On the challenge of unambiguous identification of fentanyl analogs: exploring measurement diversity using standard reference mass spectral libraries

AS Moorthy[1], EP Erisman[1], AJ Kearsley[2], Y Liang[1], E Sisco[3], WE Wallace[1]

[1]Mass Spectrometry Data Center, Biomolecular Measurement Division, National Institute of Standards and Technology
[2]Mathematical Analysis and Modeling Group, Applied and Computational Mathematics Division, National Institute of Standards and Technology
[3]Surface and Trace Chemical Analysis Group, Materials Measurement Science Division, National Institute of Standards and Technology

Abstract

Fentanyl analogs are a class of designer drugs that are particularly challenging to unambiguously identify due to the mass spectral and retention time similarities of unique compounds. In this paper, we use agglomerative hierarchical clustering to explore the measurement diversity of fentanyl analogs and better understand the challenge of unambiguous identifications using analytical techniques traditionally available to drug chemists. We consider four measurements in particular: gas chromatography retention indices, electron ionization mass spectra, electrospray ionization tandem mass spectra and direct analysis in real time mass spectra. Our analysis demonstrates how simultaneously considering data from multiple measurement techniques increases the observable measurement diversity of fentanyl analogs, which can reduce identification ambiguity. This paper further supports the use of multiple analytical techniques to identify fentanyl analogs (among other substances), as is recommended by the Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG).

Keywords: direct analysis in real time mass spectrometry (DART-MS); electron ionization mass spectrometry (EI-MS); electrospray ionization tandem mass spectrometry (ESI-MS/MS); fentanyl analogs; gas chromatography retention indices (RI); mass spectral libraries.

1. Introduction

Fentanyl analogs have been some of the more problematic drugs of abuse this past decade [1]. Historically, authorities scheduled fentanyl analogs individually, which allowed for slightly modified and unscheduled variants to be distributed faster than they could be controlled [2]. While "blanket" scheduling of fentanyl analogs by governing bodies like the United States Drug Enforcement Administration have helped limit the rise of new modified versions [3], their high potency has allowed them to permeate the drug supply chain as components in increasingly complex samples [4], complicating both law enforcement and public health responses [1,5,6].

Underlying these societal challenges is a fundamental measurement issue—fentanyl analogs are difficult to unambiguously discriminate with analytical techniques typically found in forensic laboratories. For example, consider the near-identical centroided mass spectra of cyclopropyl fentanyl and crotonyl fentanyl, measured with electron ionization mass spectrometry (EI-MS), shown in Figure **1.** Roughly speaking, in EI-MS, an analyte interacts with high energy electrons and forms *ions*. The relative abundance of these ions is then reported as a function of ion *mass-to-charge ratio (m/z)* in a data structure referred to as a mass spectrum. Because the *m/z* values depend only on the molecular constitution (*viz*., mass) of the observed ions, the mass spectra of isomeric compounds are often very similar. Spectral differences due to molecular connectivity will be reflected through changes in signal intensity that may be subtle or even indistinguishable (see Figure **1)**.
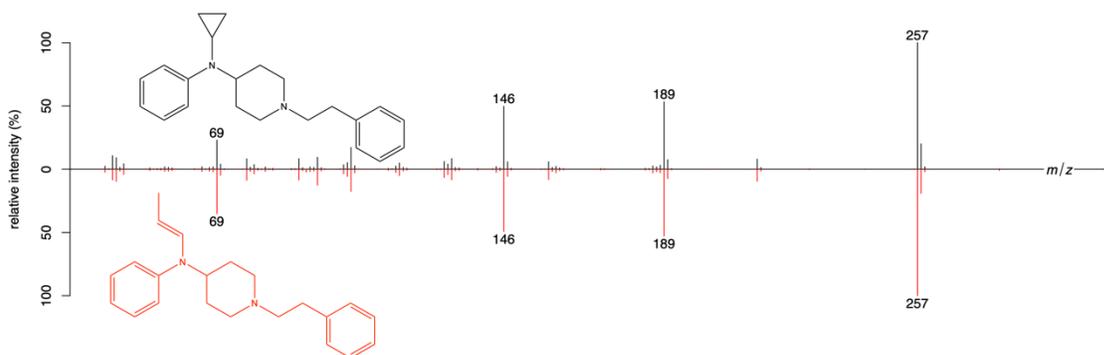
Figure 1: Head-to-tail display of cyclopropyl fentanyl (top/black) and crotonyl fentanyl (bottom/red) mass spectra with structures overlaid. These mass spectra were obtained from the SWGDRUG Mass Spectral Library (version 3.11) [7].

Due to difficulties with EI-MS, it is commonplace to leverage additional measurements, like chromatographic retention times, or other technologies altogether to help discriminate structurally similar compounds. Examples of these measurements include a bespoke gas chromatography (GC) mass spectrometry method for discriminating synthetic opioids [8], and ultra-high performance liquid chromatography tandem mass spectrometry method for discriminating between cyclopropyl fentanyl and crotonyl fentanyl in toxicology applications [9]. The Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) currently recommends reporting results from multiple techniques when identifying drugs or substances [10].

While the solution of using multiple measurement techniques for identifying compounds is conceptually straight-forward, it is difficult to know how many techniques—and which techniques in particular—are necessary to accurately identify specific compounds without first characterizing how these measurements differ across similar compounds. Therefore,

understanding the *measurement diversity* of fentanyl analogs will help us select measurement techniques that minimize identification ambiguity.

In this paper, we define the concept of measurement diversity using agglomerative hierarchical clustering, and characterize the measurement diversity of fentanyl analogs across four common analytical measurement strategies: (a) GC retention indices, (b) full scan mass spectra collected using electron ionization mass spectrometry (EI-MS), (c) precursor *m/z* and product ion mass spectra (MS^2) collected at multiple collision energies using electrospray ionization tandem mass spectrometry (ESI-MS/MS), and (d) in-source collision induced dissociation (is-CID) full scan mass spectra collected at three orifice 1 energies using direct analysis in real time mass spectrometry (DART-MS). Examples of each measurement type using fentanyl are presented in Figure **2**. We also discuss how these four measurements can be combined to increase measurement diversity and subsequently improve our ability to unambiguously identify compounds.
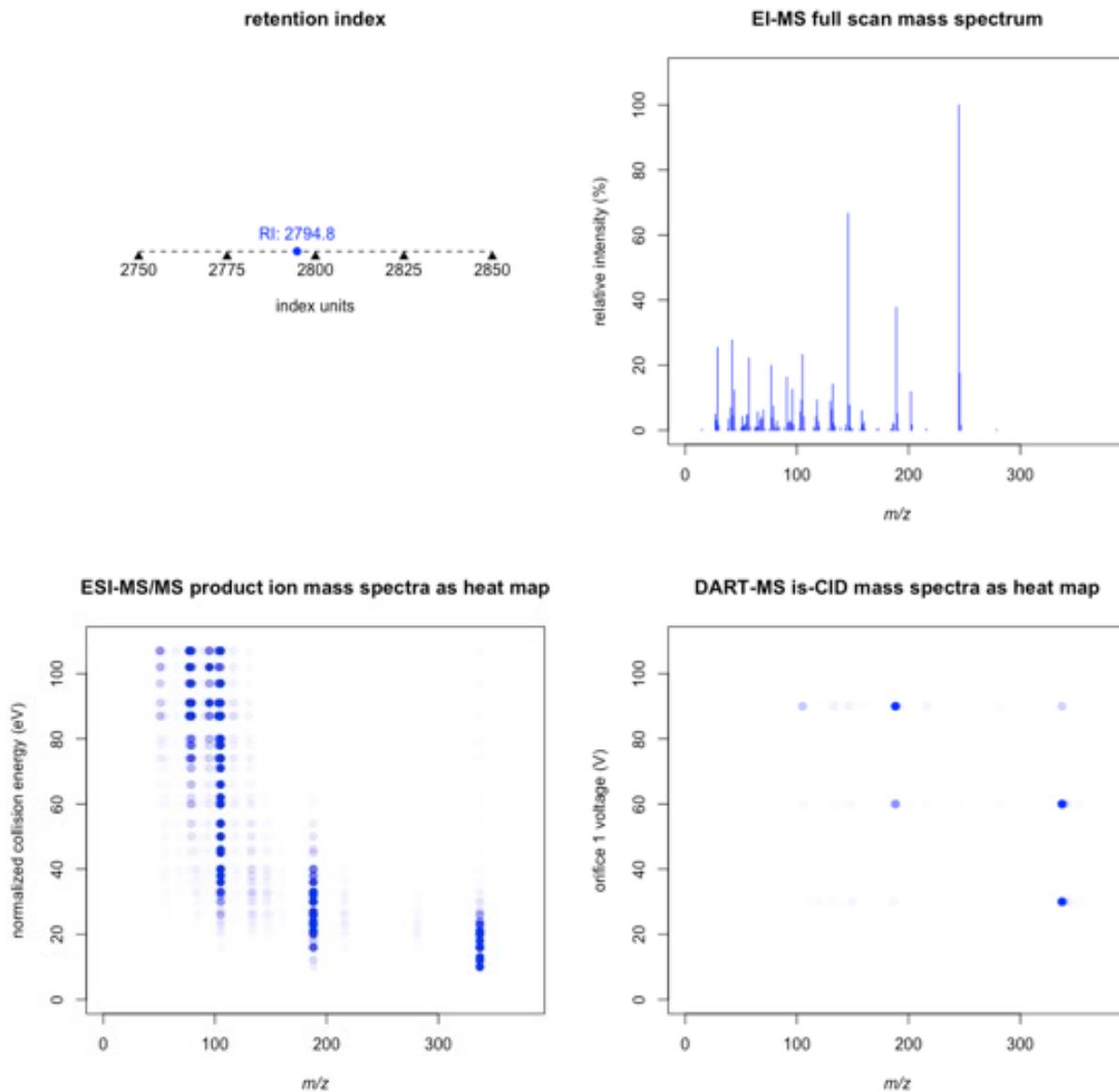
Figure 2: Example measurements of fentanyl: (**a**) retention index from NIST 23 gas chromatography methods library, (**b**) EI mass spectrum from NIST 23 EI-MS library, (**c**) set of product ion mass spectra (MS^2) measured at multiple normalized collision energies from NIST 23 MS/MS library, and (**d**) set of is-CID mass spectra measured at multiple orifice 1 energies from NIST DART-MS Forensic Database (version 7 - Grasshopper). For heat maps (panels **c** and **d**), point opacity represents relative peak intensity in the underlying mass spectra.

2. Theory and Methods

In the following sections, we first describe the mathematical foundations of clustering used to define our notion of *measurement diversity* (sections 2.1). Next, dissimilarity is computed for each measurement types (section 2.2) followed by a description of the specific data collection and analysis procedure (section 2.3) used to generate the results presented in Section 3.

2.1 Defining measurement diversity

Agglomerative hierarchical clustering (AHC) is an approach for constructing a tree-structure (dendrogram) that describes a hierarchical relationship between a set of objects [11,12]. This relationship is usually based on a mathematically defined measure of dissimilarity or distance between objects – an AHC algorithm receives as input a square and symmetric *dissimilarity matrix* that summarizes the pairwise dissimilarity between all objects in the set.

In an AHC analysis, we begin by assuming every object belongs to its own individual cluster (i.e., if there are n objects in a set, we begin with n clusters all of size 1). In each step, the least dissimilar (or most similar) clusters—which will initially be the least dissimilar individual objects—are merged to create a new cluster. This procedure continues until all objects in the set belong to a single cluster; the sequence by which clusters are merged and the dissimilarities between merged clusters are tracked throughout the process. An excellent overview of AHC algorithms and specific implementation considerations can be found in [13].

With a dendrogram, we can identify dissimilarity levels at which we can "cut" the dendrogram and evaluate the resulting clusters; the results can be interpreted through a variety of metrics. In this paper, we define a simple metric referred to as the *measurement diversity index*

$$D = \frac{N(\tau)}{n}, \tag{1}$$

where $N(\tau)$ is the number of clusters identified from a dendrogram at a specified dissimilarity level τ, and n is the number of objects in the set. If the clustering results have a low measurement diversity ($N \ll n$), it will be difficult to accurately identify individual objects. As measurement diversity increases ($N \rightarrow n$), individual objects are easier discriminate.

2.2 *Calculating dissimilarity between measurements and clusters*

As noted previously, the input to an AHC algorithm is a dissimilarity matrix that summarizes the pairwise dissimilarity between objects in a set. Thus, the first step to performing an AHC analysis is selecting a *pairwise dissimilarity* measure to describe the relationship between objects. In this paper, we are interested in discriminating fentanyl analogs by four analytical measurements.

For retention indices, the dissimilarity can be calculated using an absolute difference,

$$\phi_1(r_1, r_2) = |r_1 - r_2|, \tag{2}$$

where r_1 and r_2 are two retention indices.

The similarity between any two EI full scan mass spectra can be measured using a variety of techniques [14–22]; one long-standing approach for approximating similarity is referred to as the dot-product, or cosine similarity. We denote the cosine similarity between mass spectra x and y, measured with a low-resolution mass spectrometer and *m/z* tolerance $\epsilon_{lr} = 0$, as $\theta(x, y, \epsilon_{lr})$, and approximate the dissimilarity between any mass spectra as $\phi_2(x, y, \epsilon_{lr}) = 1 - \theta(x, y, \epsilon_{lr})$.

Specifications for computing cosine similarity with *m/z* tolerance as a parameter is provided in the Supplemental Information.

We use a two-stage approach to characterize dissimilarity between fentanyl analogs using high-resolution ESI-MS/MS measurements. We first consider the absolute difference between their precursor *m/z* values, $\Delta_{ms1} = |M_1 - M_2|$, where M_i is the precursor *m/z* of analogs 1 and 2, respectively. If $\Delta_{ms1} > \epsilon_{hr} = 0.005$, we set the dissimilarity between the analogs as 1 (i.e., the maximum possible dissimilarity). If $\Delta_{ms1} \leq \epsilon_{hr} = 0.005$, and the two analogs have at least one pair of MS^2 mass spectra that were collected at the same collision energy, we compute dissimilarity as

$$\phi_3(\boldsymbol{x}, \boldsymbol{y}) = 1 - \frac{1}{E}\sum_i^E \theta(\boldsymbol{x}_i, \boldsymbol{y}_i, \epsilon_{hr}),\qquad(3)$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are sets of MS^2 mass spectra collected at $E > 0$ different collision energies, $\boldsymbol{x}_i$ and $\boldsymbol{y}_i$ are the specific mass spectra collected at the same ith collision energy, $\epsilon_{hr} = 0.005$ is the *m/z* tolerance with a high resolution mass spectrometer, and the pairwise spectral similarity function θ is the same as employed with EI-MS mass spectra. If $\Delta_{ms1} \leq \epsilon_{hr} = 0.005$, and the compared compounds do not share any MS^2 spectra at the same collision energy, then $\phi_3(\boldsymbol{x}, \boldsymbol{y}) = 1$.

DART-MS is an ambient ionization mass spectrometry technique that is generally not preceded by a chromatography step. Accordingly, real-world mass spectra collected with DART-MS often contain signature ions originating from more than one compound, and spectral similarity or dissimilarity between an unknown mass spectrum and a reference spectrum of a pure compound is approximated using partial pattern matching approaches [23–25]. In this study, we compare pure

standard library spectra to each other, which allows us to use cosine similarity (or full pattern matching) in a manner similar to Equation (3). With the DART-MS mass spectra, $E = 3$ represents all compounds with mass spectra measured at the exact same is-CID energies.

The second step in an AHC analysis is specifying how dissimilarity is computed between clusters, and this is independent of the method selected for measuring pairwise dissimilarities between objects (described previously). As described in [13], there are several "linkage" methods for describing the dissimilarity between clusters. For example, we can approximate the dissimilarity between two clusters based on the average dissimilarity between all objects in clusters, commonly referred to as a Group Average or Unweighted Pair Group Method with Arithmetic mean (UPGMA) [26]. In this study, we consider a "complete link" method which means the dissimilarity between clusters is approximated by the maximum dissimilarity between objects in the clusters, thus we guarantee that the maximum dissimilarity between any two objects in a cluster created at a specified dissimilarity cutoff level will always be bounded by the cutoff level value itself.

2.3 Data and Analysis Details

The data used in this study was extracted from various NIST databases [27]. The EI mass spectra, ESI MS^2 mass spectra, and RI values were selected from pre-release versions of the NIST 23 EI-MS, ESI-MS/MS, and GC Methods libraries, respectively. DART-MS mass spectra were selected from the most recent release of the NIST DART-MS Forensic Database (version 7, Grasshopper) [28,29]. Articles describing how these measurements are recorded and how libraries/databases are constructed/evaluated can be found in the literature [30–33]. We only considered fentanyl analogs

for which we had all four measurements in this study; details about the complete set of 197 fentanyl analogs are in Table S1.

Data analysis was conducted using a custom script prepared in the R programming language [34]; the underlying source code is available for review by contacting the corresponding author. A schematic overview of the analysis steps is provided as Figure 3.
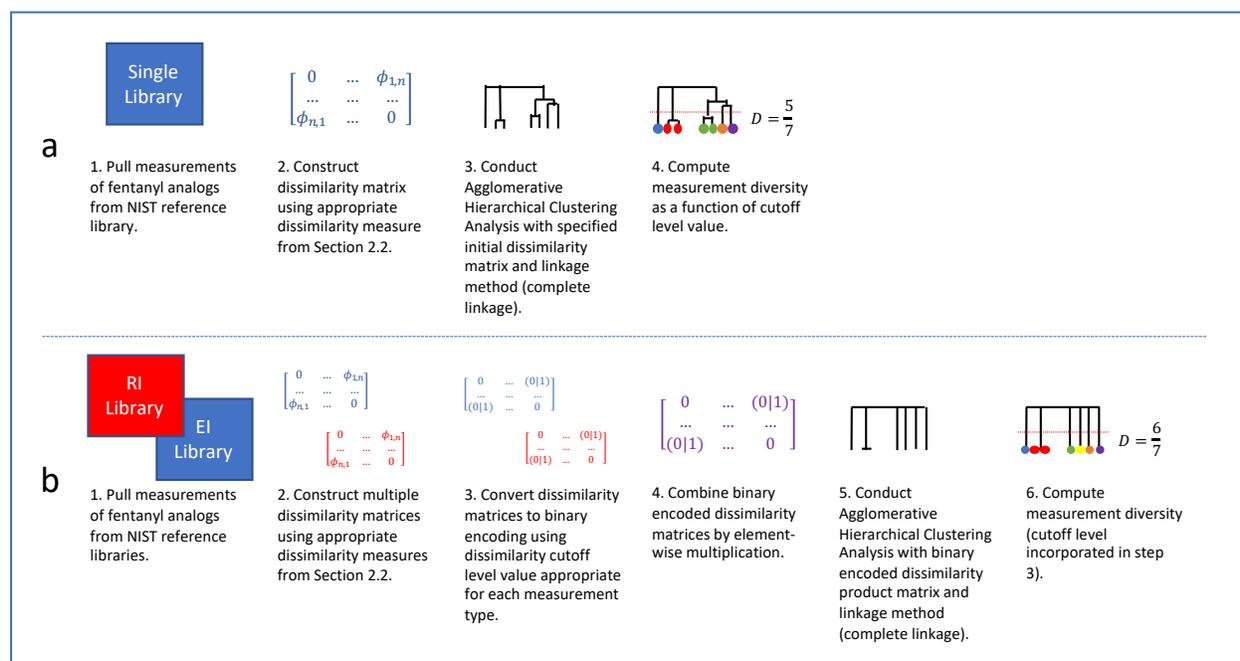
Figure 3: Schematic overview of measurement diversity calculations when using (a) reference data for a single measurement technique and (b) reference data from multiple measurement techniques.

To compute the measurement diversity using a single library, we followed the steps outlined in Figure 3a. For clustering, we used the `hclust` package available in base R with a calculated dissimilarity matrix and "complete" linkage method. Diversity calculations were done with a variety of cutoff values depending on the measurement type being considered. To compute combined diversity indices using multiple measurements, we followed the steps in Figure 3b. In this process, we need to convert traditional dissimilarity matrices into binary encoded dissimilarity

matrices based on values in the original matrices. More details on these calculations are provided during the discussion of multiple measurement comparisons in Section 3.

## 3. Results and Discussion

For each of the 197 fentanyl analogs investigated, we performed four AHC analyses using the four measurement types discussed previously. Figure **4** shows parts of the dendrogram created using retention indices as the discriminating measurement, with an overlayed magnification of 24 analogs and a table discussing three analogs (norsufentanil, N-methyl cyclopropyl norfentanyl, and cyclopropyl norfentanyl). From this, we can observe several groups consisting of two to four compounds with retention indices within 10 arbitrary units (a.u.) (e.g., cyclopropyl norfentanyl and N-methyl cyclopropyl norfentanyl) and a few compounds that would be uniquely identifiable with a 10 a.u. cutoff level for retention indices (e.g., Norsufentanil). Analogous dendrograms can be generated for EI-MS dissimilarity, ESI-MS/MS MS^2 dissimilarity and DART-MS is-CID mass spectral dissimilarity (figures not shown).

| Underlying Information for selected compounds		
Study ID	Name	R.I.
197	Norsufentanil	2073.9
158	N-Methyl cyclopropyl norfentanyl	2087.9
176	Cyclopropyl norfentanyl	2089.4
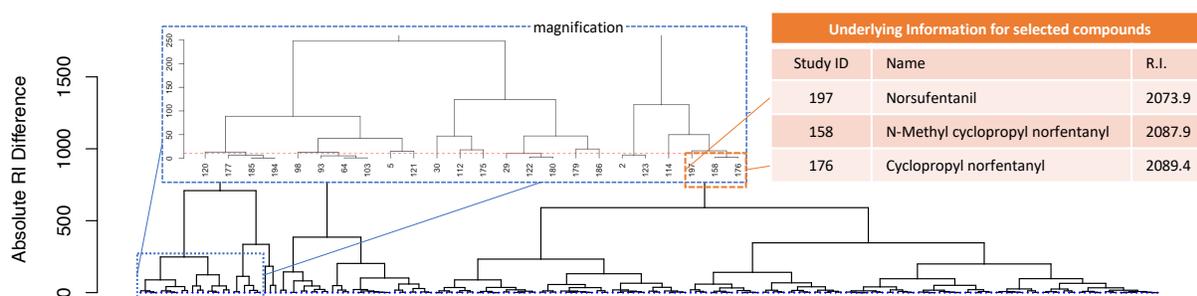
Figure 4: Dendrogram created by agglomerative hierarchical clustering of RI measurements of 197 fentanyl analogs in the NIST GC Methods library. Overlayed is a magnification of 24 of the compounds and further details about three of the compounds.

We studied how the measurement diversity changed for each measurement as a function of dissimilarity cutoff level (Figure **5**). The diversity of retention indices was never a perfect 1, even with a dissimilarity level of 0 (Figure **5a**)—there are a pair of fentanyl analogs (β-Methylfentanyl and p-Fluoro acryl fentanyl) that happen to have identical retention indices in the NIST GC-Methods library. The diversity of mass spectral measurements (Figures **5b-d**) is 1 with a dissimilarity cutoff value of 0 As expected, the diversity of all measurements decreased with increasing cutoff values.
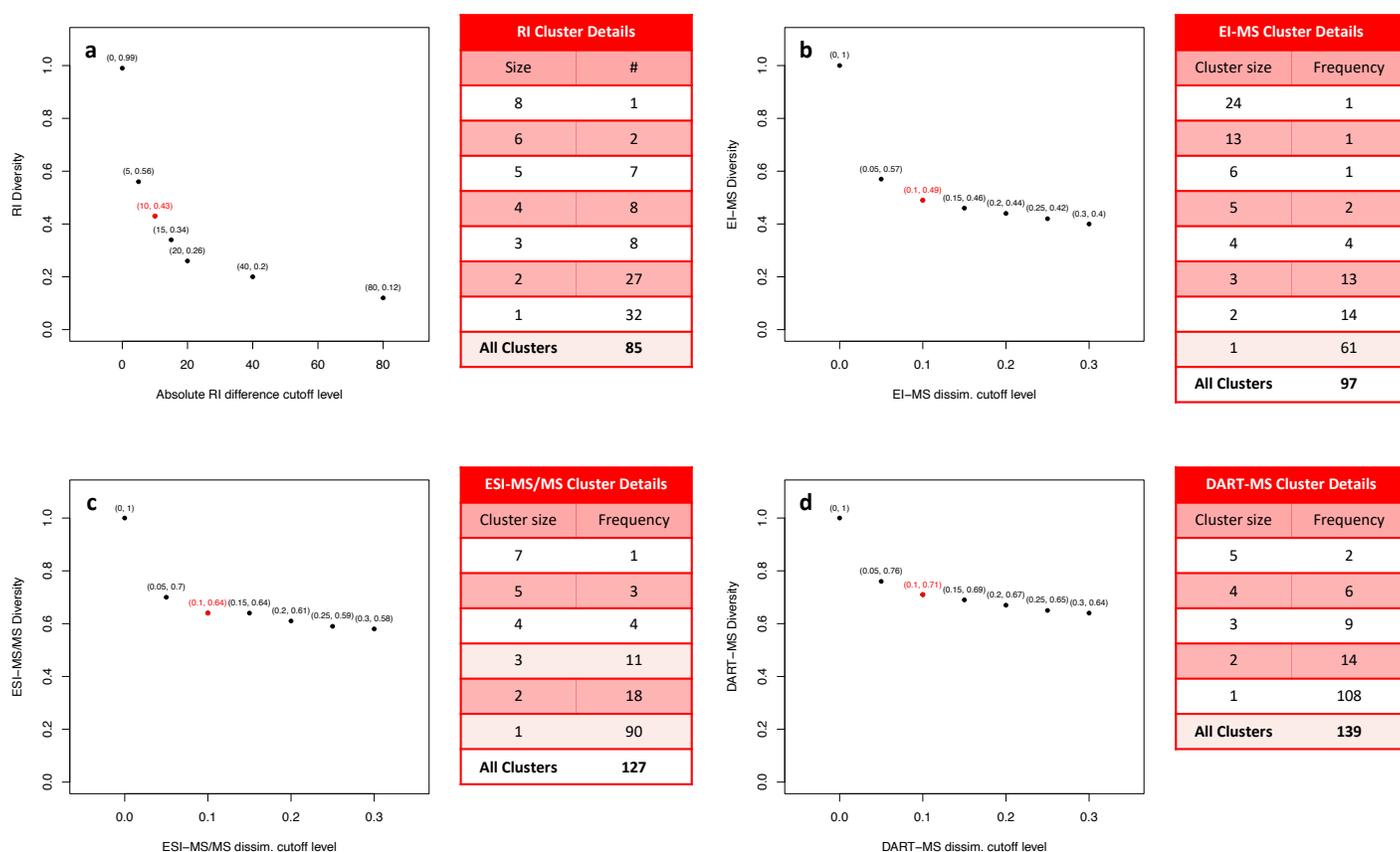
Figure 5: Measurement diversity across several dissimilarity thresholds for various measurement techniques (a) gas chromatography retention indices, (b) electron ionization full scan mass spectra, (c) electrospray ionization product ion mass spectra and (d) direct analysis in real time mass spectrometry in source collision induced dissociation mass spectra. Each figure is accompanied with a table describing the detailed clustering breakdown for analysis using the highlighted (red) dissimilarity cutoff value, where frequency and cluster size indicate how often (frequency) clusters of the specified size occur in the cluster results.

In practical application, repeat RI and mass spectral measurements are only roughly reproducible and thus, will always have dissimilarity values greater than 0 a.u. In particular, we might expect retention indices of the same molecule to differ by upwards of 10 a.u. and for the replicate mass spectra to have similarity values of approximately 0.9 a.u. (or dissimilarity values of 0.1 a.u.) even when measured under identical conditions. With measurements collected under nonidentical conditions (e.g., when building a reference library), one might observe even larger discrepancies between measurements of the same compound. Using a dissimilarity threshold of 10 a.u. for retention index and 0.1 a.u. for mass spectral measurements (red points in Figures 5a-d), we see that the measurement diversity is highest with DART-MS and ESI-MS/MS measurements, followed by EI-MS and RI measurements. This is expected since both DART-MS and ESI-MS/MS measurements contain information across multiple collision energies; lower energy spectra are dominated by peaks that provide insights about the intact molecule (e.g., molecular weight) while higher energy spectra generally have more peaks allowing us to infer potential fragmentation information. If only using single DART-MS measurements collected at a single low energy value, as is common in many forensic applications [35], measurement diversity drops to 0.5 (Figure S1). With EI-MS, mass spectra are collected at 70 eV and can contain little information about the intact molecular ion for analytes with strongly labile bonds like fentanyl analogs. Accordingly, several EI-MS of fentanyl analogs have indistinguishable mass spectra at the 0.1 dissimilarity threshold (see Figure **6**).
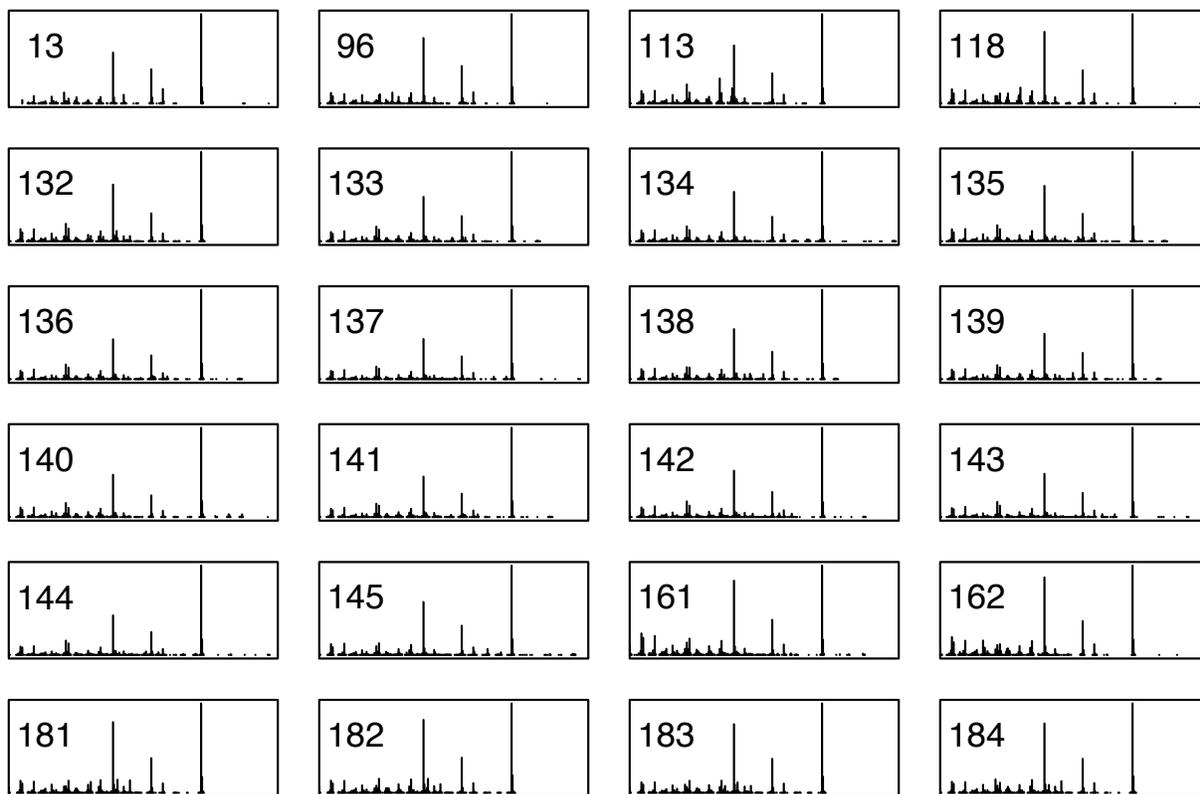
Figure 6: Sparkline-style mass spectra of fentanyl analogs with indistinguishable EI-MS measurements at a dissimilarity cut off level of 0.1 (or similarity level of 0.9). For all plots, the x-axis range is between *m/z* 40 and *m/z* 320, the y-axis is relative intensity with a range between 0 and 1, and the base peak occurs at *m/z* 245. Label numbers correspond with IDs in Table S1.

As noted earlier in the paper, the limitations of EI-MS are well-established. And given that EI-MS is usually preceded by GC, it is prudent to consider retention times (or retention indices) while trying to discriminate samples. We can simulate this experience by computing a combined RI and EI-MS dissimilarity prior to an AHC analysis (see Figure 3b). Because the measurements differ in range and interpretability, it is easiest to work with binary values based on whether the individual dissimilarities are above or below specified cutoff values. In particular, we can compute the combined dissimilarity of two fentanyl analogs using RI and EI-MS as

$$\phi_{12} = \begin{cases} 0, & \text{if } \phi_1 \leq \tau_1 \text{ and } \phi_2 \leq \tau_2 , \\ 1, & \text{otherwise}, \end{cases} \tag{4}$$

where ϕ_1 and ϕ_2 are the computed RI and EI-MS dissimilarity, respectively, and τ_1 and τ_2 are the dissimilarity cutoff value for RI and EI-MS, respectively. Using the cutoff values of $\tau_1 = 10$ and $\tau_2 = 0.1$, the measurement diversity of the fentanyl analogs is 0.89 a.u. The cluster breakdown when using combined dissimilarity is: two clusters consisting of three indistinguishable compounds, 17 clusters comprised of two indistinguishable compounds, and 157 individual clusters each containing a single compound. Following the same thought process, we can further improve our measurement diversity by combining RI with EI-MS and DART-MS, resulting in a diversity of 0.96 a.u. and only seven pairs of indistinguishable compounds. With the ESI-MS/MS contribution, measurement diversity increases further to 0.98 a.u. with only five pairs of indistinguishable compounds (see Figure **7**). Of the identified pairs of compounds that were indistinguishable with all standard library measurements, almost all were positional isomers.

While the results from this study provide us with useful intuition about the measurement diversity of fentanyl analogs, it is important to understand that the numerical results are imperfect. For instance, we are using a limited set of fentanyl analogs. At the time of the study, the NIST libraries shared complete measurements for only 197 fentanyl analogs. Also, the measurements were collected as part of general library building in the NIST Mass Spectrometry Data Center. Measurement diversity could vary when using targeted methods specific for drug analysis [36–38] or using different instrumentation with varied laboratory conditions.

In addition to the quality and comprehensiveness of the measurements, the underlying mathematical approach for approximating object and cluster dissimilarity greatly effects the computed measurement diversity. In this study, we considered traditional measures of dissimilarity for comparing retention indices and EI mass spectra, and extensions of traditional methods for comparing MS^2 collected with ESI-MS/MS and is-CID mass spectra collected with DART-MS. It

is possible that a specialized approach for approximating the dissimilarity of fentanyl measurements will improve the computed measurement diversity and thus the ease with which compounds can be uniquely identified.

ID	Compound	ID	Compound
6		8	
11		115	
37		43	
78		117	
165		169	
154*		174*	
181*		182*	
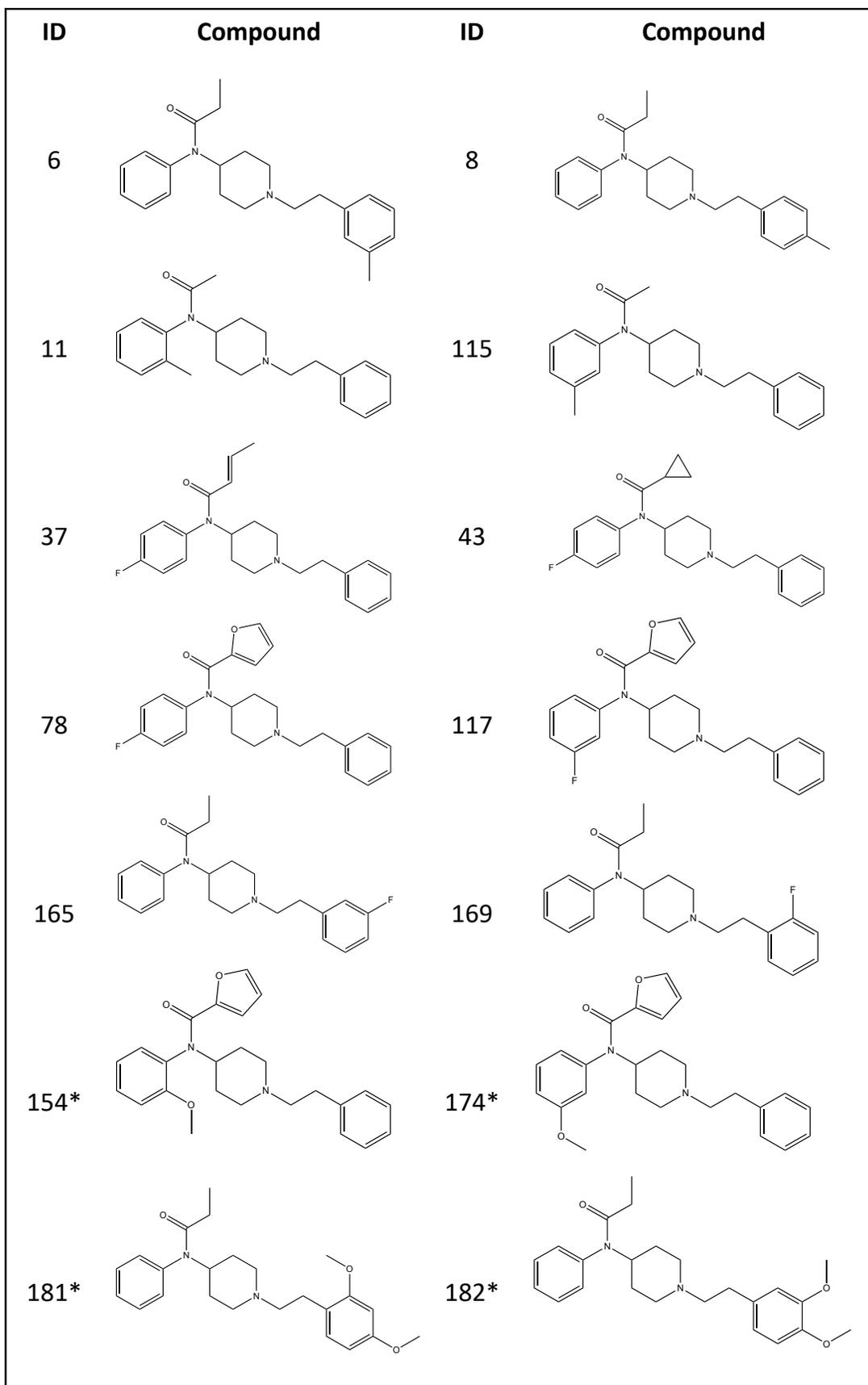
Figure 7: Pairs of fentanyl analogs that were indistinguishable using measurements from NIST libraries and discrimination requirements of (i) absolute retention index difference of 10 units, (ii) EI-MS dissimilarity of 0.1, (iii) DART-MS dissimilarity of 0.1 a.u., and (iv) ESI-MS/MS dissimilarity of 0.1. Pairs 154/174 and 181/182 (denoted by an *) are distinguishable using ESI-MS/MS but not the other three techniques.

A natural extension to this study is to characterize measurement diversity while leveraging replicate measurements per fentanyl analog. Having replicate measurements allows us to use a broader variety of mathematical tools [39–42], including machine learning [43,44]. Additionally, it would be fruitful to evaluate measurement diversity for other classes of drugs [45] and using measurements beyond retention indices and mass spectra.

4. Conclusions

In this paper, we tried to better understand the challenge of unambiguous identification of fentanyl analogs by exploring the measurement diversity of these substances in NIST mass spectral libraries. We defined measurement diversity through the results of agglomerative hierarchical clustering and were able to identify that these analytes were most diverse when measured with DART-MS. Measurement diversity improved by combining multiple measurements, but still five pairs of compounds (mostly positional isomers) were indistinguishable with the standard measurements in NIST libraries. Designing custom/targeted methods or new mathematical approaches to characterize pairwise dissimilarity—including leveraging replicate measurements—are potentially fruitful approaches that will improve our ability to identify fentanyl analogs.

5. Disclaimer

Official contribution of the National Institute of Standards and Technology (NIST); not subject to copyright in the United States. Certain commercial products are identified in order to adequately specify the procedure; this does not imply endorsement or recommendation by NIST, nor does it imply that such products are necessarily the best available for the purpose.

6. Acknowledgments

The authors would like to thank Dr. Gary Mallard (NIST) for sharing his insights about retention indices and the general challenges with compound identification using mass spectrometry.

7. References

1. Volkow ND. The epidemic of fentanyl misuse and overdoses: challenges and strategies. World Psychiatry 2021;20(2):195–6. https://doi.org/10.1002/wps.20846.

2. Armenian P, Vo KT, Barr-Walker J, Lynch KL. Fentanyl, fentanyl analogs and novel synthetic opioids: A comprehensive review. Neuropharmacology 2018;134:121–32. https://doi.org/10.1016/j.neuropharm.2017.10.016.

3. Weedn VW, Elizabeth Zaney M, McCord B, Lurie I, Baker A. Fentanyl-related substance scheduling as an effective drug control strategy. Journal of forensic sciences 2021;66(4):1186–200.

4. Fentanyl Signature Profiling Report. U.S. Department of Justice, Drug Enforcement Agency, Special Testing Laboratory, 2019.

5. Pardo B, Taylor J, Caulkins JP, Kilmer B, Reuter P, Stein BD. The future of fentanyl and other synthetic opioids. Rand Corporation, 2019.

6. Moeller K, Svensson B. "Shop Until You Drop": Valuing Fentanyl Analogs on a Swedish Internet Forum. Journal of Drug Issues 2021;51(1):181–95. https://doi.org/10.1177/0022042620964129.

7. SWGDRUG MS Library Version 3.11. 2022. https://swgdrug.org (accessed December 2, 2022).

8. Sisco E, Burns A, Moorthy AS. Development and evaluation of a synthetic opioid targeted gas chromatography mass spectrometry (GC-MS) method. Journal of Forensic Sciences 2021;66(6):2369–80. https://doi.org/10.1111/1556-4029.14877.

9. Bergh MS-S, Bogen IL, Wohlfarth A, Wilson SR, Øiestad ÅML. Distinguishing Between Cyclopropylfentanyl and Crotonylfentanyl by Methods Commonly Available in the Forensic Laboratory. Therapeutic Drug Monitoring 2019;41(4):519–27. https://doi.org/10.1097/FTD.0000000000000617.

10. Scientific Working Group for the Analysis of Seized Drugs (SWGDRUG) Recommendations Version 8.0. 2019.

11. Nielsen F. Hierarchical clustering. Introduction to HPC with MPI for Data Science 2016;:195–211.

12. Murtagh F. Counting dendrograms: A survey. Discrete Applied Mathematics 1984;7(2):191–9. https://doi.org/10.1016/0166-218X(84)90066-0.

13. Murtagh F, Contreras P. Algorithms for hierarchical clustering: an overview. Wiley Interdisciplinary Reviews Data Mining Knowledge Discovery 2012;2(1):86–97.

14. Pesyna GM, Venkataraghavan Rengachari, Dayringer HE, McLafferty FW. Probability based matching system using a large collection of reference mass spectra. Anal Chem 1976;48(9):1362–8. https://doi.org/10.1021/ac50003a026.

15. Stein SE, Scott DR. Optimization and testing of mass spectral library search algorithms for compound identification. J Am Soc Mass Spectrom 1994;5(9):859–66. https://doi.org/10.1016/1044-0305(94)87009-8.

16. Wan KX, Vidavsky I, Gross ML. Comparing similar spectra: From similarity index to spectral contrast angle. J Am Soc Mass Spectrom 2002;13(1):85–8. https://doi.org/10.1016/S1044-0305(01)00327-0.

17. Koo I, Zhang X, Kim S. Wavelet- and Fourier-Transform-Based Spectrum Similarity Approaches to Compound Identification in Gas Chromatography/Mass Spectrometry. Anal Chem 2011;83(14):5631–8. https://doi.org/10.1021/ac200740w.

18. Kim S, Koo I, Jeong J, Wu S, Shi X, Zhang X. Compound Identification Using Partial and Semipartial Correlations for Gas Chromatography–Mass Spectrometry Data. Anal Chem 2012;84(15):6477–87. https://doi.org/10.1021/ac301350n.

19. Garg N, Kapono CA, Lim YW, Koyama N, Vermeij MJA, Conrad D, et al. Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures. International Journal of Mass Spectrometry 2015;377:719-7.

20. Moorthy AS, Wallace WE, Kearsley AJ, Tchekhovskoi DV, Stein SE. Combining Fragment-Ion and Neutral-Loss Matching during Mass Spectral Library Searching: A New General Purpose Algorithm Applicable to Illicit Drug Identification. 2017. https://pubs.acs.org/doi/pdf/10.1021/acs.analchem.7b03320 (accessed May 6, 2020).

21. Moorthy AS, Kearsley AJ. Pattern similarity measures applied to mass spectra. Progress in Industrial Mathematics: Success stories. Springer International Publishing, 2021;43–54.

22. Bittremieux W, Schmid R, Huber F. Comparison of Cosine, Modified Cosine, and Neutral Loss Based Spectrum Alignment For Discovery of Structurally Related Molecules. J Am Soc Mass Spectrom 2022.

23. Abramson FP. Automated identification of mass spectra by the reverse search. Anal Chem 1975;47(1):45–9. https://doi.org/10.1021/ac60351a028.

24. Moorthy AS, Sisco E. A New Library-Search Algorithm for Mixture Analysis Using DART-MS. J Am Soc Mass Spectrom 2021;32(7):1725–34. https://doi.org/10.1021/jasms.1c00097.

25. Moorthy AS, Tennyson SS, Sisco E. Updates to the Inverted Library Search Algorithm for Mixture Analysis.

26. Gronau I, Moran S. Optimal implementations of UPGMA and other common clustering algorithms. Information Processing Letters 2007;104(6):205–10.

27. NIST Mass Spectrometry Data Center. NIST Mass Spectrometry Data Center. https://chemdata.nist.gov/ (accessed March 6, 2023).

28. Sisco E, Moorthy AS. NIST DART-MS Forensics Database (is-CID). 2020.

29. Sisco E, Moorthy AS, Tennyson SS, Corzo R. NIST/NIJ DART-MS Data Interpretation Tool. 2021.

30. Wallace WE, Ji W, Tchekhovskoi DV, Phinney KW, Stein SE. Mass Spectral Library Quality Assurance by Inter-Library Comparison. J Am Soc Mass Spectrom 2017;28(4):733–8. https://doi.org/10.1007/s13361-016-1589-4.

31. Yang X, Neta P, Stein SE. Quality Control for Building Libraries from Electrospray Ionization Tandem Mass Spectra. Anal Chem 2014;86(13):6393–400. https://doi.org/10.1021/ac500711m.

32. Stein S. Mass Spectral Reference Libraries: An Ever-Expanding Resource for Chemical Identification. Anal Chem 2012;84(17):7274–82. https://doi.org/10.1021/ac301205z.

33. Sisco E, Moorthy AS, Watt LM. Creation and Release of an Updated NIST DART-MS Forensics Database. J Am Soc Mass Spectrom 2021;32(3):685–9. https://doi.org/10.1021/jasms.0c00416.

34. R Core Team. R: A language and environment for statistical computing. 2022.

35. Sisco E, Forbes TP. Forensic applications of DART-MS: A review of recent literature. Forensic Chemistry 2021;22:100294. https://doi.org/10.1016/j.forc.2020.100294.

36. Sisco E, Burns A, Moorthy AS. A Framework for the Development of Targeted Gas Chromatography Mass Spectrometry (GC-MS) Methods: Synthetic Cannabinoids. Journal of Forensic Science 66(5):1908–18. https://doi.org/10.1111/1556-4029.14775.

37. Sisco E, Burns A, Moorthy AS. Development and Evaluation of a Synthetic Cathinone Targeted Gas Chromatography Mass Spectrometry (GC-MS) Method. Journal of Forensic Science 66(5):1919–28. https://doi.org/10.1111/1556-4029.14789.

38. Sisco E, Burns A, Moorthy AS. Development and Evaluation of a Synthetic Opioid Targeted Gas Chromatography Mass Spectrometry (GC-MS) Method. https://doi.org/10.33774/chemrxiv-2021-0pcnq.

39. Bonetti J. Mass spectral differentiation of positional isomers using multivariate statistics. Forensic Chemistry 2018;9:50–61.

40. Liliedahl RE, Davidson JT. The differentiation of synthetic cathinone isomers using GC-EI-MS and multivariate analysis. Forensic Chemistry 2021;26:100349.

41. Stuhmer EL, McGuffin VL, Waddell Smith R. Discrimination of seized drug positional isomers based on statistical comparison of electron-ionization mass spectra. Forensic Chemistry 2020;20:100261. https://doi.org/10.1016/j.forc.2020.100261.

42. Roberts MJ, Moorthy AS, Sisco E, Kearsley AJ. Incorporating measurement variability when comparing sets of high-resolution mass spectra. Analytica Chimica Acta 2022;1230:340247.

43. Koshute P, Hagan N, Jameson NJ. Machine learning model for detecting fentanyl analogs from mass spectra. Forensic Chemistry 2022;27:100379. https://doi.org/10.1016/j.forc.2021.100379.

44. Bonetti JL, Samanipour S, van Asten AC. Utilization of machine learning for the differentiation of positional NPS isomers with direct analysis in real time mass spectrometry. Analytical Chemistry 2022;94(12):5029–40.

45. Feeney W, Moorthy AS, Sisco E. Spectral trends in GC-EI-MS data obtained from the SWGDRUG mass spectral library and literature: A resource for the identification of unknown compounds. Forensic Chemistry 2022;31:100459.

Supplemental Information

Computing Cosine Similarity with an *m/z* tolerance specification

Given two mass spectra, and an accepted *m/z* tolerance of ϵ, we use the following steps to compute the similarity of the mass spectra:

1. Sort each mass spectrum such that peaks are ordered from highest to lowest intensity.

2. Create a sink variable denoted M and set it equal to zero.

3. Using one spectrum as a "query" and starting from the highest intensity peak, look for the first peak in the second "reference" spectrum within *m/z* tolerance of the query peak. If a matching peak exists, multiply the peak intensities from the query and reference spectrum at that *m/z* value ($\pm\epsilon$) and add the result to the sink variable M; remove the peak information from the second spectrum.

4. Continue until all peaks from the query spectrum have been searched against the reference.

5. Compute the sum of squared intensity for all peaks in both the original query and the reference spectra (before peaks were removed in step 3), denoted Q and L respectively.

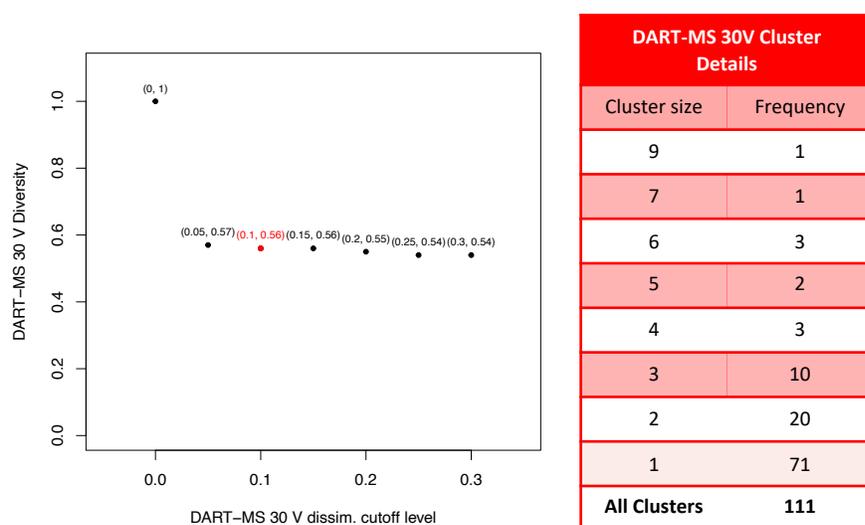6. Compute the cosine similarity between the two spectra as

$$\theta(x, y, \epsilon) = \frac{M}{\sqrt{Q} \sqrt{L}}.$$

Table S1: Summary of all fentanyl analogs considered in this study. For complete InChIKey, append "QHHAFSJGSA-N" to compound 37, "XNWCZRBMSA-N" to compound 38, "OCCSQVGLSA-N" to compound 114, "BIMFAAKUSA-N" to compound 171, and "UHFFFAOYSA-N" for all others.

Index	Name	Formula	InChIKey part 1
1	Acetanilide, N-(1-phenethyl-4-piperidyl)-	C21H26N2O	FYIUUQUPOKIKNI
2	Propanamide, N-phenyl-N-4-piperidinyl-	C14H20N2O	PMCBDBWCQQBSRJ
3	.alpha.-Methylfentanyl	C23H30N2O	NGTVDHYUFBKWID
4	para-Fluorofentanyl	C22H27FN2O	KXUBAVLIJFTASZ
5	Propanamide, N-phenyl-N-[1-(phenylmethyl)-4-piperidinyl]-	C21H26N2O	POQDXIFVWVZVML
6	1-[2-(3-Methylphenyl)ethyl]-4-(N-propanilido)piperidine	C23H30N2O	RBPHYEMEGYTXGK
7	1-[2-(3-Methylphenyl)ethyl]-4-(N-acetanilido)piperidine	C22H28N2O	RROIUIGBEXMCGR
8	Propanamide, N-phenyl-N-[1-(2-(4-methylphenyl)ethyl)-4-piperidinyl]-	C23H30N2O	AGSRCZOBGTWFFY
9	Propanamide, N-(2-methylphenyl)-N-[1-(2-phenylethyl)-4-piperidinyl]-	C23H30N2O	DPAJFOSXYXNYMA
10	1-(2-Phenylethyl)-4-(4-methyl-N-propananilido)piperidine	C23H30N2O	XHWYYMNEJCMADF
11	1-(2-Phenylethyl)-4-(2-methyl-N-acetanilido)piperidine	C22H28N2O	GRDWUDZBHWHLSH
12	1-(2-Phenylethyl)-4-(4-methyl-N-acetanilido)piperidine	C22H28N2O	JNQPTABAZAHVEN
13	.beta.-Methylfentanyl	C23H30N2O	UXIGUKSHASXDNI
14	.alpha.-Methylfentanyl acetyl analog	C22H28N2O	OKTLVZBUKMRPLL
15	Acetamide, N-phenyl-N-[1-(1-phenyl-2-propyl)-4-piperidinyl]-	C22H28N2O	MDRBZPCJVBUXSG
16	Acetyl fentanyl 4-methylphenethyl analog	C22H28N2O	BBMRIHXVAUNKEV
17	3-Fluorofentanyl analog	C22H27FN2O	SLTQVWMQISKVDN
18	1-[2-(2-methylphenyl)ethyl]-4-(N-acetanilido)piperidine	C22H28N2O	QTVCZKXTTIXGRA
19	Butyrlfentanyl	C23H30N2O	QQOMYEQLWQJRKK
20	.beta.-Hydroxythiofentanyl	C20H26N2O2S	GLAAETOTOUGGSB
21	para-Fluorobutyryl fentanyl	C23H29FN2O	QZFMCYUBPSLOBP
22	4-Methoxy-butyryl fentanyl	C24H32N2O2	FNVSEQCPMXWQKG
23	Furanylfentanyl	C24H26N2O2	FZJVHWISUGFFQV
24	Isobutyryl fentanyl	C23H30N2O	WRPFPNIHTOSMKU
25	Valeryl fentanyl	C24H32N2O	VCCPXHWAJYWQMR
26	Acrylfentanyl	C22H26N2O	RFQNLMWUIJJEQF
27	Cyclopentyl fentanyl	C25H32N2O	PEASFKSPITUZGT
28	para-Methoxyfentanyl	C23H30N2O2	QKQAUPIOVFHVJI
29	Despropionyl-2-fluorofentanyl	C19H23FN2	WUNLGTOLOUTCPE
30	Despropionyl para-fluorofentanyl	C19H23FN2	WWDHLOLWLHHFBH
31	Furanyl norfentanyl	C16H18N2O2	DDSUGTXSJFAMIM
32	para-Chlorofentanyl	C22H27ClN2O	CUGMWAHBYRKBKL
33	Cyclopropyl fentanyl	C23H28N2O	OIQSKDSKROTEMN
34	.alpha.'-Methyl butyryl fentanyl	C24H32N2O	DRAWSQJHCQAWNN
35	Methacrylfentanyl	C23H28N2O	YRRFMVAFZJGZNS
36	.alpha.-Methyl butyryl fentanyl	C24H32N2O	IPIGJTVNMCMXMT
37	p-Fluoro crotonyl fentanyl	C23H27FN2O	AXNCZLWDQLHJBM
38	Crotonyl fentanyl	C23H28N2O	VDYXGPCGBKLRDA
39	p-Methyl cyclopropyl fentanyl	C24H30N2O	HVDSXTDEYBCMRJ
40	Cyclobutyl fentanyl	C24H30N2O	XHRDMQWXMFOAIX
41	o-Methyl acrylfentanyl	C23H28N2O	OGVFVSLKWMUIGR
42	o-Fluoro acrylfentanyl	C22H25FN2O	ROBNYLIAYXEIFM
43	para-Fluoro cyclopropyl fentanyl	C23H27FN2O	CGLWFPGHBIADCN
44	Fentanyl Carbamate	C22H28N2O2	BPXVEPWHWMDYCP
45	ortho-Methyl methoxyacetyl fentanyl	C23H30N2O2	JJNJJOPCUJRZRG
46	.alpha.-methyl Thiofentanyl	C21H28N2OS	YPOXDUYRRSUFFG
47	para-Chloro cyclopropyl fentanyl	C23H27ClN2O	IPGLBSCVAGTNBX
48	para-Fluoro acrylfentanyl	C22H25FN2O	ZTLLQVADDIYHJU
49	Phenyl fentanyl	C26H28N2O	BJPDWVPQDSVQKD
50	Tetrahydrofuran fentanyl	C24H30N2O2	OHJNHKUFSKAANI
51	Methoxyacetyl fentanyl	C22H28N2O2	SADNVKRDSWWFTK
52	ortho-Fluoroisobutyryl fentanyl	C23H29FN2O	NEWBKHYGTFTPMC
53	para-fluoro Furanyl fentanyl 3-furancarboxamide	C24H25FN2O2	XXKBQRKRYOZPGV
54	para-Methyl tetrahydrofuran fentanyl	C25H32N2O2	XFSPASILKQGTMK
55	Phenylacetyl fentanyl	C27H30N2O	QIGINNCOUQLAIF
56	para-Chloroisobutyryl fentanyl	C23H29ClN2O	YWHLYGSHOQKCJG
57	meta-Fluorobutyryl fentanyl	C23H29FN2O	BSGQPGPNNBBEEQ
58	ortho-Fluoro furanyl fentanyl	C24H25FN2O2	QAURJPLHDYKGHY
59	meta-Fluoroisobutyryl fentanyl	C23H29FN2O	QBNWLXVOLIWDQF
60	para-Chloro methoxyacetyl fentanyl	C22H27ClN2O2	SKFHINXGRSUWIJ
61	.beta.'-Phenyl fentanyl	C28H32N2O	DIRAGWDYMRIDIO
62	Benzodioxole fentanyl	C27H28N2O3	ZFAAZMIOHJNKGD
63	para-Methyl isobutyryl fentanyl	C24H32N2O	SVTFAGUOPSJFBQ
64	Despropionyl p-fluorofentanyl, N-acetyl	C21H25FN2O	OXKSBDHHJOPFAH
65	Pivaloyl fentanyl	C24H32N2O	ZSFLALBHNJDBFJ
66	Furanyl fentanyl 3-furancarboxamide isomer	C24H26N2O2	AEDOTOMIDAMDFC
67	para-Chloro acrylfentanyl	C22H25ClN2O	CSFZVPAQJHDOCD
68	para-Fluoro methoxyacetyl fentanyl	C22H27FN2O2	KDXSBALZECTNCT
69	Isovaleryl fentanyl	C24H32N2O	HQXKBEMWVAZFPK

70	Tetrahydrofuran fentanyl 3-tetrahydrofurancarboxamide	C24H30N2O2	IYRYSBGEQAGZME
71	para-Methoxy acrylfentanyl	C23H28N2O2	OPFSSCQJIGSMRG
72	ortho-Methoxy butyryl fentanyl	C24H32N2O2	RWKPKWZAKCWONI
73	meta-Methyl methoxyacetyl fentanyl	C23H30N2O2	FQVRNAAHEQDAAK
74	ortho-Fluorofentanyl	C22H27FN2O	BKUWDIVZCJNXRA
75	para-Chlorobutyryl fentanyl	C23H29ClN2O	YRAMWTYYYLTADR
76	Ethoxyacetyl fentanyl	C23H30N2O2	FBHHANTXGLSYKL
77	meta-Fluoro methoxyacetyl fentanyl	C22H27FN2O2	GLYKZHGNYYYJLH
78	para-Fluoro furanyl fentanyl	C24H25FN2O2	MGEQOAJOFFXWTR
79	para-Fluoro cyclopentyl fentanyl	C25H31FN2O	FHNWCPRXMDUFOQ
80	ortho-Methyl phenyl fentanyl	C27H30N2O	RQTVMBDUXYALMZ
81	para-Fluoro valeryl fentanyl	C24H31FN2O	LRGKRSJHHVAEAD
82	para-Methyl methoxyacetyl fentanyl	C23H30N2O2	IHNULIRQUWXUEL
83	Cyclohexyl fentanyl	C26H34N2O	CWLVCTNETVWMMZ
84	para-Methoxy valeryl fentanyl	C25H34N2O2	SVNODVBERPGABW
85	.alpha.'-Methoxy fentanyl	C23H30N2O2	JFUZJQRRXWYJQC
86	para-Methyl furanyl fentanyl	C25H28N2O2	KPCWABTXRUWDSW
87	para-Methoxy tetrahydrofuran fentanyl	C25H32N2O3	PRGGNEGQOIKQSQ
88	para-Methoxy furanyl fentanyl	C25H28N2O3	JCCKUDBCTLAHGQ
89	meta-Methyl furanyl fentanyl	C25H28N2O2	MAUKPWKPNGATGI
90	para-Chloro cyclopentyl fentanyl	C25H31ClN2O	SJMSXMAHJSSXOR
91	para-Fluoro tetrahydrofuran fentanyl	C24H29FN2O2	DNWBTPZLCFRPRR
92	2,2,3,3-Tetramethyl-cyclopropyl fentanyl	C27H36N2O	BYCDHAVFKDTVAM
93	Thienyl fentanyl	C19H24N2OS	JSOSWRYHPGIWGT
94	ortho-Isopropyl furanyl fentanyl	C27H32N2O2	AYVSHUVVWFNBHR
95	para-Chloro valeryl fentanyl	C24H31ClN2O	FCGXDEMTTPNXQP
96	Thiofentanyl	C20H26N2OS	YMRFZDHYDKZXPA
97	N-benzyl Furanyl norfentanyl	C23H24N2O2	GDPJXHFICUEJBT
98	Benzyl acrylfentanyl	C21H24N2O	BANFGBDVLADRGR
99	2'-Fluoro ortho-fluorofentanyl	C22H26F2N2O	AUXYZYMWQUXVDV
100	m-Methyl cyclopropyl fentanyl	C24H30N2O	JHXYIANWOVWBEY
101	p-Methyl cyclopentyl fentanyl	C26H34N2O	YJYUFHRBRGOFJY
102	p-Methoxy acetyl fentanyl	C22H28N2O2	MEAAIIUNOQOZBW
103	Fentanyl methyl carbamate	C21H26N2O2	TYJLZTSBOBABQF
104	Cyclopentenyl fentanyl	C25H30N2O	SYSQDIUBJXIHAM
105	4'-Fluorofentanyl	C22H27FN2O	LSSHGONYDWDTQM
106	N,N-Dimethylamido-despropionyl fentanyl	C22H29N3O	CNNWEBFAZNLXKY
107	p-Methoxy methoxyacetyl fentanyl	C23H30N2O3	UYPGPJKUKWRQGQ
108	Heptanoyl fentanyl	C26H36N2O	ZLPDQEYWTXBVRY
109	Tetrahydrothiophene fentanyl	C24H30N2OS	RABQLHAVXBEJOF
110	4-Phenylfentanyl	C28H32N2O	BXCJXJLHYMWMQU
111	Phenoxyacetyl fentanyl	C27H30N2O2	IFPKCNASKXNARX
112	2,3-seco-Fentanyl	C22H30N2O	TVSLDWMMAUJGPV
113	N-(3-Ethylindole) norfentanyl	C24H29N3O	KZRCNDCSHLUGKI
114	(.+/-.)-cis-3-Methyl norfentanyl	C15H22N2O	REORAZISQPSCHR
115	Despropionyl m-methylfentanyl, N-acetyl	C22H28N2O	SZZZZTBKJBQTDR
116	Hexanoyl fentanyl	C25H34N2O	PWKBFVCAQVONRR
117	m-Fluoro furanyl fentanyl	C24H25FN2O2	NLTYWPGXHYUTRN
118	2'-Methyl fentanyl	C23H30N2O	LCAFYALTIFGYLM
119	N-Benzyl phenyl norfentanyl	C25H26N2O	TVPYSIMEYHFHLN
120	Despropionyl m-methylfentanyl	C20H26N2	DSWKXTYRMTZUHK
121	N-Benzyl p-fluoro norfentanyl	C21H25FN2O	PUFHNCRAVCTYOY
122	Despropionyl 2'-fluoro o-fluorofentanyl	C19H22F2N2	UBBYFFKICFYDHU
123	N-Methyl norfentanyl	C15H22N2O	PCMUWYHREXSXFM
124	N-Benzyl p-fluoro cyclopropyl norfentanyl	C22H25FN2O	DGFCHXNCEJHFDB
125	p-Bromofentanyl	C22H27BrN2O	UTXQWNZEVFXSIA
126	p-Toluoyl fentanyl	C27H30N2O	MYYJDRIKKKVNOD
127	3',4'-Methylenedioxy .alpha.-methyl fentanyl	C24H30N2O3	QUIXSQPQKKOKJC
128	.alpha.-Dimethyl fentanyl	C24H32N2O	FOOGWVPUGNMCNY
129	N-(6-APB) Fentanyl	C25H30N2O2	RUIOAJGEGJVHIU
130	N-(6-APDB) Fentanyl	C25H32N2O2	FCKLMVPDFBDMDA
131	N-(2-APB) Fentanyl	C25H30N2O2	QOBILHNYPYXIGT
132	2',5'-Dimethoxy fentanyl	C24H32N2O3	SAQRGDVCSASPSH
133	2',5'-Dimethoxy 4'-ethyl fentanyl	C26H36N2O3	LDTWBDPHKDXMOK
134	2',5'-Dimethoxy 4'-bromo fentanyl	C24H31BrN2O3	IWJDXHKRMXYRFD
135	2',5'-Dimethoxy 4'-chloro fentanyl	C24H31ClN2O3	UYOUMHSSQFWUFE
136	2',5'-Dimethoxy 4'-propyl fentanyl	C27H38N2O3	ILWQIOCAWRGJKL
137	2',5'-Dimethoxy 4'-propylthio fentanyl	C27H38N2O3S	SBBMOHKSHNBAAX
138	2',5'-Dimethoxy 4'-methyl fentanyl	C25H34N2O3	WOAKBYSJZBMEGV
139	2',5'-Dimethoxy 3',4'-dimethyl fentanyl	C26H36N2O3	UPSPWPZILMIZFV
140	2',5'-Dimethoxy 4'-iodo fentanyl	C24H31IN2O3	HKBAXOAWQOWSMS
141	2',5'-Dimethoxy 4'-isopropyl fentanyl	C27H38N2O3	PPKRENZXLPJJJN
142	2',5'-Dimethoxy 4'-methylthio fentanyl	C25H34N2O3S	NYGUILYQUWYLKO
143	2',5'-Dimethoxy 4'-ethylthio fentanyl	C26H36N2O3S	RVFYNEYOJXXUQT
144	2',5'-Dimethoxy 4'-isopropylthio fentanyl	C27H38N2O3S	XUHDOFPVQOMQAX

145	2',5'-Dimethoxy 4'-trifluoromethyl fentanyl	C25H31F3N2O3	QYTMFDXAWBGVFX
146	3',4',5'-Trimethoxy .alpha.-methyl fentanyl	C26H36N2O4	UOWROPJYPQKYER
147	2',5'-Dimethoxy .alpha.-methyl fentanyl	C25H34N2O3	NMYYRSZOKMKYPF
148	2',5'-Dimethoxy 4'-chloro .alpha.-methyl fentanyl	C25H33ClN2O3	MYUPALHKQYRWHK
149	2',5'-Dimethoxy 4'-bromo .alpha.-methyl fentanyl	C25H33BrN2O3	NVCARFLHKMIFTK
150	2',5'-Dimethoxy 4'-iodo .alpha.-methyl fentanyl	C25H33IN2O3	DWQJPNTVTUVHQJ
151	2',5'-Dimethoxy 4'-methyl .alpha.-methyl fentanyl	C26H36N2O3	XGDQJZIBFOIVEQ
152	2',5'-Dimethoxy 4'-ethyl .alpha.-methyl fentanyl	C27H38N2O3	QVLIEDWSQXEYBL
153	2',5'-Dimethoxy 4'-butyl .alpha.-methyl fentanyl	C29H42N2O3	JGVJLLGNTVSNSP
154	o-Methoxy furanyl fentanyl	C25H28N2O3	DCDCWVRDPIOMBB
155	o-Methyl furanyl fentanyl	C25H28N2O2	BFKNHOZXFOQILA
156	Thiophene fentanyl	C24H26N2OS	CCHPKGYUIHSQIE
157	p-Chloro cyclobutyl fentanyl	C24H29ClN2O	BCYKSTXBLMALSD
158	N-Methyl cyclopropyl norfentanyl	C16H22N2O	QCCRFOCZMVDDGA
159	p-Methyl acrylfentanyl	C23H28N2O	PAGSZSGDIQRYHE
160	o-Methyl cyclopropyl fentanyl	C24H30N2O	BFQXWRBDVXLPEX
161	2',5'-Dimethoxy 4'-nitro fentanyl	C24H31N3O5	SRAXTWKCHURJIT
162	.beta.-Hydroxyfentanyl	C22H28N2O2	JEFVHLMGRUJLET
163	meta-Fluoro valeryl fentanyl	C24H31FN2O	RBNHLBNQCFIATD
164	meta-Fluoro acrylfentanyl	C22H25FN2O	JWYRFGPPCJVCSJ
165	3'-Fluorofentanyl	C22H27FN2O	QGNPBZQBUVNBKQ
166	4-Methyl fentanyl	C23H30N2O	GNKKPEHVTFKLMN
167	Cyclopropaneacetyl fentanyl	C24H30N2O	YZIHFOIVZMMDOI
168	para-Chloroacetyl fentanyl	C21H25ClN2O	ZTXNOSVZJODMMX
169	2'-Fluorofentanyl	C22H27FN2O	HWUYTWQIMNJNNQ
170	3'-Fluoro ortho-fluorofentanyl	C22H26F2N2O	XIHMBSWTNQURLE
171	Tigloyl fentanyl	C24H30N2O	NYWRTCNBYKQMSB
172	Thiophene fentanyl 3-thiophenecarboxamide	C24H26N2OS	GLAUHJTZUGIICI
173	para-Hydroxy butyryl fentanyl	C23H30N2O2	HMPNQEXOJZKKKE
174	meta-Methoxy furanyl fentanyl	C25H28N2O3	SHGHIPYKQOXWJN
175	Despropionyl ortho-methylfentanyl	C20H26N2	UQFMMFWGILFTGJ
176	Cyclopropyl norfentanyl	C15H20N2O	MFGRYVBZTFUIQB
177	N-Benzyl meta-fluoro Norfentanyl	C21H25FN2O	IYRNFWXZMPQELH
178	N-Benzyl meta-fluoro Cyclopropyl norfentanyl	C22H25FN2O	AOLNSOXELLOWKK
179	N-Methyl ortho-methyl phenyl fentanyl	C20H24N2O	VWGMIZQFXRZGKF
180	N-Methyl meta-methyl phenyl fentanyl	C20H24N2O	UEYMYUWEUXEFSB
181	2',4'-Dimethoxy fentanyl	C24H32N2O3	DHNOAILILCSUCH
182	3',4'-Dimethoxy fentanyl	C24H32N2O3	UAMWMLGPFPCSAP
183	2',3'-Dimethoxy fentanyl	C24H32N2O3	HDNHGNZVCQDFNF
184	2',6'-Dimethoxy fentanyl	C24H32N2O3	GSKXYRVXRAGXBJ
185	Despropionyl para-methylfentanyl	C20H26N2	PPFZZKQMVFHCCX
186	N-Methyl para-methyl phenyl fentanyl	C20H24N2O	LBCNYSDAKXSWOU
187	ortho-Fluoro valeryl fentanyl	C24H31FN2O	KAZMSFHTWANUCG
188	Fentanyl	C22H28N2O	PJMPHNIQZUBGLI
189	Sufentanil	C22H30N2O2S	GGCSSNBKKAUURC
190	Carfentanil	C24H30N2O3	YDSDEBIZUNNPOB
191	Propanamide, N-(1-(2-(4-ethyl-4,5-dihydro-5-oxo-1H-tetrazol-1-yl)ethyl)-4-(methoxymethyl)-4-piperidinyl)-N-phenyl-	C21H32N6O3	IDBPHNDTYPBSNI
192	Ocfentanil	C22H27FN2O2	NYISTOZKVCMVEL
193	Benzyl carfentanil	C23H28N2O3	QKRGHVDSVANHAD
194	Remifentanil	C20H28N2O5	ZTVQQQVZCWLTDF
195	Norcarfentanil	C16H22N2O3	HFNFODVRYCEIQL
196	N-methyl Norcarfentanil	C17H24N2O3	KKEVIELPQLXPRR
197	Norsufentanil	C16H24N2O2	ULOZGJWEIWAWML

DART-MS 30V Cluster Details	
Cluster size	Frequency
9	1
7	1
6	3
5	2
4	3
3	10
2	20
1	71
All Clusters	**111**

Figure S1: Diversity results for DART-MS measurements when using only low fragmentation energy (+30 V orifice 1 voltage). The figure is accompanied with a table describing the detailed clustering breakdown for analysis using the highlighted (red) dissimilarity cutoff level value (0.1), where frequency and cluster size indicate how often (frequency) clusters of the specified size occur in the cluster results.