Can social media monitoring help identify the next EVALI? An examination of Reddit posts about vitamin E acetate and Dank Vapes

Jessica K. Pepper[a], Lauren V. Zitney[a], Alexander Preiss[a], Thomas Bukowski[a], Glen Szczypka[a], Annice E. Kim[a]

[a]RTI International, 3040 E. Cornwallis Rd., Research Triangle Park, NC, USA

Abstract

Background: Scientists identified vitamin E acetate (VEA) and "Dank Vapes" (a fake brand of tetrahydrocannabinol [THC] vaping products) as contributors to the 2019–2020 outbreak of e-cigarette, or vaping, product use-associated lung injury (EVALI). On social media, people who post about vaping or THC discussed the causes of EVALI. We examined whether Reddit conversations may have served as early signals of the outbreak.

Methods: We collected Reddit posts from March 2018 to February 2020 on vaping- and THC-related subreddits that mentioned VEA or Dank Vapes. We identified peaks in post volume, examined post content, and used natural language processing to identify terms most characteristic of posts.

Results: There were almost no posts about VEA before EVALI. Subsequently, there were two peaks, both referencing media coverage of scientific findings that linked VEA to EVALI. Discussion regularly referenced concerns about the legitimacy of Dank Vapes before EVALI; peaks in posts were largely unrelated to scientific findings or media coverage of those findings. The terms most characteristic of VEA posts were EVALI-related; those most characteristic of Dank Vapes posts were about quality or legitimacy.

Conclusions: Although posts about VEA and Dank Vapes did not predict the outbreak, the public health community could use social media to encourage people who vape or use THC to report future health concerns (e.g., through FDA's Safety Reporting Portal). Researchers and regulators could also use social media to see if potentially problematic products, such as Dank Vapes, have a history of concern among individuals who use those products.

1. Introduction

In July 2019, the news media began covering a mysterious lung illness associated with vaping (Scutti, 2019). The Centers for Disease Control and Prevention (CDC) later coined a name for the illness: "EVALI," or e-cigarette, or vaping, product use-associated lung injury (Siegel et al., 2019). As of February 18, 2020, there have been 2,807 cases of EVALI in the United States, including 68 deaths (CDC, 2020). Although there is no single definitive cause of EVALI, the CDC identified vitamin E acetate (VEA) as "strongly linked" to the outbreak. VEA, which can be used as a thickener in tetrahydrocannabinol (THC) vapes, was detected in product samples and patient lung fluid (CDC, 2020). Other public health organizations identified "Dank Vapes" as a potential cause (Navon et al., 2019). Dank Vapes is not a specific product; it is branded, empty packaging that individuals can fill with any THC product (Blood, 2019). There is no centralized company behind the brand and consumers typically acquire Dank Vapes from informal sources, giving Dank Vapes a reputation as a "fake brand" (Blood, 2019). The news media covered EVALI, including its links to VEA and Dank Vapes, extensively during the outbreak (e.g., Algiers et al., 2021; Aubrey, 2019; Edwards, 2019; Sun, 2019; Sun and McGinley, 2019).

Understanding perspectives about EVALI among people who vape is challenging because THC use is illegal in many states. Researchers have used Reddit, a pseudonymous social media platform, to understand illicit vaping-related behaviors (Brett et al., 2019; Zhan et al., 2019). Reddit is a network of online communities organized by topic; forums for a given topic are known as subreddits.

This exploratory study aims to understand trends in the volume of Reddit conversations about VEA and Dank Vapes as the EVALI crisis evolved, identify the causes of peaks in the volume of conversations, and characterize the words and phrases that distinguish VEA and Dank Vapes conversations. We limited our analysis to subreddits focused on vaping and THC to focus on conversations among those who likely use THC/vaping products and could have been affected by EVALI. We sought to understand whether Reddit conversations might have served as early signals of the outbreak (i.e., explicitly linking VEA and Dank Vapes to health concerns) or whether posts were largely reactions to the outbreak after it was identified.

2. Materials and Methods

2.1 Identification of Reddit Conversations

We used Brandwatch, a social media listening software, to collect publicly available Reddit posts and comments that mentioned VEA or Dank Vapes. Brandwatch is a secure web application available only to registered users. Reddit data obtained through Brandwatch contain no personally identifying information. We performed data collection from May 20,

2021, to June 10, 2021. We searched for Reddit conversations posted between March 31, 2018, one year prior to the start of CDC reporting on EVALI cases, and February 16, 2020, one week after CDC ceased reporting on EVALI cases. Discussion of COVID-19, an unrelated illness affecting the lungs, became widespread in the United States after the search's end date.

We used an iterative process to develop a search query intended to minimize noise in the data while maximizing the collection of relevant conversations. The final query searched for original posts that mentioned VEA or Dank Vapes and responses to those posts. It excluded irrelevant phrases (e.g., "I am a bot, and this action was performed automatically"). The full search query appears in Supplemental File 1.

We downloaded all posts and associated metadata for our study time frame. Next, we collapsed the data by subreddit. We eliminated irrelevant subreddits (e.g., r/ SkincareAddiction, which contained posts discussing vitamin E as a moisturizer); general subreddits (e.g., r/news); and subreddits that accounted for less than 0.5% of downloaded posts. We did this separately for VEA and Dank Vapes posts and then combined the results into a single list (see Supplemental File 1). The final list had 21 subreddits, including four that were regional (e.g., r/FLMedicalTrees) and three focused on nicotine products (e.g., r/juul).

2.2   Volume Analysis

We identified peaks in daily conversation for VEA and Dank Vapes using a peak detection algorithm in the SciPy Python package (Virtanen et al., 2020). See Supplemental File 2 for details. We examined the driving factors behind each peak by collapsing that day's data by Reddit thread title, sorting by frequency of posts, selecting the threads with the most posts, and reading the original post in each of those threads. In most cases, the original post contained a clear connection to a media event (e.g., linked to a news article).

2.3   Identification of Characteristic Terms

We used natural language processing tools from several Python packages to identify terms that distinguished VEA and Dank Vapes conversations. First, we processed the text of the posts using the spaCy package's small English model, which converts natural language into structured data (Montani et al., 2020). Next, we loaded the processed posts, which were each categorized as related to VEA or Dank Vapes, into Scattertext, a package designed to find terms that distinguish collections of texts (Kessler, 2017). We used Scattertext to make a list of terms found in the posts and count term frequencies. We then calculated scaled F-scores for each identified term for the two categories, VEA and Dank Vapes. The scaled F-score (range 0–1) assumes that the terms that are characteristic of a category have high category-specific precision and high category-specific frequency (Kessler, 2018). Terms with scaled F-scores closer to 1 for a category are more characteristic of that category. We identified the top 10 characteristic terms for each category as those with the highest scaled F-scores, omitting terms that were variations of the query terms (e.g., "dank" and "vit e"). Finally, we calculated the difference in rank frequency both between the two categories (VEA and Dank Vapes), and between all posts and general English. We used these rank

differences to visualize terms by the extent to which they characterized Dank Vapes posts compared to VEA posts and all posts compared to general English. See Supplemental File 2 for details.

3. Results

3.1 Trends in Volume of Mentions of VEA and Dank Vapes

On vaping-related subreddits, there were only six posts about VEA prior to the September 5, 2019, article that gave rise to the September 6 peak, with 283 VEA-related original posts and replies in one day. Three of the most popular posts linked to a September 5 *Washington Post* article (Figure 1: Media Event A) (Sun, 2019). The article cites a New York State Department of Health press release linking VEA to vaping-related lung illness (New York State Department of Health, 2019).

VEA-related conversation declined and then peaked again on November 8, 2019, with 233 posts. The post driving this peak linked to an Ars Technica article (Figure 1: Media Event B) (Mole, 2019), which reported on CDC data released that day showing that VEA had been detected in all of the lung fluid samples from 29 EVALI patients (Blount et al., 2019). Conversations about VEA declined throughout the rest of the time frame.

In contrast, conversations about Dank Vapes began appearing regularly in October 2018, albeit at low frequency (usually fewer than 25 posts per day) (Figure 1). On July 26, 2019, there was a peak of 102 posts driven by three Reddit threads discussing a news story from the day prior (Figure 1: Media Event C) about a man put into a medically induced coma after suffering lung damage from vaping (Sears, 2019). Notably, the news story did not mention Dank Vapes by name; Redditors initiated those threads because a picture and video clip depicted Dank Vapes packaging.

Another peak of 121 posts on September 13, 2019, was not linked to discrete external events. The peak was driven by two organic threads discussing the legitimacy of Dank Vapes.

Four threads drove the peak of 109 posts on September 27, 2019. One thread linked to a *Washington Post* article (Figure 1: Media Event D) published that day connecting vaping-related lung injuries to THC products (Sun and McGinley, 2019). The article cited a just-released CDC report (Ghinai et al., 2019). A second thread linked to a September 25, 2019, news story (Figure 1: Media Event E) covering the seizure of illegal Dank Vapes THC cartridges in Virginia (Freeman, 2019). Two other threads contributing to the peak were organic conversations related to the legitimacy of Dank Vapes and the organization of posts within the subreddit.

The final Dank Vapes-related peak on December 24, 2019, had 80 posts and discussed whether Dank Vapes were fake. The original poster initiated the thread after seeing Dank Vapes advertised on Snapchat.

3.2 Characteristic Phrases for VEA and Dank Vapes

Table 1 shows the top 10 characteristic terms by scaled F-score for VEA and Dank Vapes posts. Supplemental Table 1 includes the full list of characteristic terms with scaled F-scores ≥ 0.95. Supplemental Figure 1 visualizes all terms by their association with each category (VEA or Dank Vapes) and with the entire collection of posts.

Most of the top 10 terms for VEA are relevant to EVALI illness (e.g., "samples," as in lung fluid samples; "illnesses"; "found in"; and "the lungs") or to scientific and regulatory bodies (e.g., "cdc" and "fda"). Posts with the characteristic term "mct" (medium-chain triglycerides) questioned whether the additive MCT oil causes EVALI. Posts with the characteristic term "tobacco" typically contrasted the harm caused by vaping and EVALI with the harm caused by tobacco use.

Most of the top 10 terms for Dank Vapes are related to the product's questionable legitimacy (e.g., "are fake," "fakes," and "the packaging") or to other potentially fake brands (e.g., "exotic carts" and "mario carts"), rather than to EVALI. Posts about Dank Vapes' legitimacy sometimes included concerns about danger, e.g., whether Dank Vapes contain pesticides.

4. Discussion

Scientific reports described both VEA and Dank Vapes as causes of EVALI, but the timing and content of Reddit discussions about VEA differed from those about Dank Vapes. For VEA, there was a clear, consistent pattern: the scientific community published a finding, media outlets summarized it, and Reddit conversation followed within 24 hours. Reddit conversations about VEA were entirely reactive, and the speed of response suggests that people on relevant subreddits are highly attuned to news coverage of vaping. Most of the top 10 terms associated with VEA were clear references to the scientific findings about health risks.

The causes of Dank Vapes-related peaks were variable and rarely rooted in scientific information. Redditors discussed the illegitimacy and potential danger of Dank Vapes as part of organic conversations about multiple potentially fake brands. Terms like "are fake" were characteristic of Dank Vapes conversations even before the first relevant media event. Perhaps because of pre-existing concern about Dank Vapes, Redditors noticed the product in news stories even when never mentioned by name. Although conversations expressing concerns about Dank Vapes occurred prior to the EVALI outbreak, they were focused on legitimacy and were not distinct signals of the upcoming outbreak.

4.1 Limitations

This study only captures the sentiments of individuals who discuss these topics on relevant subreddits. It does not represent all people who vape or use THC. We limited our study to topically-related subreddits in an attempt to capture conversations among those for whom EVALI is potentially personally relevant; however, users of these subreddits might be less at risk for EVALI than other people who vape because they are more knowledgeable about vaping/THC. They could also be community outsiders (e.g., people who post about but do not use THC). We could not capture posts that were deleted before our search. Our search

query may have missed conversations that used different wording to discuss VEA or Dank Vapes.

4.2 Conclusions

Redditors were concerned about fake and potentially dangerous brands like Dank Vapes but expressed no concerns about the health effects of VEA prior to media coverage of its link to EVALI. Our results do not suggest that monitoring Reddit data could serve as an early warning system to predict future outbreaks. However, public health organizations could use Reddit to communicate about the importance of reporting product-related health concerns. For example, organizations could reach out to moderators of relevant subreddits and encourage them to post links to FDA's Safety Reporting Portal. In addition, if the scientific community identifies a product as potentially problematic, researchers and regulators could use Reddit to understand if there is a history of concern about that product among its users.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Role of Funding Source

5. References

Algiers O, Wang Y, Laestadius L, 2021. Content analysis of U.S. Newspaper coverage of causes and solutions to vaping-associated lung injury. Subst Use Misuse. 56, 522–528. 10.1080/10826084.2021.1883663 [PubMed: 33627031]

Aubrey A, 2019. Many vaping illnesses linked to black market 'dank vapes' or other THC products, https://www.npr.org/sections/health-shots/2019/09/27/765151187/many-vaping-illnesses-linked-to-black-market-dank-vapes-or-other-thc-products. Accessed on July 13, 2021.

Blood M, 2019. Dank, the brand that isn't, scrutinized amid vaping illness. AP News. AP News, New York, NY.

Blount BC, Karwowski MP, Morel-Espinosa M, Rees J, Sosnoff C, Cowan E, Gardner M, Wang L, Valentin-Blasini L, Silva L, De Jesús VR, Kuklenyik Z, Watson C, Seyler T, Xia B, Chambers D, Briss P, King BA, Delaney L, Jones CM, Baldwin GT, Barr JR, Thomas J, Pirkle JL, 2019. Evaluation of bronchoalveolar lavage fluid from patients in an outbreak of e-cigarette, or vaping, product use-associated lung injury - 10 states, August-October 2019. MMWR Morb Mortal Wkly Rep. 68, 1040–1041. 10.15585/mmwr.mm6845e2 [PubMed: 31725707]

Brett EI, Stevens EM, Wagener TL, Leavens ELS, Morgan TL, Cotton WD, Hébert ET, 2019. A content analysis of JUUL discussions on social media: Using Reddit to understand patterns and perceptions of JUUL use. Drug Alcohol Depend. 194, 358–362. 10.1016/j.drugalcdep.2018.10.014 [PubMed: 30472576]

Centers for Disease Control and Prevention, 2020. Outbreak of lung injury associated with the use of e-cigarette, or vaping, products, https://www.cdc.gov/tobacco/basic_information/e-cigarettes/severe-lung-disease.html. Accessed on July 13, 2021.

Edwards E, 2019. Counterfeit 'dank vapes' among products linked to lung illnesses. NBC News. NBC News, New York City, NY.

Freeman J,V, 2019. Virginia police seize over 1,000 'Dank' vape cartridges linked to lung illnesses. WTVR Richmond. WTVR Richmond, Richmond, VA.

Ghinai I, Pray IW, Navon L, O'Laughlin K, Saathoff-Huber L, Hoots B, Kimball A, Tenforde MW, Chevinsky JR, Layer M, Ezike N, Meiman J, Layden JE, 2019. E-cigarette product use, or vaping, among persons with associated lung injury - Illinois and Wisconsin, April-September 2019. MMWR Morb Mortal Wkly Rep. 68, 865–869. 10.15585/mmwr.mm6839e2 [PubMed: 31581166]

Kessler J, 2017. Scattertext: a browser-based tool for visualizing how corpora differ. Association for Computational Linguistics, Stroudsburg, PA.

Kessler J, 2018. Lexicon mining, language visualization and semiotic squares in Python. Presented at the Puget Sound Python Programming Group. https://github.com/JasonKessler/PuPPyTalk. Accessed on September 7, 2021.

Mole B, 2019. Vitamin E acetate found in all lung fluid samples tested from injured vapers. https://arstechnica.com/science/2019/11/vitamin-e-acetate-found-in-all-lung-fluid-samples-tested-from-injured-vapers/. Accessed on July 13, 2021.

Montani I, Honnibal M, Van Landeghem SV, Boyd A, Peters H, et al., 2020. spaCy: Industrial-strength natural language processing in Python, Zenodo. Berlin, DC.

Navon L, Jones CM, Ghinai I, King BA, Briss PA, Hacker KA, Layden JE, 2019. Risk factors for e-cigarette, or vaping, product use-associated lung injury (EVALI) among adults who use e-cigarette, or vaping, products - Illinois, July-October 2019. MMWR Morb Mortal Wkly Rep. 68, 1034–1039. 10.15585/mmwr.mm6845e1 [PubMed: 31725708]

New York State Department of Health, 2019. New York State Department of Health announces update on investigation into vaping-associated pulmonary illnesses. Department warns against use of black market vaping products: Lab test results show high levels of vitamin E acetate, now focus of investigation. https://www.health.ny.gov/press/releases/2019/2019-09-05_vaping.htm. Accessed on July 13, 2021.

Proferes N, Jones N, Gilbert S, Fiesler C and Zimmer M, 2021. Studying Reddit: a systematic overview of disciplines, approaches, methods, and ethics. Social Media + Society. 7(2), 20563051211019004.

Scutti S, 2019 8 Wisconsin teens hospitalized with severe lung damage due to vaping, doctors suspect, https://www.cnn.com/2019/07/26/health/wisconsin-8-teens-lung-damage/index.html. Accessed on September 28, 2021.

Sears A, 2019. Burlington man suffers lung damage from vaping, put into medically-induced coma, https://www.fox6now.com/news/burlington-man-suffers-lung-damage-from-vaping-put-into-medically-induced-coma. Accessed on July 13, 2021.

Siegel DA, Jatlaoui TC, Koumans EH, Kiernan EA, Layer M, Cates JE, Kimball A, Weissman DN, Petersen EE, Reagan-Steiner S, Godfred-Cato S, Moulia D, Moritz E, Lehnert JD, Mitchko J, London J, Zaki SR, King BA, Jones CM, Patel A, Delman DM, Koppaka R, Lung Injury Response Clinical Working Group, Lung Injury Response Epidemiology/Surveillance Group, 2019. Update: interim guidance for health care providers evaluating and caring for patients with suspected e-cigarette, or vaping, product use associated lung injury — United States, 10 2019. https://www.cdc.gov/mmwr/volumes/68/wr/mm6841e3.htm?s_cid=mm6841e3_w. Accessed on September 28, 2021.

Sun L, 2019. Contaminant found in marijuana vaping products linked to deadly lung illnesses, tests show, https://www.washingtonpost.com/health/2019/09/05/contaminant-found-vaping-products-linked-deadly-lung-illnesses-state-federal-labs-show/?outputType=amp&utm_source=reddit.com. Accessed on July 13, 2021.

Sun L, McGinley L, 2019. Most vaping-related lung injuries linked to marijuana products, CDC says, https://www.washingtonpost.com/health/2019/09/27/most-vaping-related-lung-injuries-linked-marijuana-products-cdc-says/. Accessed on July 13, 2021.

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat I, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, SciPy 1.0 Contributors, 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods. 17, 261–272. [PubMed: 32015543]

Zhan Y, Zhang Z, Okamoto JM, Zeng DD, Leischow SJ, 2019. Underage JUUL use patterns: Content analysis of Reddit messages. J Med Internet Res. 21, e13038. 10.2196/13038 [PubMed: 31502542]

Highlights

- Vitamin E acetate (VEA) and Dank Vapes were causes of the vaping illness "EVALI"

- On vaping-related Reddit communities, Redditors did not discuss VEA before EVALI

- Redditors expressed concern about Dank Vapes as a fake brand prior to EVALI

- Monitoring Reddit may not provide early warning of future outbreaks like EVALI

- The public health community could use Reddit to encourage reporting safety concerns
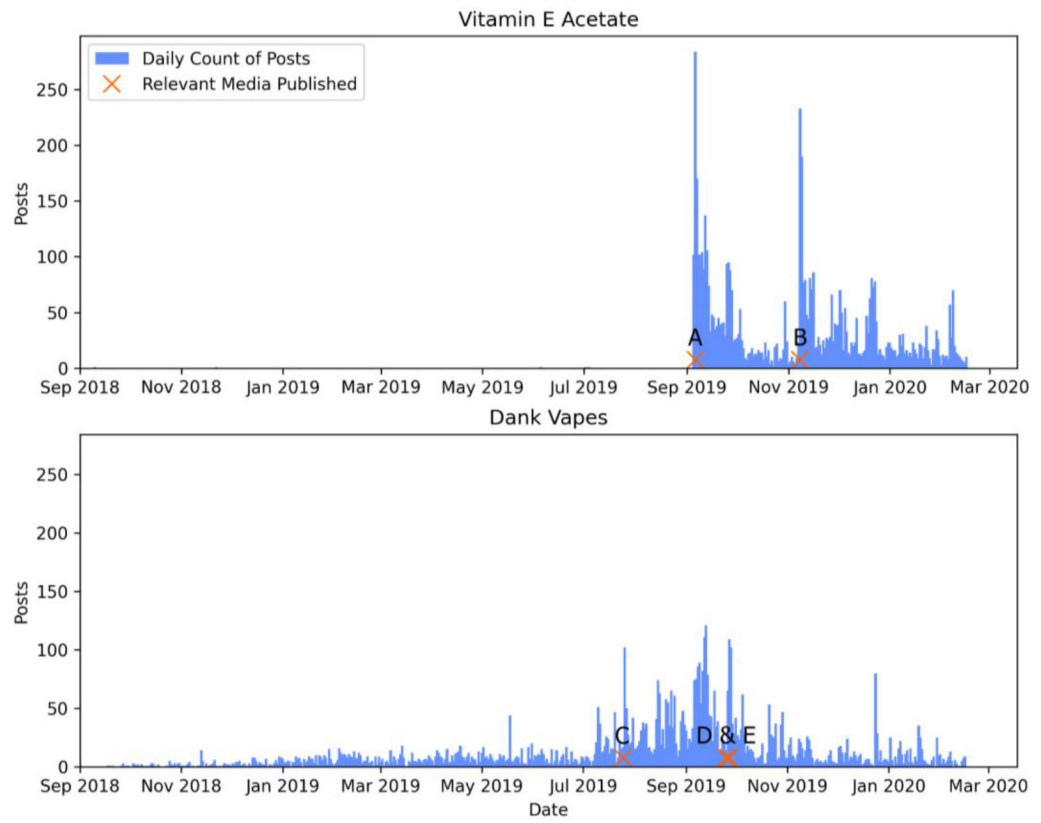
Figure 1.
Volume of mentions of VEA and Dank Vapes on Reddit (September 2018 – March 2020)

Table 1.

Top 10 characteristic terms for VEA and Dank Vapes by F-score (highest to lowest) and paraphrased excerpts from representative posts

Characteristic Term	VEA F-Score (range 0–1)	VEA Frequency	Dank Vapes F-Score (range 0–1)	Dank Vapes Frequency	Paraphrased Excerpt from Representative Post	Date Posted
VEA						
samples	0.997	220	0.003	3	According to the cdc, every sick person had VEA in the juice. And patients who first claimed they hadn't vaped thc lied after testing lung samples.	01/01/2020
tobacco	0.994	366	0.006	12	Of people who hit carts with vitamin E acetate, a tiny percentage get sick. Compare that to the percentage of people who die from alcohol, tobacco, and pills.	09/09/2019
illnesses	0.994	268	0.006	9	Vaping illnesses and respiratory problems we see now seem to be directly linked to vitamin e acetate and bootleg cartridges.	01/13/2019
met	0.994	257	0.006	9	Online, I saw carts with coconut-derived MCT oil as the base. I wonder…is Vitamin E acetate the cause of these illnesses or is vaping coconuts the problem?	09/16/2019
found in	0.989	188	0.011	9	Vitamin E was found in liquid samples people were using, but a recent study found many of those same samples contained arsenic too. I'd rather not inhale that cocktail!	10/20/2019
the lungs	0.987	168	0.013	7	Samples from patients across the US provide clear evidence of VEA at the primary site of injury within the lungs.	11/08/2019
the cdc	0.986	295	0.014	22	If the cdc is concerned that vitamin E acetate is the real cause of these lung illnesses the test results should be made public and should be written in a way that even a 7 year-old could understand.	09/08/2019
nicotine	0.985	700	0.015	55	Nicotine products were tested, and none contained alpha tocopherol, gamma tocopherol, MCT, THC, or VEA.	11/27/2019
cdc	0.985	417	0.015	33	I'm pretty sure the CDC reported that the problem was Vitamin E acetate in black market THC vape carts.	11/27/2019
fda	0.983	195	0.017	16	Check out the CDC and the FDA findings: the whole problem was illegal THC carts containing Vitamin E acetate.	12/28/2019
Dank Vapes						
exotic	0.001	2	0.999	246	A lot of fake carts that aren't even real brands like brass knuckles, dr z., dank vapes, and exotic are sold at legal dispensaries	01/13/2020
exotic carts	0.002	1	0.998	166	You should be concerned about Mario Carts, Exotic Carts, and any other similar carts.	02/06/2019
mario	0.004	12	0.996	275	I've never had a problem with Dank Vapes. But I tried a Mario cart once and will never buy that again.	08/03/2019
boxes	0.005	10	0.995	188	I've used Dank Vapes carts a lot and have only come across a few bad ones. But you can buy empty boxes that say Dank Vapes online and fill the carts with whatever you want before selling them.	04/19/2019
mario carts	0.006	9	0.994	167	Just stick with medical brands, stay away from dank vape carts or Mario carts.	08/15/2019
are fake	0.007	13	0.993	193	Dank Vapes are fake, and you're just inhaling pesticides. The THC level is definitely not what the packaging says.	04/18/2019

Characteristic Term	VEA F-Score (range 0–1)	VEA Frequency	Dank Vapes F-Score (range 0–1)	Dank Vapes Frequency	Paraphrased Excerpt from Representative Post	Date Posted
the packaging	0.009	25	0.991	270	I don't see dank vapes at dispensaries, but people can buy the packaging and empty carts online at ebay or wherever. So anyone can fill dank vapes carts with whatever they want.	07/09/2019
fakes	0.009	20	0.991	206	Used a bunch of other fakes, but never actual dank Vapes.	09/26/2019
illegal state	0.010	34	0.990	329	I've been hearing carts like dank vapes are filled with pesticides. But when you live in an illegal state, you don't know what you're actually buying.	03/20/2019
packaging	0.011	108	0.989	957	Be careful with dank vapes. It's just packaging and carts that get filled by your plug or whoever sells to them.	12/19/2018

MCT = medium-chain triglycerides. VEA = vitamin E acetate.

Note: To maintain privacy of Reddit users, posts have been paraphrased rather than quoted. For more information, please see related literature such as Proferes et al., 2021 (full citation in references).