

# UC San Diego

## UC San Diego Previously Published Works

### Title

Identification and Characterization of Synthetic Nicotine Product Promotion and Sales on Instagram Using Natural Language Processing.

### Permalink

<https://escholarship.org/uc/item/9km4j1cr>

### Journal

Nicotine & Tobacco Research, 26(5)

### Authors

Shah, Neal

Li, Zhuoran

McMann, Tiana

et al.

### Publication Date

2024-04-22

### DOI

10.1093/ntr/ntad222

Peer reviewed

# Identification and Characterization of Synthetic Nicotine Product Promotion and Sales on Instagram Using Natural Language Processing

Neal A. Shah, MS<sup>1,2</sup>, Zhuoran Li, MS<sup>3,4</sup>, Tiana McMann, MA<sup>2,4,5</sup>, Alec J. Calac, BS<sup>2,6</sup>,,  
Nicolette Le, MA<sup>5</sup>, Matthew C. Nali, BA<sup>1,2,4</sup>,, Raphael E. Cuomo, PhD<sup>1,2</sup>,,  
Tim K. Mackey, MAS, PhD<sup>3,4,5</sup>,

<sup>1</sup>Department of Anesthesiology, University of California, San Diego School of Medicine, San Diego, CA, USA

<sup>2</sup>Global Health Policy and Data Institute, San Diego, CA, USA

<sup>3</sup>San Diego Supercomputer Center, University of California, San Diego, CA, USA

<sup>4</sup>S-3 Research, San Diego, CA, USA

<sup>5</sup>Global Health Program Department of Anthropology, University of California, San Diego, La Jolla, CA, USA

<sup>6</sup>The Herbert Wertheim School of Public Health and Human Longevity Science, University of California, San Diego, CA, USA

**Corresponding Author:** Timothy K. Mackey, MAS, PhD, Global Health Program Department of Anthropology, University of California, San Diego, 9500 Gilman Drive, Mail Code: 0505, La Jolla, CA 92093, USA. E-mail: [tmackey@ucsd.edu](mailto:tmackey@ucsd.edu)

## Abstract

**Introduction:** There has been a rapid proliferation of synthetic nicotine products in recent years, despite newly established regulatory authority and limited research into its health risks. Previous research has implicated social media platforms as an avenue for nicotine product unregulated sales. Yet, little is known about synthetic nicotine product content on social media. We utilized natural language processing to characterize the sales of synthetic nicotine products on Instagram.

**Methods:** We collected Instagram posts by querying Instagram hashtags (eg, “#tobaccofreenicotine”) related to synthetic nicotine. Using Bidirectional Encoder Representations from Transformers, collected posts were categorized into thematically related topic clusters. Posts within topic clusters relevant to study aims were then manually annotated for variables related to promotion and selling (eg, cost discussion, contact information for offline sales).

**Results:** A total of 7425 unique posts were collected with 2219 posts identified as related to promotion and selling of synthetic nicotine products. Nicotine pouches (52.9%,  $n = 1174$ ), electronic nicotine delivery systems (30.6%,  $n = 679$ ), and flavored e-liquids (14.1%,  $n = 313$ ) were most commonly promoted. About 16.1% ( $n = 345$ ) of posts contained embedded hyperlinks and 5.8% ( $n = 129$ ) provided contact information for purported offline transactions. Only 17.6% ( $n = 391$ ) of posts contained synthetic nicotine-specific health warnings.

**Conclusions:** In the United States, synthetic nicotine products can only be legally marketed if they have received premarket authorization from the Food and Drug Administration (FDA). Despite these prohibitions, Instagram appears to be a hub for potentially unregulated sales of synthetic and “tobacco-free” products. Efforts are needed by platforms and regulators to enhance content moderation and prevent unregulated online sales of existing and emerging synthetic nicotine products.

**Implications:** There is limited clinical understanding of synthetic nicotine’s unique health risks and how these novel products are changing over time due to regulatory oversight. Despite synthetic nicotine-specific regulatory measures, such as the requirement for premarket authorization and FDA warning letters issued to unauthorized sellers, access to and promotion of synthetic nicotine is widely occurring on Instagram, a platform with over 2 billion users and one that is popular among youth and young adults. Activities include direct-to-consumer sales from questionable sources, inadequate health warning disclosure, and exposure with limited age restrictions, all conditions necessary for the sale of various tobacco products. Notably, the number of these Instagram posts increased in response to the announcement of new FDA regulations. In response, more robust online monitoring, content moderation, and proactive enforcement are needed from platforms who should work collaboratively with regulators to identify, report, and remove content in clear violation of platform policies and federal laws. Regulatory implementation and enforcement should prioritize digital platforms as conduits for unregulated access to synthetic nicotine products and other future novel and emerging tobacco products.

## Introduction

The advent and proliferation of synthetic nicotine products, or non-tobacco nicotine products (NTNs), including NTN nicotine pouches, electronic nicotine delivery systems (ENDS), e-liquids, and nicotine salts, presents novel challenges to global regulatory bodies that have oversight over tobacco-derived products. In the United States, the Food and Drug

Administration (FDA)’s regulatory authority was extended to cover ENDS in 2016, but with this change, only covered products that contained nicotine made or derived from tobacco.<sup>1</sup> As such, synthetic nicotine product manufacturers and retailers introduced these products into the retail market and continued to promote and sell NTNs without specific FDA regulations or oversight existing at the time.<sup>1</sup>

The emergence of synthetic nicotine products, combined with a lack of transparency of manufacturers' formulation processes, has led to limited understanding of the specific dangers and health concerns of NTN. Previous research has shown that the synthetic nicotine in these products is not equal to tobacco-derived nicotine. Instead, there are varying concentrations of *R*- and *S*-nicotine enantiomers in synthetic nicotine compared to the 99% *S*-nicotine profile of tobacco-derived nicotine; notably, some NTN contain a racemic mixture of 50% *R*-nicotine and 50% *S*-nicotine.<sup>1</sup> Further, the metabolic and physiological effects of high exposure to *R*-nicotine remain poorly understood.<sup>2,3</sup> Research limited to animal models has shown a difference in the metabolic breakdown of *R*-nicotine compared to *S*-nicotine, but the physiological effects of these metabolites are not well understood. Previous testing analyses demonstrated incongruencies between the labeled nicotine concentration of some synthetic nicotine ENDS and true nicotine concentrations, and significantly greater concentrations of *R*-nicotine in comparison to tobacco-derived ENDS, potentially exposing individuals to higher concentrations of nicotine than advertised.<sup>3</sup> Further, claims that a product is "tobacco-free" may reduce harm perceptions, may be incorrectly interpreted, and may increase intent to use these products.<sup>4</sup>

Despite limited knowledge of synthetic nicotine's effects, NTN products have continued to proliferate. In February 2021, Puff Bar, the most popular ENDS product among high school- and middle school-aged students, announced a new synthetic nicotine formulation.<sup>5</sup> This announcement was the first instance of a major manufacturer announcing a shift to synthetic nicotine formulations, with increased availability of synthetic nicotine products in the marketplace occurring soon thereafter. Furthermore, in recent years, NTN nicotine pouches—such as NIIN and Rush—have introduced synthetic nicotine formulations.<sup>1,4</sup>

In April 2022, Federal law was passed that clarified that the FDA had regulatory authority over the manufacturing, distribution, and marketing of tobacco products containing nicotine from any source, explicitly including products made with NTN.<sup>6</sup> Despite federal requirements to seek premarket authorization by May 2022 for these products and warning letters issued to manufacturers and retailers, numerous NTN products continue to be promoted, bought, and sold via different channels of promotion and access.<sup>7</sup>

The concerns regarding the continued availability of synthetic nicotine products mirror concerns about proliferation of naturally occurring nicotine ENDS products amidst increased popularity among youth and young adult populations. By 2019, the percentage of US middle school- and high school-aged students who reported using ENDS had peaked to an all-time high.<sup>8</sup> Research has also demonstrated a substantial increase in the use of nicotine pouches between 2016 and 2020.<sup>8</sup>

Digital communication platforms—specifically, social media—have been implicated as a leading driver behind the promotion of nicotine ENDS products.<sup>9</sup> Furthermore, nicotine pouch manufacturers, such as ZYN, have focused on increasing their share of online advertising compared to their use of traditional advertising models (eg, television, print magazine).<sup>10</sup> Instagram, a popular image and video sharing platform with an estimated global userbase of over 2 billion allows users to upload and share photo or videos, represents one of those digital promotional

conduits.<sup>11</sup> For each post that is shared, there is image or video content, along with a user-provided text description. The content for each post can fall under any theme—including promotion and advertisements. Within the social media space, according to a 2021 Pew Research survey on social media usage, Instagram was the most popular social media platform for youth and young adults; 76% of survey participants between the age of 18 and 24 responded that they use Instagram daily.<sup>12</sup>

Much of the success of JUUL, a popular tobacco-derived ENDS product among youth and young adults, was predicated on its early adoption of Instagram for aggressive promotion tactics.<sup>13</sup> Previous research has demonstrated that the exposure to nicotine promotion on social media has profound effects on youth and young adults. One study demonstrated that visual exposure to ENDS content by adolescents was associated with an increased urge for initiation of ENDS products and ENDS promotion-related posts on Instagram were an effective means at attracting the attention of younger users.<sup>14,15</sup> Additionally, younger users were more likely to have a more positive perception of ENDS and a decreased perception of the risks of ENDS use following visual exposure to ENDS content.<sup>16</sup>

In response, social media companies implemented new guidelines to restrict promotional content of nicotine and tobacco products on their platforms. Specifically, Meta (the parent company of Facebook and Instagram) implemented community guidelines which prohibited the "promotion of certain regulated products, such as tobacco or vaping products" on Instagram.<sup>17</sup> However, research has shown that promotion of nicotine products, even including the sales of prohibited ENDS products on Instagram, continues via user-generated comments and via influencers who are often popular among youth and young adults.<sup>9,16,18</sup>

Given ongoing challenges regarding the process of premarket approval for NTN products, in tandem with the rising popularity and influence of social media platforms and their role in providing access to novel and emerging nicotine products, this study sought to identify and characterize the use of Instagram for product promotion and sales specific to NTN. Previous studies have identified marketing and sales promotion of naturally occurring ENDS products on Instagram.<sup>9</sup> However, there are no studies to our knowledge that have specifically examined both the promotion and offer for sale of novel NTN while also characterizing specific marketing and selling features that may lead to consumer harm. In this study, we focus on the United States as our country of interest and first collected data from the Instagram platform using common synthetic nicotine hashtags and then utilized a pretrained natural language processing (NLP) algorithm and manual content coding approach to identify and characterize posts specific to the promotion and sales of NTN. We conclude with some recommendations on how to enhance future tobacco regulatory science efforts to address challenges faced by social media-based promotion and unregulated sales online.

## Methods

To enable the identification and characterization of Instagram posts related to synthetic nicotine product promotion and sales, a multistep methodology was used with the following workflow: data collection, data cleaning, topic

modeling utilizing NLP, and human annotation for content analysis.

### Data Collection

The purpose of the data collection phase was to obtain an initial corpus of Instagram posts associated with synthetic nicotine user-generated discussions. All posts on Instagram are categorized into lists which can be queried based on specific “hashtags” (eg, “#”). Users can include hashtags as part of the text description on Instagram posts, and other users can query the specific hashtag and all public posts containing that specific hashtag will be returned through Instagram’s internal search engine.

We obtained a list of hashtags which were specifically focused to return Instagram posts related to synthetic nicotine products. In order to develop the list of hashtags, we examined previous research which included queried hashtags related to nicotine products. Additionally, authors conducted an exploratory manual search analysis for Instagram posts related to synthetic nicotine product promotion in the English language and using a US-based IP address and tabulated frequently utilized hashtags in each post to create a final list of study search hashtags. The complete list of 11 hashtags utilized were #puffbar, #tfnliquids, #tfnnicotine, #tfnvape, #tfnvapor, #tobaccofreedisposable, #tobaccofreenicotine, #tobaccofreevape, #syntheticnicotine, #syntheticnicotinevape, #zyn.

The dataset was time-restricted, representing all posts between September 2020 to the end of the data collection timeframe—December 2022. September 2020, 5 months before Puff Bar’s announcement of shifting to a synthetic nicotine formulation, was chosen as the starting point for data collection in this analysis to observe how synthetic nicotine product promotion and sales-related posts on Instagram evolved amidst this announcement and also to assess how changes in the regulatory environment potentially impacted online promotion and access.<sup>1</sup>

These date-restricted, returned Instagram posts following our hashtag querying were collected utilizing a web scraping tool. A web scraping tool functions as an automated software bot to extract all of the queried Instagram posts and their associated metadata ([Supplementary Material 1](#)). The web scraping tool utilized in this study was created using the computer programming language Python (Version 3.7.3), specifically using the BeautifulSoup library package (Version 4.9.3). This web scraping tool collected numerous variables and metadata of each Instagram post, including: the text description of the posts, the list of associated hashtags, the date and time, total number of likes, and total number of comments. These data were stored in a JavaScript Object Notation file.

### Data Cleaning

To facilitate subsequent NLP topic modeling after data collection, multiple minor modifications had to occur to prepare the dataset for optimized processing. In the text description of each Instagram post, the stop words “is,” “the,” and “and” were removed. Furthermore, punctuation marks, emojis/emoticons, numbers, and embedded hyperlinks were removed. This intermediary data cleaning step yielded only the relevant text content of each Instagram post to undergo NLP.

### NLP Topic Modeling

Given the volume of data returned by our search parameters, complete human annotation for each post was not feasible. Instead, we explored an NLP approach to categorize relevant Instagram posts into relevant thematic clusters and outliers.

We chose a pretrained NLP algorithm, the multilanguage Bidirectional Encoder Representations from Transformers (BERT). BERT is a self-supervised NLP model which comprehends the contextual meaning of specific phrases and written texts, itself having been trained by a large corpus of written text data as it appears on the Internet. Instead of conducting language processing in the standard left-to-right sequential manner, BERT evaluates each word in a phrase of text bidirectionally. The bidirectional functionality is aimed to increase the processing ability, accuracy, and contextual understanding of a given phrase. For this study, we utilized the specific multilanguage BERT model (BERTopic Version 0.6.0 with Python Version 3.7), which is trained to make the same contextual inferences across different language inputs. BERT’s utility has previously been validated across prior social media research, and a specific model evaluation study supported BERT’s functionality to conduct content analysis over other machine learning models.<sup>19–21</sup>

After the cleaned dataset was inputted into the multilanguage BERT model to undergo NLP, the outputted data included the sum total of posts sorted into a predetermined number of  $k$  clusters ( $k = 20$ ). Within each cluster, individual Instagram posts deemed thematically related by BERT are included. The remainder of Instagram posts which were not considered thematically related were automatically placed into an “outlier” pile by BERT. Furthermore, all posts which were not unique, such as duplicates of Instagram posts by the same user, were removed by BERT.

### Human Annotation

Clusters outputted by BERT were sorted by relevance to study aims of identifying promotion and sales of synthetic nicotine products by human annotators (NAS, TM, NL, AJC, and MCN). Though the BERT model only outputted clusters based on the text of the post, the outputted clusters themselves included links to the original Instagram posts—including text and image content. Therefore, after BERT clustering, all original text and hyperlinks were rematched to original posts in order to effectuate manual coding of posts for external link content or images as necessary. Given the scope of our study, posts including other discussions of NTN-related themes—such as advocacy, commentary on policies, and adverse events—were excluded. Furthermore, posts that were unrelated to synthetic nicotine products were excluded, even if it was relevant to promotion of other nicotine products. BERT clusters containing themes relevant to study aims were chosen for content analysis, and then a codebook was inductively developed.

For each post, authors collected multiple variables beyond whether or not the post was related to synthetic nicotine product promotion and sales. Variables collected were only from the primary post including the text and image content, so no comments or other interactions related to the post were used for data analysis purposes. Data regarding which specific synthetic nicotine product were being promoted, the presence of contact information or embedded hyperlinks for purported off-platform sales, specific discussions of cost of the product, the presence and classification of health warnings, images

associated with product being promoted or sold, and language the post was written in, were all included within the codebook. Additionally, an analysis of the interactive impact of each post was conducted via combining the “likes” and “comments” to yield a single interactive score. To conduct content analysis on foreign language posts, coders translated the text content of the Instagram post using Google Translate, though the primary focus of this study was on the United States and English-language posts. A full list of the codebook, including the inclusion and exclusion criteria, is shown in [Supplementary Table 1](#).

Each coder independently coded the dataset, and the consistency was evaluated. There was a high level of intercoder reliability ( $\kappa = 98\%$ ), with each inconsistency in coding being discussed together until an agreement was made over the proper classification.

## Results

### Data Overview

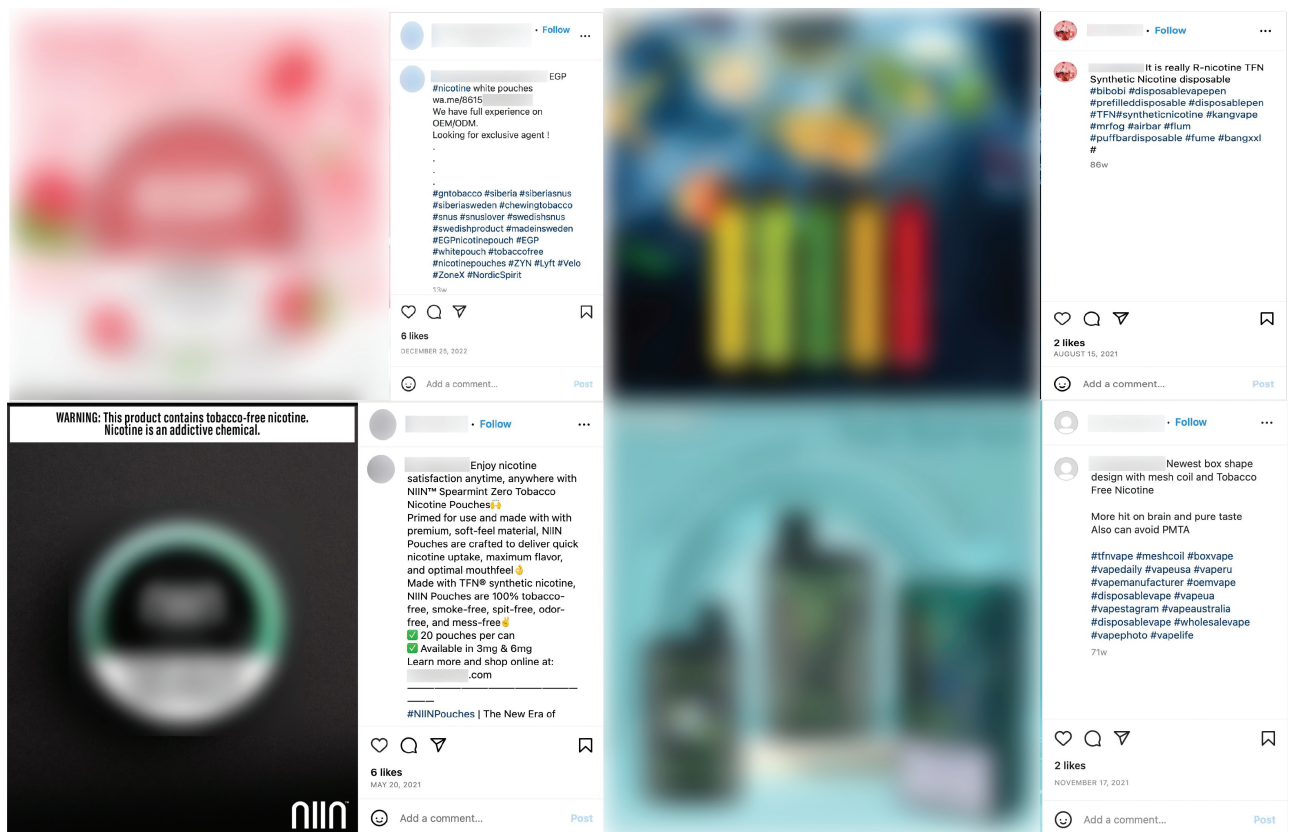
The dataset of date-restricted, hashtag-queried posts on Instagram collected by the web scraping tool was 8947 Instagram posts. Of these, 83.0% ( $n = 7425$ ) of the posts collected were unique Instagram posts which had undergone multilanguage BERT analysis. About 57.6% ( $n = 4274$ ) of the posts which had undergone multilanguage BERT analysis were sorted into the 20 topic clusters—therefore, 42.4% ( $n = 3151$ ) of Instagram posts that were inputted into BERT did

not exhibit thematic coherence or relation to other posts and were placed in a separate outlier pile.

Human coders selected topic clusters representing 72.9% ( $n = 3115$ ) of total clustered Instagram posts as potentially relevant to study aims to undergo manual human annotation based on outputted word frequencies and co-occurrences visualized from BERT topic outputs ([Figure 1](#)). Of these 3115 posts, 68.0% ( $n = 2219$ ) of Instagram posts were confirmed as related to the promotion and sales of synthetic nicotine products. Therefore, of the total corpus of Instagram posts which had undergone BERT analysis ( $n = 7425$ ), 28.9% ( $n = 2219$ ) were identified as related to the promotion and sales of NTN products. There was heterogeneity in the types of synthetic nicotine products: with the most common product being oral nicotine pouches with synthetic nicotine formulations ( $n = 1174$ , 52.9%), followed by synthetic nicotine ENDS ( $n = 679$ , 30.6%), flavored e-liquids ( $n = 313$ , 14.1%), synthetic nicotine salts ( $n = 32$ , 1.4%), and miscellaneous parts related to synthetic nicotine ENDS products such as atomizers and heating elements ( $n = 3$ , 0.14%). The full list of descriptive statistics can be found on [Table 1](#).

### Sales Characteristics

Of these promotion and sales-related posts, 5.8% ( $n = 129$ ) provided some sort of contact information in the text description for purported off-platform sales. The contact information provided was primarily phone numbers and usernames to contact sellers via third-party encrypted messaging platforms,



**Figure 1.** Examples of synthetic nicotine products promoted on Instagram: (a) Top left: Example of embedded hyperlink in oral nicotine pouch promotion. (b) Top right: Example of synthetic nicotine ENDS promotion, text related to R-nicotine content. (c) Bottom left: Example of synthetic nicotine health warning in oral nicotine pouch promotion. (d) Bottom right: Example of synthetic nicotine ENDS promotion, text related to avoiding regulation. ENDS = electronic nicotine delivery systems.

**Table 1.** Thematic Characteristics and Features of Instagram Posts

Instagram posts	<i>n</i>	% of promotion and sales-related posts
Total	8947	—
Analyzed via multilanguage BERT algorithm	7425	—
Promotion and sales related	2219	—
Related to oral nicotine pouches	1174	52.9%
Related to ENDS	679	30.6%
Related e-liquids	313	14.1%
Related to nicotine salts	32	1.4%
Related to other, miscellaneous components of ENDS	3	0.14%
Provided contact information	129	5.8%
Provided embedded hyperlinks	358	16.1%
Provided direct prices for products offered for sale	222	10.0%
Provided health warning label	1009	45.5%
Provided health warning label: general nicotine warning	618	27.9%
Provided health warning label: specific NTN language	391	17.6%

BERT = Bidirectional Encoder Representations from Transformers; ENDS = electronic nicotine delivery systems; NTN = non-tobacco nicotine product.

such as Telegram, WhatsApp, and WeChat. Occasionally, contact information was provided in the form of a specific Instagram username to send a private, direct message to for further instruction; however, this only represented a small amount of contact information content that was observed. Instead of contact information, 16.1% ( $n = 358$ ) of posts contained embedded hyperlinks within the text description for purported off-platform sales. About 10.0% ( $n = 222$ ) posts contained direct pricing for synthetic nicotine products being offered for sale in the form of some currency value in either the text description or image content of the post. The discussion regarding costs was also heterogeneous—costs were either related to a specific product being offered for sale, a sale price, or a price for a bundle of synthetic nicotine product being offered for sale.

### Temporal Analysis

There was temporal variation in the number of posts over the data collection timeframe (Figure 2). The largest number of posts related to promotion of nicotine pouches occurred between October 2020 and January 2021 (dates prior to clarification of FDA's regulatory authority over NTNs), whereas the largest number of posts related to synthetic nicotine ENDS promotion occurred between October 2022 and November 2022 (dates following after FDA's regulatory authority was extended and also occurring after the May 14, 2022, deadline to submit a premarket application to FDA for synthetic nicotine products and also the July 13, 2022, date to remove NTNs from the market if they were not approved). Furthermore, the largest number of posts related to promotion of synthetic nicotine e-liquids and flavors occurred between April 2022 and May 2022 (dates immediately preceding the due date for

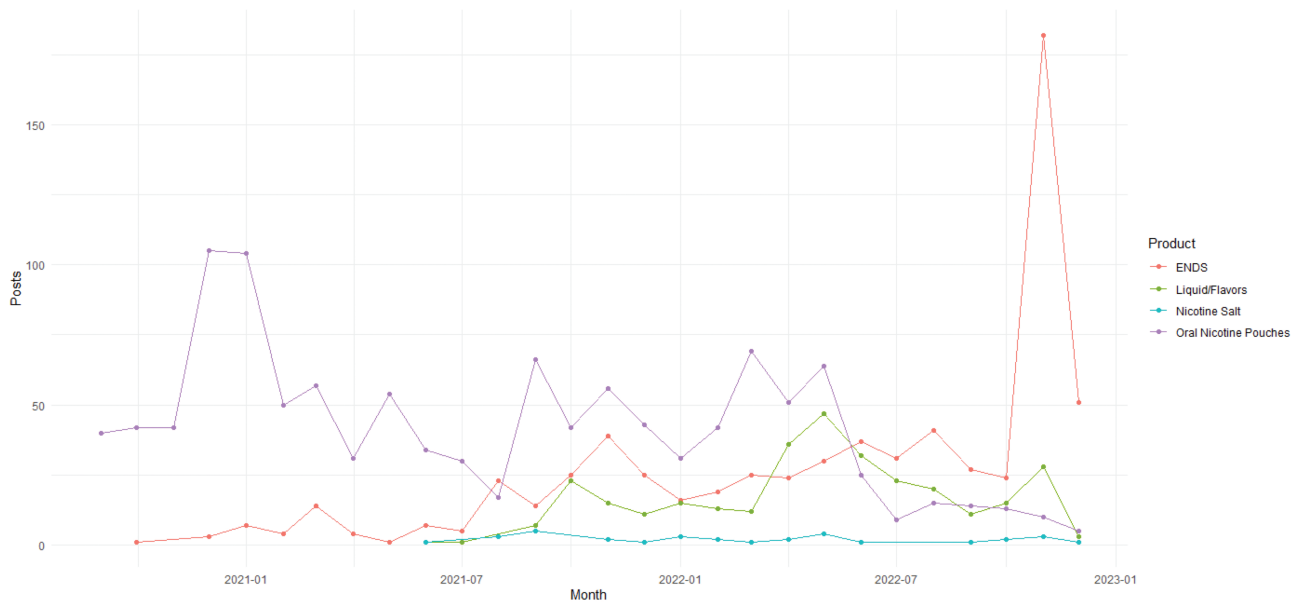
premarket application). The macrorends depicted in Figure 2 illustrate an increase in synthetic nicotine ENDS products and synthetic nicotine e-liquids after the announcement of Puff Bar switching to a synthetic nicotine formulation in March 2021. Furthermore, there was no immediate decrease of posts following the announcement by the FDA to regulate synthetic nicotine products in March 2022; however, there was a decrease in oral nicotine pouch and e-liquid posts in the following months, but there was no overall decrease in posts related to synthetic nicotine ENDS.

### Thematic Characteristics

The text and image content of Instagram posts related to the promotion and selling of synthetic nicotine products shared some similar characteristics and followed a consistent theme (see Figure 1 for de-identified examples). The image content of these posts usually followed specific, predictable patterns: the image content was most often the NTN product being promoted or offered for sale. However, the manner in which the product was presented varied, with some posts utilizing stock photos of the product and others showing a picture of the product sitting on a table or in the purported seller's hand. This theme held consistent across the range of NTN products being promoted and offered for sale. In addition, some posts contained text that was embedded within the image portion of the post—these often included the associated health warning labels, embedded hyperlinks, or a display of cost. Though embedded hyperlinks were often seen as displayed within the image content of the post, there was no post with contact information in the image content; instead, this appeared in the text of the post.

The text content of the posts varied, but were thematically related to product promotion and sales topics as outputted by the BERT algorithm. Generally, the text content of each post included the name or type of NTN product being sold, or a brief description advertising the product that was being sold. In some examples, unique characteristics of synthetic nicotine products related to *R*-nicotine content and ability to circumvent FDA regulations were included as part of the promotional message (Figure 1b and d). However, for many posts, the brief promotion message was only used, and no explicit mention of the product that was shown in the image was in the text content. Notably, all of the posts contained numerous hashtags no matter the length of the description. These hashtags often included multiple specific NTN-related hashtags, along with some general nicotine, tobacco, ENDS, and nicotine pouch-related hashtags. Cost discussion, contact information, and embedded hyperlinks were also generally in the text, rather than the image portion of each Instagram post (Figure 1a and b). Analyses of posts from different languages also conformed to these similar patterns in both the image and text content with no major exceptions.

Furthermore, 45.5% ( $n = 1009$ ) of Instagram promotion and selling posts included some form of health warning labeling, indicating that the majority of posts related to the promotion and sales of synthetic nicotine products did not include any health warning labeling. About 61.3% ( $n = 618$ ) of posts which contained health warning labeling included general warnings about nicotine being an addictive chemical, and 38.8% ( $n = 391$ ) of health warning labels were specifically related to synthetic nicotine (Figure 1a). There was variance in the nature of health warning labeling as well—some of the health warning labels utilized federally



**Figure 2.** Temporal analysis of synthetic nicotine product posts.

standardized language incorporated within the image content of the post. Others advertised the products being sold in a way to explicitly show the health warning label on the product itself as part of the product image. Some posts did not include any health warning label on the image; rather, user-generated language was written on the text description of the post itself. These were inconsistent and did not use standardized language. For instance, example posts included the language: “warning: 21+” or “warning: this product contains nicotine.”

There was variation in the interaction score for posts related to selling. The highest number of interactions observed in a single post was 1019 combined likes and comments, while the lowest had no comments or interactions. In the entire dataset, the mean interaction score of the posts was 29.9. The mean interaction score of the top 5, top 10, and top 50 posts with the most interaction was 845.2, 792.4, and 551.7, respectively. Interactions varied by type of product being sold as well. The highest median number of interactions were for oral nicotine products ( $n = 19$ ), whereas the median interaction score for ENDS and e-liquids were the same ( $n = 5$ ). For nicotine salts, the interaction score was slightly lower ( $n = 4$ ).

## Discussion

The results of this study indicate that promotion and selling of NTN products is widely occurring on Instagram when searching for common hashtags associated with NTNs. From a total of 7425 posts which had undergone BERT analysis, 28.9% ( $n = 2219$ ) were directly related to the promotion and sales of NTN products. Approximately half of sales (52.9%,  $n = 1174$ ) were for oral nicotine products, followed by ENDS products (30.6%,  $n = 679$ ), e-liquids (14.1%,  $n = 313$ ), and salts (1.4%,  $n = 32$ ). A subset of the total of promotion and sales-related posts included contact information leading to encrypted third-party messaging platforms for the purported buying and selling of products (5.8%,  $n = 129$ ), and embedded hyperlinks leading to third-party websites for direct-to-consumer selling (16.1%,  $n = 358$ ). Finally,

oral nicotine products had the greatest number of median interactions ( $n = 19$ ) in relation to the ENDS products ( $n = 5$ ), and e-liquids ( $n = 5$ ).

Currently, the FDA requires any covered tobacco product, including products with NTN formulations, to display the following health warning: “WARNING: This product contains nicotine. Nicotine is an addictive chemical.”<sup>22</sup> Despite this, the majority of promotion and selling-related Instagram posts (54.6%,  $n = 1211$ ) included no form of health warning labels or descriptions. Of the posts which included health warnings, a subset (17.6%,  $n = 391$ ) included specific descriptors related to synthetic nicotine and NTN products, stating that products were “tobacco-free nicotine,” “non-tobacco nicotine,” or “synthetic nicotine.” Currently, the FDA has permitted health warning labels to add these descriptors to the standardized, mandated health warning label. However, recent research examining the response of young adults to these descriptors in ENDS and oral nicotine pouches showed that these descriptors were associated with participants believing these products were less harmful and less addictive.<sup>23</sup>

Prior research has shown that visual exposure to promotional ENDS content is correlated with increased desire, positive sentiment, and willingness to use nicotine products.<sup>14,15</sup> Previously, studies have also indicated that there is a growing and diverse online digital marketplace for the proliferation of ENDS and flavored e-liquids designed to appeal to youth and young adult consumers, including specifically on Instagram.<sup>9</sup> Our findings demonstrate that NTN products are no exception, comprising a total of 44.7% of all promotion and sales-related posts we reviewed. In 2021, Puff Bar, which manufacturers disposable, NTN ENDS devices, was labeled as the most popular ENDS product among middle school- and high school-aged students in the United States, even surpassing JUUL, concomitantly with studies finding that it has been promoted on popular social media platforms like TikTok.<sup>5,24,25</sup>

Existing FDA regulations are aimed at halting the production and distribution of NTNs unless they receive premarket approval. As of 2022, all NTN products must obtain a

premarket authorization; without it, no products are allowed to be promoted or offered for sale. After obtaining premarket authorization, proposed NTN products then undergo review for authorization to promote and sell their products. FDA enforcement of these regulations toward manufacturers and retailers promoting and selling NTN products without the appropriate authorizations may be given warning letters. Puff Bar has been issued multiple warning letters by the FDA for the manufacture and sale of unauthorized nicotine and NTN ENDS, including various flavored products. Notably, since Puff Bar's initial announcement to begin offering an NTN formulation, the number of Instagram posts we observed in our dataset related to the promotion and sale of NTN ENDS and NTN e-liquids experienced an overall increase (Figure 2).

Furthermore, our findings evidence exposure to content actively promoting and offering direct-to-consumer sale of NTN products throughout various study time periods, including after clarification of FDA's regulatory oversight authority of these products, imposition of requirements for manufacturers to submit premarket applications, and issuance of warning letters to manufacturers and retailers. This represents a shift from sellers turning to Instagram to simply promote NTN products to increase appeal, to more overt acts of illegal marketing and sale via unregulated social media channels that are not subject to robust enforcement action compared to retail storefronts. This circumvention of federal and possibly state law is likely made possible by a combination of anonymity afforded by social media platforms, an unregulated e-commerce ecosystem that enables social media sellers to advertise to large audiences and then enter into sales transactions using third-party encrypted messaging platforms or a third-party website, and a lack of proactive monitoring and enforcement actions by platforms to remove illegal content and prohibited products. Our findings also align with a recent independent report by the Reagan-Udall Foundation that concluded that enforcement of illegal ENDS that remain on the market needs better prioritization by FDA and that failure to remove products may be misconstrued by sellers as authorizing lawful marketing.<sup>26</sup>

Results of this study show that Instagram is a modality of potentially unregulated or unapproved sales of NTN products, including via high-risk transaction processes such as buying directly from a seller via a third-party encrypted messaging applications or third-party websites. General characteristics of marketing and availability align with prior findings that have detected popular synthetic nicotine brands actively being offered for sale on major online e-commerce sites, though user-generated content on social media platforms are generally subject to fewer restrictions compared to these marketplaces.<sup>27</sup> Additionally, research is needed to better understand how oral nicotine pouches are bought and sold on social media; with much of the research focused on ENDS products, our findings reveal that there is a significant presence of oral synthetic nicotine formulations being promoted and offered for sale on Instagram; significantly, our findings suggest posts promoting these oral nicotine pouches in general have greater median interaction levels of interaction by users than ENDS and e-liquid posts have. Future tobacco regulatory science efforts must be adaptive to the introduction of novel and emerging tobacco and nicotine-derived products and the growing and evolving digital landscape that promotes their use and sale through unregulated direct-to-consumer channels.

## Multilanguage BERT Subanalysis

Though the focus of this study was on the United States and English-language posts, notably, Instagram posts related to the promotion and offering for sale of NTNs were found in multiple languages. In total, posts from a total of 15 different languages were observed including: Arabic, Czech, Danish, English, Finnish, French, Italian, Mandarin, Polish, Portuguese, Russian, Swedish, Turkish, and Ukrainian. The most frequent languages observed were English, which comprised of 91.9% of the dataset ( $n = 2041$ ) and Polish (7.0%,  $n = 155$ ). The remainder of languages each represented a small volume of the total posts reviewed—the sum total of posts in all other languages besides English and Polish made up only 1.0% ( $n = 23$ ) of the dataset. Though we could not discern from available metadata what geographical regions these posts originate from, we note there is limited research on country-specific regulations for NTN products. This points to the need for future studies that specifically examine cross-border marketing of NTNs and other novel nicotine products on social media, the impact of national tobacco control policies on platform content moderation practices, and the availability of prohibited NTN, tobacco, and nicotine products that may be purchased online and imported from other less-restrictive markets.

## Limitations

This study has certain limitations. This study limited data collection to a set of common NTN-related hashtags following manual exploratory queries that may not be representative of all discussions, promotion, and sale of synthetic nicotine products. Future studies should use a more comprehensive lists of hashtags associated with NTN brands, product types, and popular NTN-related Instagram accounts (including social media influencers) to assess the full scope and magnitude of this activity on Instagram. Data from this study also only included text from Instagram posts and did not include analysis of user comments in response to posts. Future studies should also analyze user comments to assess intent to purchase or enter into a transaction with a seller. Importantly, this study did not attempt, and it is also extremely difficult to accurately assess the effects of these NTN-specific Instagram posts on use of or risk perceptions about NTNs. Additional research specific to self-reported user behavior on social media platforms is needed to better assess this relationship. Furthermore, though NLP is an effective means at thematically categorizing text content on social media discussions into specific topic clusters, it cannot analyze image content. Therefore, if any posts included information related to promotion of synthetic nicotine products only on the image content, but not the text description or accompanying hashtags, these would have been discarded as outliers by the multilanguage BERT algorithm. Incorporation of image classification using deep learning approaches in tandem with NLP specific to text or the use of multimodal models in future studies may improve the ability to process and contextually analyze a greater variety of Instagram posts. Finally, human coders did not actually attempt to purchase synthetic nicotine products offered for sale through embedded hyperlinks, nor did coders directly message potential sellers on the third-party encrypted messaging platforms to inquire about purchasing. These decisions were made due to the need for compliance with local laws and regulations. Therefore,



authors cannot verify with full certainty that products that are purported to be offered for sale were actually attempted to be sold.

## Supplementary Material

Supplementary material is available at *Nicotine and Tobacco Research* online.

## Funding

This study was funded by the University of California Tobacco-related Disease Research Program award T32IP4788.

## Declaration of Interests

MCN, ZL, TM, and TKM are employees of the minority owned small business S-3 Research LLC. S-3 Research is a company that was previously funded by the National Institutes of Health—National Institute of Drug Abuse through a Small Business Innovation and Research contract for opioid-related social media research and technology commercialization. Author reports no other conflict of interest associated with this manuscript.

## Author Contributions

Neal Shah (Conceptualization [lead], Data curation [supporting], Formal analysis [lead], Investigation [lead], Methodology [lead], Software [equal], Validation [lead], Writing—original draft [lead], Writing—review & editing [lead]), Zhuoran Li (Data curation [lead], Investigation [equal], Methodology [lead], Software [lead], Visualization [equal]), Tiana McMann (Investigation [equal], Methodology [equal], Validation [equal], Writing—review & editing [supporting]), Alec Calac (Investigation [equal], Methodology [equal], Validation [equal]), Nicolette Le (Investigation [equal], Methodology [equal], Validation [equal]), Matthew Nali (Investigation [equal], Methodology [equal], Validation [equal]), Raphael Cuomo (Data curation [equal], Visualization [equal]), Joshua Yang (Conceptualization [equal], Writing—original draft [equal], Writing—review & editing [equal]), and Tim Mackey (Conceptualization [equal], Investigation [equal], Methodology [equal], Supervision [lead], Writing—review & editing [equal])

## Data Availability

De-identified data are available by request.

## Disclaimer

The opinions expressed are those of the authors alone.

## Patient Consent for Publication

Not applicable.

## Ethical Compliance

Not applicable/not required for this study. All information collected from this study was from the public domain and the study did not involve any interaction with users. Any user

identifiable information was aggregated and removed from the study results.

## References

- Jordt S. Synthetic nicotine has arrived. *Tob Control*. 2021;32(e1):113–e117.
- Nwosu CG, Crooks PA. Species variation and stereoselectivity in the metabolism of nicotine enantiomers. *Xenobiotica*. 1988;18(12):1361–1372.
- Hellinghausen G, Lee JT, Weatherly CA, et al. Evaluation of nicotine in tobacco-free-nicotine commercial products. *Drug Test Anal*. 2016;9(6):944–948.
- Morean ME, Bold KW, Davis DR, et al. “Tobacco-free” nicotine pouches: risk perceptions, awareness, susceptibility, and use among young adults in the United States. *Nicotine Tob Res*. 2023;25(1):143–150.
- Park-Lee E, Ren C, Sawdey MD, et al. Notes from the field: e-cigarette use among middle and high school students—National Youth Tobacco Survey, United States, 2021. *MMWR Morb Mortal Wkly Rep*. 2021;70(39):1387–1389.
- Food and Drug Administration. *Regulation and Enforcement of Non-Tobacco Nicotine Products*. Washington D.C., United States of America: United States Food and Drug Administration; 2023.
- American Lung Association. *State of Tobacco Control—2023 Report*. United States of America: American Lung Association; 2023.
- Becker TD, Rice TR. Youth vaping: a review and update on global epidemiology, physical and behavioral health and clinical considerations. *Eur J Pediatr*. 2022;181(2):453–462.
- Shah N, Nali M, Bardier C, et al. Applying topic modelling and qualitative content analysis to identify and characterize ENDS product promotion and sales on Instagram. *Tob Control*. 2021;32(e2):e153–e159.
- Duan Z, Henriksen L, Vallone D, et al. Nicotine pouch marketing strategies in the USA: an analysis of Zyn, On! And Velo. *Tob Control*. 2022. doi:10.1136/tc-2022-057360
- Statista. Number of Instagram Users Worldwide from 2020 to 2025. <https://www.statista.com/statistics/183585/instagram-number-of-global-users/>. Accessed April 13, 2023.
- Brooke A, Monica A. Social Media Use in 2021. Pew Research Center; 2021. <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/>. Accessed April 13, 2023.
- Czaplicki L, Tulsani S, Kostygina G, et al. #Toolittleoolate: JUUL-related content on Instagram before and after self-regulatory action. *PLoS One*. 2020;15(5):e0233419.
- King AC, Smith LJ, Fridberg DJ, et al. Exposure to electronic nicotine delivery systems (ENDS) visual imagery increases smoking urge and desire. *Psychol Addict Behav*. 2016;30(1):106–112.
- Klein EG, Czaplicki L, Berman M, Emery S, Schillo B. Visual attention to the use of #Ad versus #Sponsored on e-cigarette influencer posts on social media: a randomized experiment. *J Health Commun*. 2020;25(12):925–930.
- Vassey J, Valente T, Barker J, et al. E-cigarette brands and social media influencers on Instagram: a social network analysis. *Tob Control*. 2021;32(e2):e184–e191.
- Facebook. Instagram: Community Guidelines. <https://facebook.com/help/instagram/477434105621119>. Accessed April 13, 2023.
- Lee J, Suttiratana SC, Sen I, Kong G. E-cigarette marketing on social media. *Curr Addict Rep*. 2023;10:29–37.
- Chintalapudi N, Battineni G, Amenta F. Sentimental analysis of COVID-19 tweets using deep learning models. *Infect Dis Rep*. 2021;13(2):329–339.
- Dadgar S, Neshat M. A novel hybrid multi-model deep learning for detecting hashtag incongruity on social media. *Sensors (Basel)*. 2022;22(24):9870.
- Bello A, Ng SC, Leung MF. A BERT framework to sentiment analysis of tweets. *Sensors (Basel)*. 2023;23(1):506.

22. Food and Drug Administration. "Covered" Tobacco Products and Roll-Your-Own Cigarette Tobacco Labeling and Warning Statement Requirements. Washington D.C., United States of America: United States Food and Drug Administration; 2023.
23. Keller-Hamilton B, Curran H, Stevens EM, et al. Effects of "tobacco free" language in warning labels on perceptions of electronic cigarettes and nicotine pouches among young adult men: a randomized trial. *Subst Use Misuse*. 2023;58(10):1302–1306.
24. Morales M, Fahrion A, Watkins SL. #NicotineAddictionCheck: puff bar culture, addiction apathy, and promotion of e-cigarettes on TikTok. *Int J Environ Res Public Health*. 2022;19(3):1820.
25. TechCrunch. TikTok Is Being Used by Vape Sellers Marketing to Teens. <https://techcrunch.com/2021/01/26/tiktok-is-being-used-by-vape-sellers-marketing-to-teens/>. Accessed August 1, 2023.
26. Silvis L, Axelrad J, Flanagan K, et al. *Operational Evaluation of Certain Components of FDA's Tobacco Program: A Report of the Tobacco Independent Expert Panel*. Washington D.C., United States of America: Reagan-Udall Foundation; 2022.
27. Ramamurthi D, Chau C, Lu Z, et al. *Marketing of "Tobacco-Free" and "Synthetic Nicotine" Products: White Paper—Stanford Research into the Impact of Tobacco Advertising*. Palo Alto, CA, United States of America: Stanford University School of Medicine; 2023.