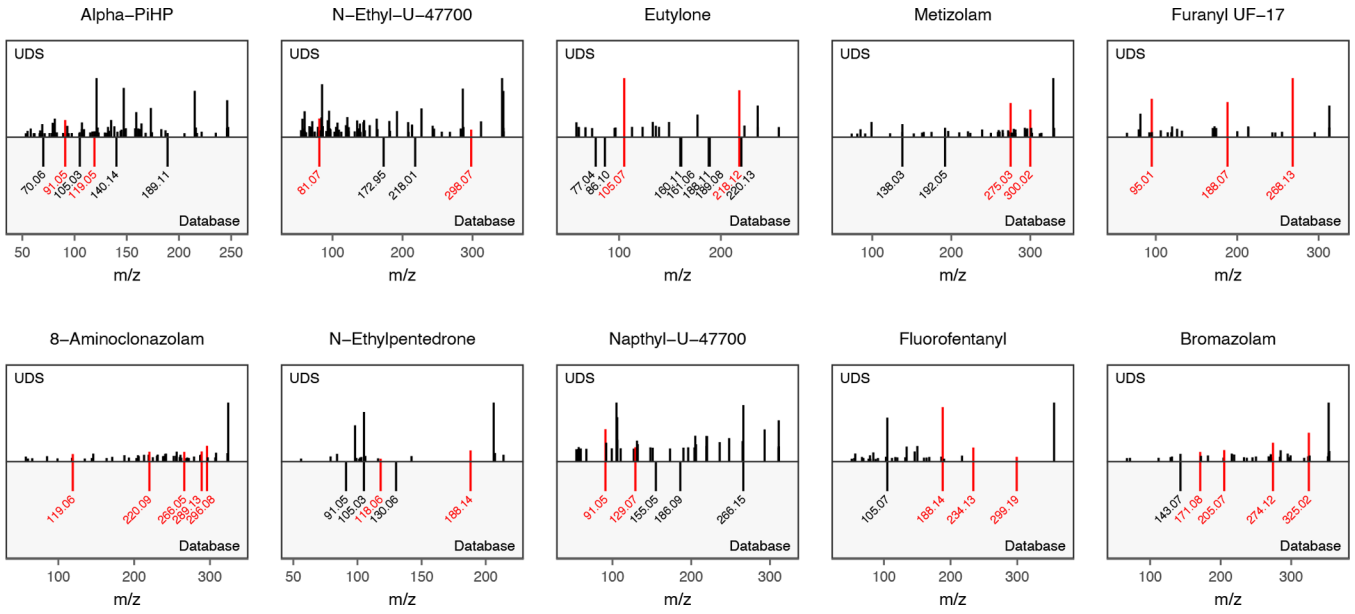


Figure 3. Representative MS/MS spectra for emerging NPSs tentatively identified based on the presence of diagnostic fragment ions from the suspect database. Mirror plots show experimental spectra acquired from urine samples, top; opposite annotated fragment ions from the suspect database, bottom. Matched fragment ions are shown in red. Experimental fragment ion intensities are square-root-transformed to improve the visualization of low-intensity fragment ions.

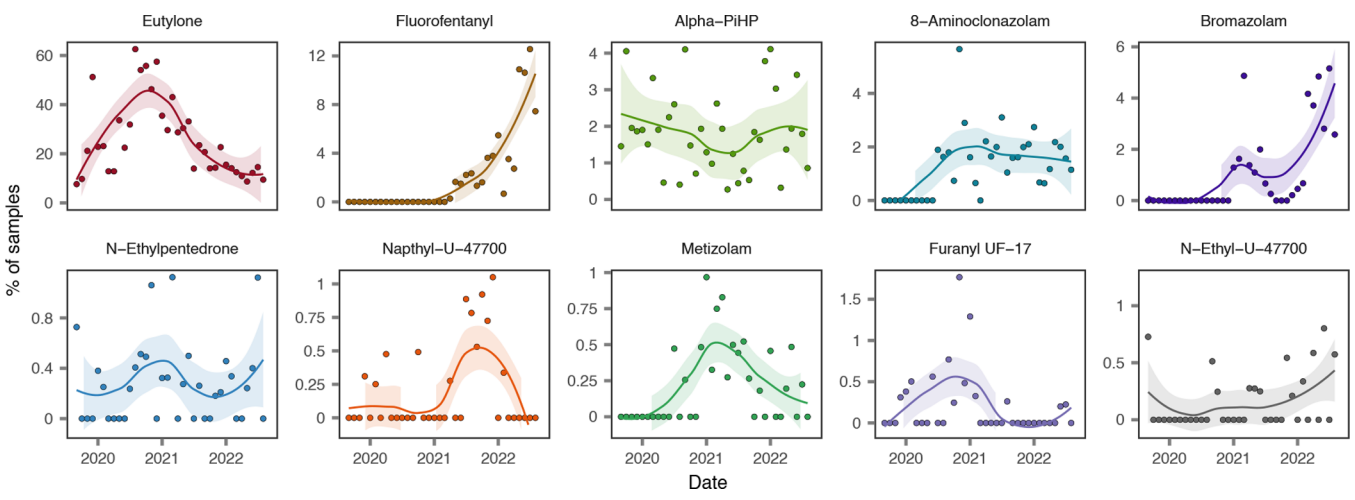


Figure 4. Proportion of samples per month in which an emerging NPS was tentatively identified over the 3 year study period, shown for 10 emerging NPSs tentatively identified in at least 20 samples.

large number of common fragment ions with at least one isobaric NPS in our database (Figure S3); for example, α-PiHP shares six fragment ions with the isomeric compound α-PHP. In the case of metonitazene, one cluster of tentative identifications, eluting between 5.50 and 5.57 min, was consistent with retention time predictions, whereas the second cluster (eluting between 6.97 and 7.04 min) was not. Other compounds eluted at retention times that were markedly distinct from the predictions, including naphthyl-U-47700, ADB-HEXINACA, or etaqualone. The notion that tentative identifications of eutylone, α-PiHP, and metonitazene actually reflected a mixture of compounds was further corroborated by the distributions of spectral similarity for these compounds, each of which exhibited at least two clusters of mutually dissimilar MS/MS spectra.

On the other hand, mass spectral and chromatographic data reinforced the reliability of several tentative NPS identifications. For instance, fluorofentanyl was tentatively identified in 301 samples, and all of the matching features eluted between 5.88 and 6.11 min, within the predicted retention time window of 5.25 min ±60 s. Moreover, fluorofentanyl identifications were characterized by a low mass error and demonstrated a high degree of MS/MS similarity to one another with a mean pairwise dot-product of 0.88. Similarly, bromazolam was tentatively identified in 145 samples, with all of the corresponding features eluting between 6.97 and 7.11 min, just outside of the predicted retention time window of 5.66 min ±60 s. Furanyl UF-17 was tentatively identified in 24 samples, with all of these identifications eluting within the predicted retention time window of 4.89 min ±60 s and all three diagnostic fragment ions matched in every sample.

Refinement of Tentative NPS Identifications by Epidemiological Data. To provide further context for these identifications, the proportion of samples in which

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

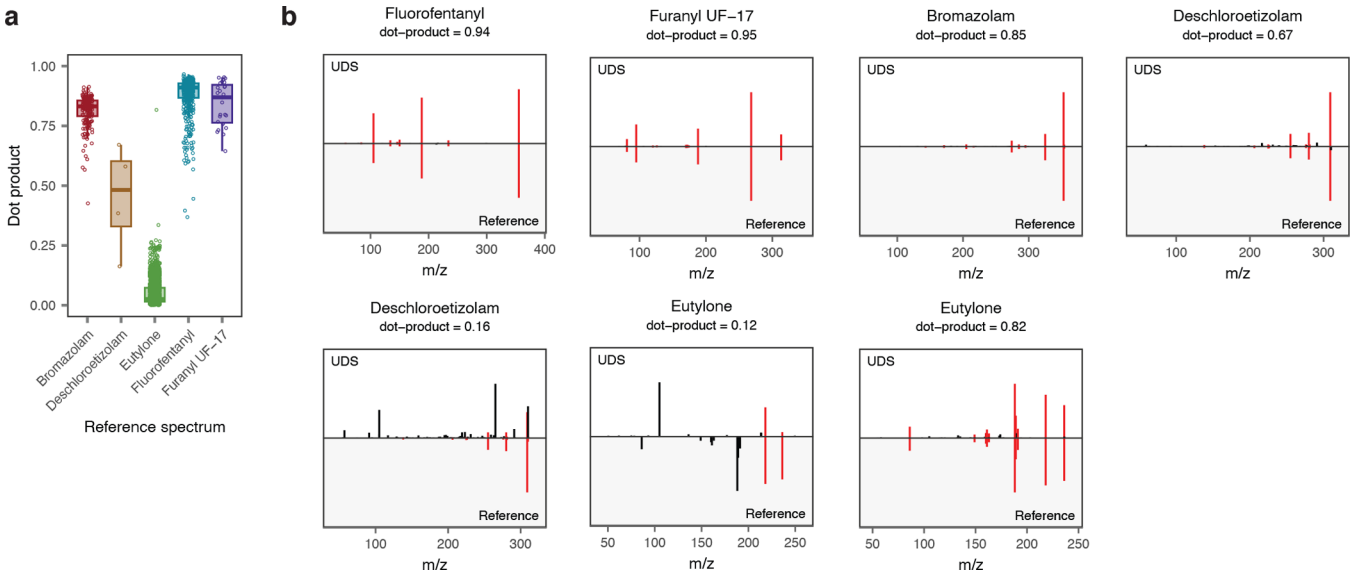


Figure 5. (a) Dot-products between reference spectra for five emerging NPSs and spectra tentatively identified in urine drug screen data. (b) Mirror plots showing examples of tentative NPS identifications (top) that were supported or rejected by similarity to reference spectra (bottom). For fluorofentanyl, furanyl UF-17, and bromazolam, representative spectra are shown. For deschloroetizolam, spectra with both high and low similarities are shown. For eutylone, a representative spectrum with low similarity to the reference spectrum is shown, in addition to an outlier spectrum with high similarity to the reference standard. Fragment ions present in both the experimental and reference spectra are shown in red. Fragment ion intensities are square-root-transformed to improve the visualization of low-intensity fragment ions.

each NPS was tentatively identified was calculated for each month over the 3 year study period (Figure 4). Interestingly, we found that NPSs with subjectively reliable identifications also tended to display clear temporal trends in their prevalence. For instance, we observed a dramatic increase in the number of tentative fluorofentanyl identifications toward the end of the study period in mid-2022. Similarly, we observed a biphasic increase in the number of tentative bromazolam identifications, with the first wave peaking in early 2021, followed by the second wave beginning in early 2022. Our data set also resolved the sudden appearances of 8-aminoclonazolam and metizolam in British Columbia in 2020, as well as a discrete peak in the prevalence of naphthyl-U-47700 in late-2021, plausibly reflecting a single introduction to the province. On the other hand, NPSs whose tentative identifications were felt to be less reliable, such as α-PiHP or N-ethylpentedrone, did not demonstrate such clear temporal trends. Eutylone presented a notable exception, with a clear peak in the proportion of tentatively positive samples in late 2020−early 2021, followed by a gradual decline. This observation raises the possibility that features tentatively identified as eutylone might instead represent structurally related drugs of abuse (e.g., other synthetic cathinones), as misidentifications of endogenous metabolites would not be expected to show such a clear temporal trend.

Validation of Tentative NPS Identifications with Reference Standards. Our results to this point support the notion that integration and manual inspection of mass spectral, chromatographic, and epidemiological data can help prioritize the acquisition of reference standards for NPSs that are not currently reported to clinicians. However, none of these data allow for the definitive identification of the NPSs that were tentatively identified by the presence of two or more fragment ions. To begin to address this gap, we acquired reference standards for five compounds, including eutylone, fluorofentanyl, bromazolam, furanyl UF-17, and deschloroetizolam (Table S2). For each sample in which these NPSs were tentatively identified, we compared the experimentally observed spectra to the reference spectrum for the relevant compound (Figure 5). This procedure provided supportive evidence for the tentative identifications of fluorofentanyl (median dot-product of 0.91), furanyl UF-17 (0.87), and bromazolam (0.83), albeit falling short of definitive identification. Of note, neither the observed MS/MS data nor our chromatographic method allowed us to differentiate between the *ortho*, *meta*, and *para* isomers of fluorofentanyl. We purchased reference materials for *para*-fluorofentanyl because this compound had previously been detected in postmortem cases across the United States by laboratories able to differentiate between fluorofentanyl isomers, whereas the *ortho* and *meta* isomers had not, but we underscore that the data presented here do not allow us to unambiguously identify the compound as such. Identifications of deschloroetizolam exhibited lower dot-products overall (median of 0.48), but manual inspection supported the reliability of these identifications with the exception of one misidentified spectrum. As expected, the vast majority of spectra tentatively identified as eutylone showed poor matches to the reference spectrum with a median dot-product of 0.03. However, a single spectrum demonstrated a convincing match to the reference standard, with a dot-product of 0.82 and a retention time of 4.15 min, compared to 4.08 min for the reference material.

Benchmarking the Sensitivity and Specificity of Retrospective Data Analysis. Together, the analyses described above support the feasibility of retrospectively reanalyzing population-scale mass spectrometry data to prioritize emerging NPSs for assay development. Nonetheless, we found that a substantial proportion of the NPSs that were tentatively identified with our permissive initial filter could be excluded as likely false-positives. We therefore sought to more systematically benchmark the sensitivity and specificity of the diagnostic fragment ion approach. To this end, we asked

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

whether identifications made based on diagnostic fragment ions could recapitulate adjudicated identifications made on the basis of authentic reference standards within the same set of clinical urine samples. We identified a total of 91 illicit drugs that were both reported to clinicians by the Provincial Toxicology Centre and which were present in the HighResNPS database,[10,15] a database of NPS fragment ions crowdsourced from dozens of laboratories around the world. For each of these 91 drugs, we repeated the analysis described above, using a threshold of two or more fragment ions, to identify each compound in the data from urine drug screens. We then computed the sensitivity and specificity with which each drug was identified using the identifications based on authentic reference standards as the ground truth.

Across all 91 drugs, identifications made on the basis of two or more diagnostic fragment ions achieved a median sensitivity of 45% and a median specificity of 97%, yielding a median accuracy of 97% per drug (Figures 6 and S4). The high specificity of this approach is a particularly desirable characteristic for a screening test in that this would allow laboratories to avoid unnecessary assay development by minimizing the number of false-positive identifications. The specificity would expectantly be further increased by the manual inspection of the results and removal of identifications that are subjectively deemed to be unreliable, as demonstrated above for NPSs such as eutylone and α-PiHP. Moreover, manual inspection of putative false-positives suggested that some of these may actually correspond to *bona fide* identifications that were not reported to clinicians at the time, as shown in Figure S5, for an identification of LSD that was supported by seven matching fragment ions. On the other hand, the sensitivity of suspect screening based on fragment ions was somewhat lower than one study had previously reported.[16] This discrepancy might reflect the differences between the two studies on how fragment ion data were used to make tentative identifications (a dot-product threshold optimized on the data set at hand in the study of Colby et al. versus a predefined number of matching fragment ions here) or the use of published MS/MS data collected with variable mass spectrometric setups in our analysis (as opposed to an in-house library in Colby et al.).

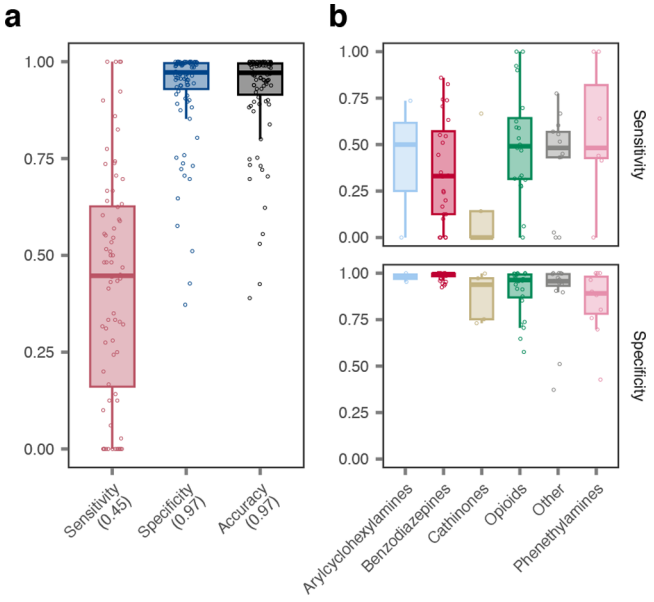


Figure 6. Evaluation of the retrospective screening approach for identifying clinically reported drugs using fragment ions from the HighResNPS database. (a) Test characteristics of the HighResNPS fragment ions for identifying 91 illicit drugs as compared to clinical reports from the same urine samples. *X*-axis text shows the median sensitivity, specificity, and accuracy of the approach across all 91 drugs. (b) Sensitivity and specificity of the HighResNPS fragment ions, shown separately for drugs of different EMCDDA categories. "Other" includes drugs annotated in HighResNPS as belonging to the following categories: piperidines and pyrrolidines, plants and extracts, indolalkylamines, or unknown.

DISCUSSION

The rapid pace at which NPSs are introduced and withdrawn from the illicit market makes it impractical for analytical laboratories to obtain reference materials for each of these compounds. As a result, these laboratories must make difficult and often subjective decisions about which reference materials to acquire. Our study aimed to explore whether retrospective suspect screening of population-scale mass spectrometry data could prioritize emerging NPSs for assay development in a more data-driven manner. To address this possibility, we conducted a reanalysis of mass spectrometry data from 12,727 urine drug screens performed over a 3 year period in one Canadian province. We compiled a suspect database including diagnostic fragment ions from public databases, monographs, and the scientific literature that were used to tentatively identify NPS, and we showed that these tentative identifications could be further refined by integrating mass spectral, chromatographic, and epidemiological data. The refined identifications were then used to direct the acquisition of new reference materials, which allowed us to provide further evidence for a number of tentative identifications. Moreover, these reference materials were used to develop new assays that were subsequently incorporated into routine workflows in our center. Finally, we systematically quantified the sensitivity and specificity of our approach by retrospectively comparing a crowd-sourced database of diagnostic fragment ions against historical adjudicated reports.

We emphasize that none of the NPS identifications presented in this paper would meet the standards for reporting to clinicians or authorities. Definitive identification in a reportable context requires matches to contemporaneously analyzed reference material, generally in the form of spiked-in reference standards.[30,31] That the tentative identifications described here do not meet this evidentiary threshold is consistent with the goal of our study, which was to facilitate operational decisions regarding the acquisition of new reference materials. Our approach represents a form of exploratory data analysis that is designed to identify trends at the scale of tens of thousands of samples and not definitively identify compounds in individual samples. To this end, we suggest that retrospective screening of emerging NPSs could be performed by individual laboratories on a regular basis (for example, every month) or on an *ad hoc* basis in response to alerts from clinicians or public health officials. If tentative identifications by fragment ion-based suspect screening raise the level of suspicion that a new NPS is circulating in the community, this would then motivate the purchase of reference materials for that NPS. However, re-extraction of historic samples and comparison against newly purchased materials would be required for definitive identification in those specific samples. We make all of our source code publicly available via GitHub to support retrospective screening efforts in other

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

Table 1. Suggested Criteria for the Prioritization of Reference Material Acquisition Based on Retrospective Data Analysis, with Examples of Supportive Vs Nonsupportive Data (and Compounds Meeting These Thresholds) under Each Criterion

Criterion	Definition	Supportive	Not supportive
Prevalence	Number of samples in which the compound was tentatively identified	✓ Detected in 10+ samples *Naphthyl-U-47700*	✗ Detected in <2 samples *MDMB-4en-BINACA*
Matched fragments	Number of fragments from biological samples matched to fragments from the suspect database	✓ 4+ fragments or >50% of fragments matched *8-aminoclonazolam*	✗ 2 fragments or <30% of fragments matched *N-ethyl-U-47700*
Fragment uniqueness	Number of fragments from the suspect database shared with other compounds in in-house libraries or HighResNPS	✓ 0-2 non-unique fragments *Flubromazepam*	✗ 3+ non-unique fragments *α-PiHP*
Retention time (RT)	Chromatographic coherence of tentative identifications and consistency with predicted or published RT data (if available)	✓ Single RT cluster, consistent with predicted or published RT *Furanyl UF-17*	✗ Two or more distinct chromatographic clusters *Eutylone*
Mass accuracy	Mass accuracy of precursor ion identification, in ppm	✓ Single cluster of errors centred around zero *Metizolam*	✗ Large mass error or multiple clusters *Benocyclidine*
Spectral coherence	Degree of (dis)similarity between tentatively identified spectra, as quantified by the dot-product	✓ Single cluster of dot-products with high median value *Fluorofentanyl*	✗ Multiple clusters of dot-products and/or low median value *Metonitazene*
Temporal trend	Frequency of tentative identifications over time	✓ Clear temporal trend *Bromazolam*	✗ No clear trend despite sufficient sample size *N-ethylpentedrone*
Clinical or forensic value	Potential of the drug in question to contribute to clinical presentations or public safety concerns	✓ High potential to cause death or public safety concern *Etodesnitazene*	✗ Less potential to cause death or public safety concern *5-methoxy MiPT*

laboratories, and we suggest a set of guidelines to prioritize the acquisition of reference materials on the basis of the data in Table 1.

Some families of NPSs are more readily detected in urine as their metabolites. Our approach allows for the tentative identification of these NPSs, as long as reference materials for their metabolites and their MS/MS spectra have been made available. For example, 8-aminoclonazolam was tentatively identified in 171 samples. This compound is a metabolite of clonazolam,[32] which was identified in just a single sample. These data would suggest that, at least in the context of urinary screening, the acquisition of reference materials for 8-aminoclonazolam should be prioritized over those for clonazolam itself. However, some challenges specific to screening for NPS metabolites should be noted: the metabolism of NPSs that have just emerged on the illicit market may be poorly understood, and metabolites of NPSs must be recognized as potential targets and included in the suspect database in order to enable their tentative identification.

The concept of retrospective data analysis to identify the emerging NPSs has previously been explored by a number of studies.[13,17−23] For example, Noble et al.[23] retrospectively screened 2,339 forensic blood samples for a set of 50 fentanyl analogues, using manually predicted fragment ions based on a proposed fragmentation pathway for fentanyl analogues to achieve identification without reference standards. Kriikku et al.[20] reanalyzed data from 1,836 forensic urine samples after acquiring reference material for the synthetic opioid U-47700, allowing them to identify two additional positive cases. Gundersen et al.[19] reanalyzed data from 1,314 forensic samples using accurate mass, retention time, and (in some cases) diagnostic fragment ion data from the HighResNPS database, although they dismissed the majority of identifications as probable false-positives. Axelsson et al.[17] researched data from 14,367 oral fluid samples against a database of reference spectra acquired for 48 NPSs. Pan et al.[22] reanalyzed data from 13,514 forensic blood samples to search for 47 designer benzodiazepines, using data from the HighResNPS database, and a training set of 13 common benzodiazepines to establish mass error, intensity, and retention time filters. However, many of these studies relied on data from newly acquired reference materials or focused on screening for a specific family of NPSs or even a single drug. Moreover, prior studies did not consider MS/MS data in retrospective analysis, did not propose automated approaches for data interpretation, or did not formally assess the accuracy of their approaches.

Our study builds on these important efforts in several directions. One key strength of our study is its scale: our retrospective analysis of 12,727 urine samples, which were

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

searched for 83 drugs, allowed us to achieve a comprehensive survey of emerging NPSs in one geographical area that has rarely been achieved.[17,22,33] Moreover, our study relied on published mass spectrometry data rather than newly acquired reference standards as the basis for our approach, and our analysis was not restricted to a single drug or class of NPSs. We developed an integrative strategy to identify likely false-positives and carefully characterized these through manual inspection. In particular, we show that combining information on matched fragment ions, mass accuracy, (predicted) retention time, and spectral similarity into interpretable visualizations (Figure 2) can greatly facilitate subjective judgments about which identifications are most likely to be reliable, an observation that we subsequently validated by acquiring new analytical reference materials. We benchmark the accuracy of retrospective suspect screening using diagnostic fragment ions from public databases, taking advantage of the large sample size afforded by three years of clinical reports. We show that the fragment ion approach is highly specific on a per-drug basis, supporting its use to prioritize emerging NPS for assay development. Finally, we have developed an efficient software implementation using open-source R packages that we expect will enable retrospective analyses of emerging NPSs on an ongoing basis.

Our approach also has a number of limitations. First, our approach does not permit de novo identification of entirely unknown substances: for an NPS to be tentatively identified, reference spectra must have been acquired in another laboratory. Second, and as noted above, our method has the potential to result in both false-positives and false-negatives when compared against gold-standard workflows based on authenticated reference materials. We searched only for $[M + H]^+$ adducts, and we searched primarily for NPS themselves (as opposed to their metabolites, which are more abundant in the urine for some families of NPSs). These are both factors that may have contributed to false-negatives. Moreover, fragmentation patterns can differ across instruments and configurations, which could also have caused false-negatives. Our full-scan mass spectrometry method used data-dependent acquisition with an inclusion list, meaning that NPSs represented by low-intensity precursor ions may not have been selected for fragmentation, particularly if they coeluted with ions in the inclusion list. Together, these factors likely explain at least some fraction of the false-negatives observed in our HighResNPS benchmarking experiment and underscore the principle that failure to detect an NPS using suspect screening approaches does not necessarily indicate that it was not present in the sample of interest.

On the other hand, we found that our approach could also lead to false-positive identifications. Filtering tentatively identified NPSs to remove compounds that do not match published retention time data, do not exhibit coherent retention times, or have mutually dissimilar MS/MS spectra can help increase the confidence in the remaining identifications, although these filters cannot rule out false-positives entirely. In some cases, two NPSs may be so similar that only separation on a chromatographic column with reference standards for both compounds would allow them to be differentiated (e.g., α-PiHP and α-PHP). Notwithstanding these limitations, our data suggest that the majority of false-positives can be subjectively identified using the integrative strategies we propose here and subsequently removed. This process does require manual review of tentative identifications by an expert analyst, but we have found that this can be accomplished within a reasonable amount of time (on the order of a few hours for the number of NPS samples analyzed in the present study). Moreover, we emphasize that our approach is not intended to replace standard workflows based on comparisons to authenticated reference standards. Instead, our goal is to contribute useful information toward data-driven decision-making in analytical laboratories, and we highlight that a platform does not need to achieve perfect sensitivity or specificity in order to contribute useful information toward this objective.

We benchmarked the performance of retrospective suspect screening using diagnostic fragment ions from public databases (as opposed to data acquired in-house from reference materials). Our aim in this analysis was not to identify illicit drugs but rather to estimate the sensitivity and specificity of our approach by taking advantage of a large resource of historical adjudicated reports. We observed that a criterion of two or more fragment ions for tentative identification yielded a median per-drug sensitivity of 45% and specificity of 97%. However, this benchmark also has limitations. Several compounds were included in the historical adjudicated reports only if they were quantified above a given limit of reporting, meaning there are likely cases in which these compounds were present in the samples but did not meet this threshold and therefore were not reported. The requirement for detection above a quantitative threshold likely explains the much lower specificity we observed for drugs such as cocaine, benzoylecgonine, morphine, or methamphetamine and suggests that our benchmark provides a relatively pessimistic estimate for these compounds. On the other hand, the estimated specificity of our approach is dictated in part by the search space of our benchmark, which encompassed only 91 illicit drugs. The specificity would be expected to decline with the addition of more drugs to this search space, particularly drugs with high spectral similarities to one another.

Our study opens a number of interesting directions for future work. The availability of computational methods to predict compound fragmentation in silico[34−36] raises the possibility of searching for computationally *predicted* fragment ions, rather than experimentally observed ones. This could conceivably enable retrospective data analysis before mass spectrometry data are even available for compounds that have just emerged on the illicit market. Moreover, the potential of anticipating the chemical structures of NPS that are most likely to emerge on the market next using chemical language models[37,38] further raises the possibility of combining MS/MS spectrum prediction with in silico structure generation in order to search for as-of-yet unknown drugs. More immediately, deploying the approach described here on a regular basis would provide a platform for continuous retrospective surveillance that would allow clinicians or scientists to rapidly ask whether a given NPS is being distributed on the illicit market locally immediately after its detection by authorities elsewhere in the world. Efforts are now underway to this end, with a view toward routinely deploying this system for surveillance of all clinical and forensic samples screened at the British Columbia Provincial Toxicology Centre.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.3c03451.

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

Additional materials and methods showing the composition of the suspect database, the performance of the retention time prediction method, and the sensitivity and specificity of the retrospective suspect screening approach (PDF)
Suspect database of 83 NPS (XLSX)
Mass spectrometry data acquired from reference standards (XLSX)

AUTHOR INFORMATION

Corresponding Author

Michael A. Skinnider − *Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada; Lewis-Sigler Institute of Integrative Genomics and Ludwig Institute for Cancer Research, Princeton University, Princeton, New Jersey 08544, United States;* orcid.org/0000-0002-2168-1621; Email: skinnider@princeton.edu

Authors

Sandrine A. M. Mérette − *Provincial Toxicology Centre, Provincial Health Services Authority, Vancouver, British Columbia V5Z 4R4, Canada*

Daniel Pasin − *Forensic Laboratory Division, Office of the Chief Medical Examiner, San Francisco, California 94124, United States*

Jason Rogalski − *Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada*

Leonard J. Foster − *Michael Smith Laboratories, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada; Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver, British Columbia V6T 1Z3, Canada;* orcid.org/0000-0001-8551-4817

Frank Scheuermeyer − *Department of Emergency Medicine, St. Paul's Hospital and the University of British Columbia, Vancouver, British Columbia V6Z IY6, Canada; Centre for Health Evaluation and Outcome Sciences, St. Paul's Hospital, Vancouver, British Columbia V6Z IY6, Canada*

Aaron M. Shapiro − *Provincial Toxicology Centre, Provincial Health Services Authority, Vancouver, British Columbia V5Z 4R4, Canada; Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia V6T 2B5, Canada;* orcid.org/0000-0002-9920-1763

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.3c03451

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was enabled in part by computational resources provided by WestGrid, Compute Canada, and Advanced Research Computing at the University of British Columbia.

REFERENCES

(1) Peacock, A.; Bruno, R.; Gisev, N.; Degenhardt, L.; Hall, W.; Sedefov, R.; White, J.; Thomas, K. V.; Farrell, M.; Griffiths, P. *Lancet* **2019**, *394* (10209), 1668−1684.
(2) Baumann, M. H.; Solis, E.; Watterson, L. R.; Marusich, J. A.; Fantegrossi, W. E.; Wiley, J. L. *J. Neurosci.* **2014**, *34* (46), 15150−15158.
(3) Underwood, E. *Science* **2015**, *347* (6221), 469−473.
(4) Madras, B. K. *Curr. Top. Behav. Neurosci.* **2016**, *32*, 1−18.
(5) Kronstrand, R.; Guerrieri, D.; Vikingsson, S.; Wohlfarth, A.; Gréen, H. *Handb. Exp. Pharmacol.* **2018**, *252*, 495−541.
(6) Baumann, M. H.; Volkow, N. D. *Neuropsychopharmacology* **2016**, *41* (3), 663−665.
(7) Pasin, D.; Cawley, A.; Bidny, S.; Fu, S. *Anal. Bioanal. Chem.* **2017**, *409* (25), 5821−5836.
(8) Fabregat-Safont, D.; Sancho, J. V.; Hernández, F.; Ibáñez, M. *J. Mass Spectrom.* **2021**, *56* (7), No. e4673.
(9) Wohlfarth, A.; Weinmann, W. *Bioanalysis* **2010**, *2* (5), 965−979.
(10) Davidsen, A.; Mardal, M.; Linnet, K.; Dalsgaard, P. W. *PLoS One* **2020**, *15* (11), No. e0242224.
(11) Gebel Berg, E. *ACS Cent. Sci.* **2016**, *2* (6), 363−366.
(12) Ifeagwu, S. C.; Raithelhuber, M.; Crean, C.; Gerostamoulos, D.; Chung, H.; Tettey, J. N. *Forensic Sci. Int.* **2017**, *274*, 2−6.
(13) Partridge, E.; Trobbiani, S.; Stockham, P.; Charlwood, C.; Kostakis, C. *J. Anal. Toxicol.* **2018**, *42* (9), 655−660.
(14) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Nishioka, T. *J. Mass Spectrom.* **2010**, *45* (7), 703−714.
(15) Mardal, M.; Andreasen, M. F.; Møllerup, C. B.; Stockham, P.; Telving, R.; Thomaidis, N. S.; Diamanti, K. S.; Linnet, K.; Dalsgaard, P. W. *J. Anal. Toxicol.* **2019**, *43* (7), 520−527.
(16) Colby, J. M.; Thoren, K. L.; Lynch, K. L. *J. Anal. Toxicol.* **2018**, *42* (4), 207−213.
(17) Axelsson, M. A. B.; Lövgren, H.; Kronstrand, R.; Green, H.; Bergström, M. A. *Basic Clin. Pharmacol. Toxicol.* **2022**, *131* (5), 420−434.
(18) Campos-Mañas, M. C.; Ferrer, I.; Thurman, E. M.; Sánchez Pérez, J. A.; Agüera, A. *Sci. Total Environ.* **2019**, *664*, 874−884.
(19) Gundersen, P. O. M.; Broecker, S.; Slørdal, L.; Spigset, O.; Josefsson, M. *Forensic Sci. Int.* **2020**, *311*, No. 110274.
(20) Kriikku, P.; Pelander, A.; Rasanen, I.; Ojanperä, I. *Forensic Sci. Int.* **2019**, *300*, 85−88.
(21) Krotulski, A. J.; Varnum, S. J.; Logan, B. K. *J. Forensic Sci.* **2020**, *65* (2), 550−562.
(22) Pan, M.; Rasmussen, B. S.; Dalsgaard, P. W.; M?llerup, C. B.; Nielsen, M. K. K.; Nedahl, M.; Linnet, K.; Mardal, M. *Front. Chem.* **2022**, *10*, 868532.
(23) Noble, C.; Weihe Dalsgaard, P.; Stybe Johansen, S.; Linnet, K. *Drug Test. Anal.* **2018**, *10* (4), 651−662.
(24) Pasin, D.; Møllerup, C. B.; Rasmussen, B. S.; Linnet, K.; Dalsgaard, P. W. *Anal. Chim. Acta* **2021**, *1184*, No. 339035.
(25) Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Mallick, P. *Nat. Biotechnol.* **2012**, *30* (10), 918−920.
(26) Smith, C. A.; Want, E. J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. *Anal. Chem.* **2006**, *78* (3), 779−787.
(27) Tautenhahn, R.; Böttcher, C.; Neumann, S. *BMC Bioinf.* **2008**, *9*, 504.
(28) Rainer, J.; Vicini, A.; Salzer, L.; Stanstrup, J.; Badia, J. M.; Neumann, S.; Stravs, M. A.; Verri Hernandes, V.; Gatto, L.; Gibb, S.; Witting, M. *Metabolites* **2022**, *12*, 2.
(29) Møllerup, C. B.; Dalsgaard, P. W.; Mardal, M.; Linnet, K. *Drug Test. Anal.* **2017**, *9* (7), 1052−1061.
(30) Rivier, L. *Anal. Chim. Acta* **2003**, *492* (1−2), 69−82.
(31) Partridge, E.; Trobbiani, S.; Stockham, P.; Scott, T.; Kostakis, C. *J. Anal. Toxicol.* **2018**, *42* (4), 220−231.
(32) Meyer, M. R.; Bergstrand, M. P.; Helander, A.; Beck, O. *Anal. Bioanal. Chem.* **2016**, *408* (13), 3571−3591.
(33) Mardal, M.; Dalsgaard, P. W.; Rasmussen, B. S.; Linnet, K.; Møllerup, C. B. *Anal. Chem.* **2023**, *95*, 4592−4596.
(34) Wang, F.; Liigand, J.; Tian, S.; Arndt, D.; Greiner, R.; Wishart, D. S. *Anal. Chem.* 2021.

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX

(35) Ruttkies, C.; Schymanski, E. L.; Wolf, S.; Hollender, J.; Neumann, S. *J. Cheminformatics* **2016**, *8*, 3.
(36) Allen, F.; Pon, A.; Wilson, M.; Greiner, R.; Wishart, D. *Nucleic Acids Res.* **2014**, *42* (Web Server issue), W94−9.
(37) Skinnider, M. A.; Stacey, R. G.; Wishart, D. S.; Foster, L. J. *Nat. Mach. Intell.* **2021**, *3* (9), 759−770.
(38) Skinnider, M. A.; Wang, F.; Pasin, D.; Greiner, R.; Foster, L. J.; Dalsgaard, P. W.; Wishart, D. S. *Nat. Mach. Intell.* **2021**, *3* (11), 973−984.

https://doi.org/10.1021/acs.analchem.3c03451
Anal. Chem. XXXX, XXX, XXX−XXX