Emerging Trends in Drugs, Addictions, and Health 4 (2024) 100157

Contents lists available at ScienceDirect

Emerging Trends in Drugs, Addictions, and Health

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

Ozempic (*Glucagon-like peptide 1 receptor agonist*) in social media posts: Unveiling user perspectives through Reddit topic modeling

Seraphina Fong [a], Alessandro Carollo [a], Lambros Lazuras [b], Ornella Corazza [a,c], Gianluca Esposito [a,c,*]

[a] *Department of Psychology and Cognitive Science, University of Trento, 38068 Rovereto, Italy*
[b] *School of Sport & Exercise Science, University of Lincoln, LN6 7TS Lincoln, United Kingdom*
[c] *School of Life and Medical Sciences, University of Hertfordshire, AL10 9AB Hatfield, United Kingdom*

ARTICLE INFO

Keywords:
Ozempic
Topic modeling
Natural language processing
Reddit
BERTopic
Addiction science

ABSTRACT

Semaglutide, a Glucagon-like peptide 1 (GLP-1) receptor agonist marketed under the brand name Ozempic, is originally prescribed for diabetes treatment and obesity management. However, healthy individuals without a medical cause use Ozempic without medical supervision to improve their physical appearance - a trend that has proliferated through social media, news coverage, and relevant celebrity endorsements. Thus, exploring social media posts can provide insight into understanding individuals' experiences, beliefs, motivation, as well as misconceptions about Ozempic. To do so, this study utilizes BERTopic, a natural language processing approach for topic modeling, to analyze 46,491 Reddit posts from three subreddits (r/ozempic, r/ozempicforweightloss, r/semaglutide) dated between April 2019 and December 2023. The analysis revealed various discussion topics, including using Ozempic for weight loss, dosaging, insurance denial due to lack of a diabetes diagnosis, weight loss tracking, and side effect management. Overall, the overarching theme centered on the off-label use of Ozempic and its GLP-1 agonist counterparts for weight loss purposes. Moreover, awareness on the health hazards associated with the off-label and unsupervised use of Ozempic as an image enhancer do not frequently appear in the social media discussions. These findings, supported by a dynamic topic modeling analysis, offer ecological insights into the experiences and opinions of community members in Ozempic-related subreddits, reinforcing the growing evidence of the drug's increasing popularity for weight management as well as the role played by social media. The study also shows how information campaigns about the health risks associated with the off-label use of Ozempic by healthy individuals without a medical cause may help counterbalance the lack of risk awareness detected in social media discussions.

1. Introduction

Glucagon-like peptide 1 (GLP-1) receptor agonists semaglutide (brand name Ozempic) is a drug class originally used for the treatment of diabetes and for the medical management of obesity (Han, et al., 2024; Suran, 2023). As a general principle, GLP-1 receptor agonists are designed to imitate the GLP-1 hormone, which boosts insulin production when blood sugar levels are high (Khawagi et al., 2023; Suran, 2023). The GLP-1 hormone additionally slows down gastric emptying, extending the sensation of fullness after meals, reduces appetite, and therefore acts as an appetite suppressant (Khawagi et al., 2023; Suran, 2023).

As a result of its weight-loss properties, the use of Ozempic outside of clinical and prescribed contexts has risen, particularly among individuals seeking to enhance their physical appearance without medical oversight or justification (Han et al., 2024). This practice carries risks, as GLP-1 receptor agonists can lead to adverse side effects such as gastrointestinal issues, inflammation, hypoglycemia, elevated heart rate, myocardial infarction, ulcerative lesions, and atherosclerosis (Bartal et al., 2024; Lopes et al., 2023). Moreover, comprehensive understanding of the function of and the potential long-term adverse effects of GLP-1 receptor agonists, especially when used at higher doses for weight loss, is still lacking (Myerson and Paparodis, 2024).

The emergence of substances like Ozempic that are used off-label, without a medical cause and medical supervision to enhance body image has expanded the literature on human enhancement drugs, such

* Correspondence author.
E-mail address: gianluca.esposito@unitn.it (G. Esposito).

https://doi.org/10.1016/j.etdah.2024.100157
Received 27 June 2024; Received in revised form 8 August 2024; Accepted 30 August 2024
Available online 3 September 2024
2667-1182/© 2024 Published by Elsevier Ltd on behalf of International Society for the Study of Emerging Drugs. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

as anabolic steroids and peptide hormones used by sportspeople (Carollo et al., 2024), and led to the conceptualization of the image- and/or performance-enhancing drugs (IPEDs) classification (Simonato et al., 2022).

The off-label use of Ozempic for body image enhancement purposes has recently been growing exponentially, driven by social media, news outlets, and celebrity influence (Han et al., 2024). This increase in off-label use has also contributed to supply shortages for individuals with diabetes (Bartal et al., 2024; Basch et al., 2023). Han et al. (2024) performed a Google Trends analysis, revealing that searches for Ozempic and other GLP-1 receptor agonists have grown exponentially over the past five years (2018–2023) in the United States. The average relative search value (RSV) index of the search term "Ozempic" increased from 10.25 RSV in December 2021 to 84.75 RSV in February 2023 (where 100 RSV indicates the highest popularity in searching a given term during a particular time and location). They further note that the hashtag "#Ozempic" on both TikTok and Instagram (with over 690 million views and 50,000 posts, respectively) features users' successful experiences with using Ozempic for weight loss. Their findings indicated an increasingly prevalent and positive public interest in the drug class, as well as the impact that social media and celebrities have on the perception of it.

Social media appears to have a significant role in boosting the off-label use of Ozempic. For this reason, several studies have utilized Reddit and other social media platforms to assess the usage and impact of Ozempic. For instance, through a descriptive analysis of 100 TikTok videos tagged with "#Ozempic", Basch et al. (2023) found that >50 % of the videos included mention of one taking or planning to take Ozempic, as well as mentioning its use for weight loss. Moreover, results showed that the 100 videos (with 86 of them uploaded by non-professional consumers) were collectively watched almost 70 million times. In line with Han et al. (2024)'s findings, this once again highlights the impact that social media can have on increasing and shaping public interest and perceptions of the drug. In another study, Quddos et al. (2023) analyzed Reddit using a machine learning approach to classify and quantify discussions regarding user experiences with reduced alcohol consumption after Ozempic usage. As the drug class has also been associated with triggering depression, Arillotta et al. (2023) examined social media users' comments (from Reddit, Youtube, and Tiktok) to better understand their perceptions of Ozempic's effect on mood, mental health, and behavior. A more recent study by Arilotta et al. (2024) used a mixed-methods (quali-quantitative and thematic analyses) approach to analyze online Reddit discussions posted from December 2019 to June 2023 to understand the effects of GLP-1 receptor agonists on substance-related and behavioral addictions (e.g., alcohol, psychostimulants, compulsive shopping). The authors' thematic analysis was conducted by grouping Reddit comments based on keywords and analyzing these groups using a specialized software that leverages ChatGPT, followed by manual revision of the generated themes and biases. Subsequently, the themes were further refined with ChatGPT 3.5, and analyzed through a phenomenological qualitative media approach.

In contrast, the current study aims to use a BERTopic (Grootendorst, 2022), a natural language processing-based approach), to conduct a thematic analysis of Ozempic-related Reddit data. In this way, we aim to outline the main topics of discussion in posts from Ozempic-related forums on Reddit to gain general insights from naturalistic online discourse between 2019 and 2023. In the current work, we implement topic modeling via BERTopic and subsequently conduct dynamic topic modeling to understand how the frequency of discussions about the identified topics varies over time (Blei and Lafferty, 2006). This study seeks to contribute to the scientific literature by exploring how Ozempic is used or misused by the general public on Reddit, aiming to provide policymakers with insights necessary for understanding and regulating its use.

2. Method

The overall methodology of the present work follows that of Fong et al. (2024).

2.1. Data collection

This study was approved by the University of Trento Ethical Committee (2024–40 ESA). The data collected were anonymously posted on Reddit.

The present work utilizes data sourced from Reddit, an interactive online social news platform. Reddit enables registered users to establish or participate in >100,000 online communities, known as "subreddits", that are dedicated to specific interests (Anon., Reddit Homepage, 2024). These subreddits serve as forums for users to exchange ideas and engage in discussions in line with a given subreddit's theme (Pont-Fernandez et al., 2023; Smith et al., 2021). For instance, the subreddit "r/gaming" (with "r/" denoting a subreddit) offers a platform for individuals to participate in discussions about gaming.

Reddit was selected as the primary data source for this study for two primary reasons. Firstly, unlike Twitter, where users are confined to a 280-character limit, Reddit allows individuals to express more detailed opinions or participate in nuanced discussions across diverse subjects. Secondly, Reddit's platform offers users anonymity, which can thereby provide a rich ecological source of human opinions, attitudes, and experiences (Wanchoo et al., 2023). Previous research has successfully used Reddit data to investigate public opinions on several topics of discussion, such as COVID-19 vaccines (Melton et al., 2021) or ChatGPT (Xu et al., 2024).

To scrape Reddit data, earlier research (e.g., Wanchoo et al., 2023; Yao et al., 2023) employed the Pushshift Reddit application programming interface (API) alongside the PRAW and/or PSAW Python packages (Baumgartner et al., 2020) for Reddit data extraction. However, changes in Reddit API in June 2023 resulted in access to the Pushshift API being restricted to approved Reddit moderators. Consequently, the Reddit data utilized in this study was sourced from a comprehensive public dataset encompassing posts and comments collected by two approved Reddit moderators from the top 40,000 subreddits (Watchful1, n.d.). The dataset includes posts and comments ranging from June 2005 to December 2023. Within the overall dataset, we identified 3 subreddits of interest for discussions around ozempic: r/ozempic, r/ozempicforweightloss, r/semaglutide. Only specific subreddits related to Ozempic were selected to ensure high homogeneity and low noise in the dataset. The title, main body ('selftext') of the posts, and the dates when they were published were extracted. Prior to data preprocessing, the dataset consisted of $N = 52{,}773$ posts.

2.2. Data preprocessing

Corresponding titles and main bodies of the posts were also merged as users would often post pictures (e.g., weight loss progress photos) in place of text. Moreover, photos (i.e., links to photos with the *.png* file format as indicated by Reddit), empty posts (null or containing empty strings), emojis, and those indicated as [removed] or [deleted] were discarded in order to only retain meaningful text.

All remaining posts were converted to lowercase. Additional preprocessing steps such as tokenization or stop word removal were not conducted to preserve each post's original structure. This approach enables transformer-based topic models to consider the contextual relationships among words within sentences and therefore results in more accurate and representative embeddings (i.e., numeric representations of text posts) (Grootendorst, n.d-c; Yao et al., 2023).

In order to implement dynamic topic modeling (see Section 2.3.1 for more details), the dates on which each post was created were converted from unix timestamps to Coordinated Universal Time (UTC).

After preprocessing the data, the dataset contained $n = 46{,}491$ posts

spanning from April 2019 to December 2023.

2.3. Topic modeling

The present work employs a Natural Language Processing (NLP) technique known as topic modeling to identify the primary topics discussed in Ozempic-related subreddits. The topic modeling approach utilizes statistical modeling in order to discern patterns within a corpus of unstructured text data and subsequently derive topics of discussion (Egger and Yu, 2022). We specifically use BERTopic (Grootendorst, 2022) as previous research has demonstrated its effectiveness in generating interpretable topics from unlabelled Reddit posts (Choi and Jang, 2023; Ng et al., 2023; Pleasants et al., 2023; Yao et al., 2023).

BERTopic leverages embeddings created from Bidirectional Encoder Representations from Transformers (BERT) (Devlin et al., 2018) and class-based term frequency-inverse document frequency (c-TF-IDF) to cluster posts into semantically similar topics (Ng et al., 2023). The BERT embeddings allow BERTopic to produce meaningful topic representations by considering the context and meaning of words (Ng et al., 2023; Grootendorst, 2022). C-TF-IDF determines the importance of a word to a cluster of posts and calculates their frequency to generate the topics (Yao et al., 2023).

Python version 3.10.12 was used for preprocessing and training BERTopic. The model was trained using the BERTopic python library (version 0.16.0) (Grootendorst, 2022). An overview of the hyperparameters set for the model has been included in Table S1 in the supplementary materials. Most of the default out-of-the-box values were retained. However, in order to make the results reproducible by avoiding stochastic behavior (i.e., generating different results each time the model is run), we added a random state of 42 for the Uniform Manifold Approximation and Projection (UMAP) model. Additional stop words were also removed during tokenization to reduce the amount of noise added to the topics (i.e., 'http', 'https', 'amp', 'com') (Grootendorst, n.d-b).

BERTopic extracts the topics and represents them with a set of relevant keywords and representative posts. These extracted topics are automatically labeled Topic 0 to *N*, where *N* is the number of topics generated minus 1. In order to generate more explanatory and descriptive topic labels, the keywords and samples of representative posts were given to ChatGPT 3.5 with the following prompt (Grootendorst, n.d.-a):

"I have a topic that contains the following documents:
[DOCUMENTS]
The topic is described by the following keywords: [KEYWORDS]
Based on the information above, extract a short topic label in the following format: topic: <topic label>"

2.3.1. Dynamic topic modelling

To understand the possible fluctuations and temporal fluctuations in the frequency of the discussions about specific topics on Reddit, we implement Dynamic Topic Modelling (DTM) (Blei and Lafferty, 2006; Grootendorst, M., 2022; Jung et al., 2023).

A c-TF-IDF representation is computed for each topic and timestamp using BERTopic (version 0.16.0)'s *topics_over_time* module subsequent to model training. In order to increase computational efficiency given limited computational resources, the number of bins generated when computing the topic representations was set to 10.

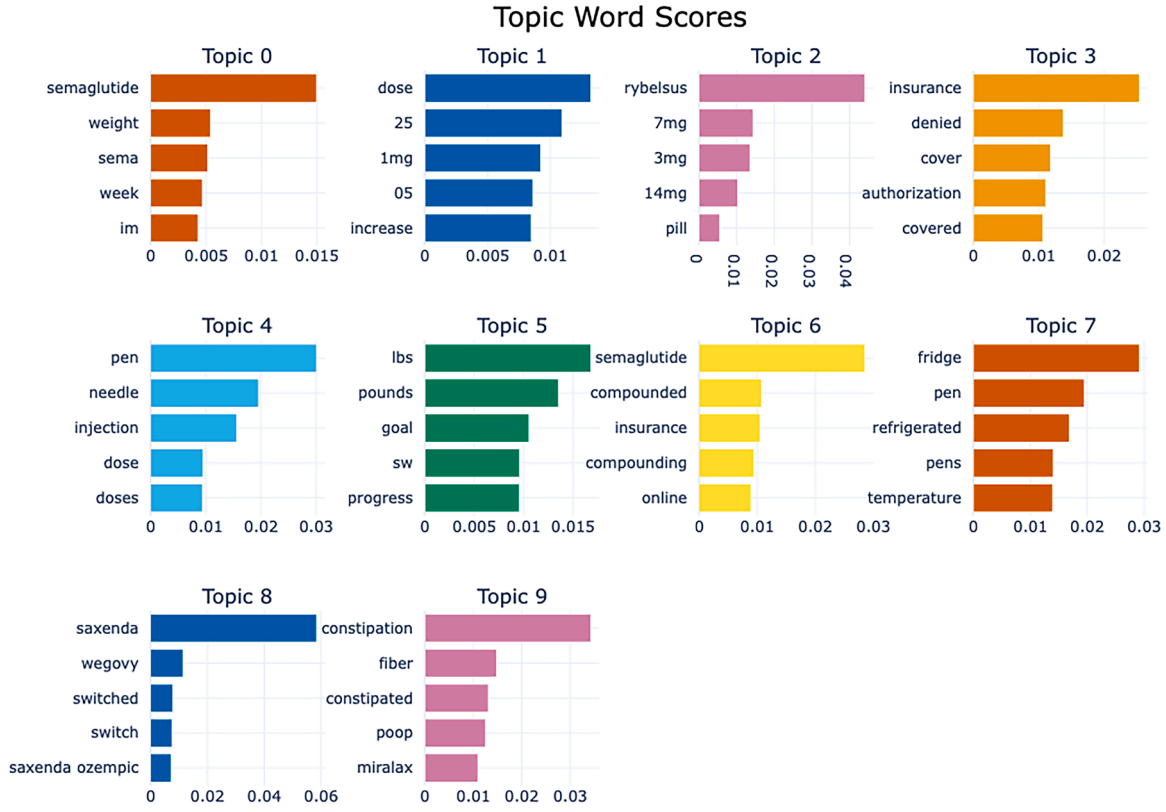

Fig. 1. Top 5 representative words based on class-based term frequency-inverse document (c-TF-IDF) scores for the top 10 topics generated from the dataset. The higher the c-TF-IDF score, the more relevant the word is in the context of a given topic. The ChatGPT-generated labels for each of the topics are as follows: Topic 0 (Semaglutide and Weight Loss Experiences), Topic 1 (Semaglutide Dosage Considerations), Topic 2 (Transitioning from Rybelsus to Ozempic), Topic 3 (UInsurance Denial for Ozempic Treatment), Topic 4 (Pen Usage and Dosage Administration), Topic 5 (Weight Loss Progress), Topic 6 (Semaglutide Use), Topic 7 (Ozempic Storage and Travel), Topic 8 (Switching from Saxenda to Wegovy), Topic 9 (Managing constipation on Ozempic).

3. Results

3.1. Main topics of discussion around GLP-1 agonists

Fig. 1 visualizes the top 10 most representative topics and the top 5 words with the highest c-TF-IDF scores that emerged from training a BERTopic model on the dataset. The c-TF-IDF score quantifies the relevance of a word within a given topic relative to other words (for example, the representation of topic 3 "Insurance Denial for Ozempic Treatment" relies heavily on the presence of the word 'insurance' in a post; Yao et al., 2023).

Table 1 provides an overview of the top 10 thematic topics that emerged, along with respective topic labels generated by ChatGPT 3.5, top 10 representative words with the corresponding c-TF-IDF scores, and the number of posts per topic.

BERTopic classified 24,636 posts out of the total 46,491 posts in the dataset as outliers and did not cluster them into particular topics. With the present hyperparameters, 200 topics were generated in total. Only the top 10 most frequent topics that emerged will be summarized.

The most frequently discussed topic, labeled "Semaglutide and Weight Loss Experiences" (Topic 0), contained 1336 posts. This topic was characterized by discussions centered on personal experiences with semaglutide in relation to weight loss. Key terms such as 'semaglutide', 'weight', 'lost', 'appetite' frequently appeared, highlighting users' emphasis on their successes and challenges with the medication for weight management.

The second most prevalent topic, "Semaglutide Dosage Considerations" (Topic 1), included 1047 posts. This topic focused on the various dosages of semaglutide, with keywords like 'dose,' '25,' '1mg,' and 'increase' reflecting users' concerns about appropriate dosing and its effects on their treatment.

"Transitioning from Rybelsus to Ozempic" (Topic 2) was the third most discussed topic, with 991 posts. This topic covered user experiences and queries related to switching from Rybelsus to Ozempic, with representative keywords such as 'rybelsus,' '7mg,' '3mg,' and 'pill,' indicating a focus on the practical aspects and effectiveness of the switch.

The fourth most frequent topic, "Insurance Denial for Ozempic Treatment" (Topic 3), contained 941 posts. Discussions in this topic revolved around issues with insurance coverage and authorization for the off-label use of Ozempic, with key terms like 'insurance,' 'denied,' 'cover,' and 'authorization' highlighting users' frustrations with obtaining coverage. "Pen Usage and Dosage Administration" (Topic 4) was the fifth most common topic, encompassing 816 posts. This topic focused on the practical aspects of using the Ozempic pen, with keywords such as 'pen,' 'needle,' 'injection,' and 'dose' reflecting users' concerns and experiences with the administration of the medication.

The topic "Weight Loss Progress" (Topic 5), with 762 posts, dealt with tracking weight loss, featuring keywords like 'lbs,' 'pounds,' 'goal,' and 'progress.' Discussions in this topic revolved around users sharing their weight loss milestones and progress updates.

"Semaglutide Use" (Topic 6) was another notable topic, comprising 664 posts. Discussions here focused on obtaining semaglutide, including compounded forms and prescription issues, with keywords such as 'semaglutide,' 'compounded,' 'insurance,' and 'prescription' prevalent.

"Ozempic Storage and Travel" (Topic 7), with 592 posts, addressed concerns about the proper storage and management of Ozempic during travel. Key terms such as 'fridge,' 'refrigerated,' 'temperature,' and 'traveling' were commonly included in the related discussions.

The ninth most discussed topic, "Switching from Saxenda to Wegovy" (Topic 8), included 476 posts. This topic involved conversations about transitioning between these two medications, with representative keywords like 'saxenda,' 'wegovy,' 'switch,' and 'shortage' highlighting users' experiences and challenges with different GLP-1 agonists.

Finally, "Managing Constipation on Ozempic" (Topic 9) was also discussed frequently. It contained 476 posts. This topic focused on users'

Table 1
Top 10 thematic topics that emerged with respective topic labels, top 10 keywords, and number of posts for the dataset.

Topic number	Topic label generated by ChatGPT 3.5	Representative words (c-TF-IDF score)	Number of representative posts in topic
0	Semaglutide and Weight Loss Experiences	'semaglutide' (0.0150), 'weight' (0.0054), 'sema' (0.0051), 'week' (0.0047), 'im' (0.0043), 'weeks' (0.0041), 'started semaglutide' (0.0040), 'lost' (0.0038), 'appetite' (0.0036), 'loss' (0.0035)	1336
1	Semaglutide Dosage Considerations	'dose' (0.0132), '25′ (0.0109), '1mg' (0.0092), '05' (0.0086), 'increase' (0.0085), 'weeks' (0.0083), 'dosage' (0.0081), '025′ (0.0079), 'appetite' (0.0078), 'week' (0.0076)	1047
2	Transitioning from Rybelsus to Ozempic	'rybelsus' (0.0441), '7mg' (0.0144), '3mg' (0.0135), '14mg' (0.0102), 'pill' (0.0054), 'taking' (0.0051), 'taking rybelsus' (0.0051), 'rybelsus 3mg' (0.0049), 'mg' (0.0047), 'started rybelsus' (0.0046)	991
3	Insurance Denial for Ozempic Treatment	'insurance' (0.0254), 'denied' (0.0138), 'cover' (0.0118), 'authorization' (0.0111), 'covered' (0.0106), 'pa' (0.0101), 'prior authorization' (0.0096), 'prior' (0.0096), 'coverage' (0.0080), 'approved' (0.0079)	941
4	Pen Usage and Dosage Administration	'pen' (0.0300), 'needle' (0.0195), 'injection' (0.0156), 'dose' (0.0095), 'doses' (0.0094), 'liquid' (0.0092), 'left' (0.0090), 'injected' (0.0085), 'button' (0.0081), 'inject' (0.0078)	816
5	Weight Loss Progress	'lbs' (0.0168), 'pounds' (0.0135), 'goal' (0.0105), 'sw' (0.0096), 'progress' (0.0096), 'cw' (0.0093), 'months' (0.0078), 'gw' (0.0073), 'lost' (0.0072), 'weight' (0.0071)	762
6	Semaglutide Use	'semaglutide' (0.0285), 'compounded' (0.0107), 'insurance' (0.0105), 'compounding' (0.0094), 'online' (0.0089), 'compound' (0.0086), 'compounded semaglutide' (0.0075), 'get semaglutide' (0.0072), 'prescription' (0.0067), 'semiglutide' (0.0065)	680
7	Ozempic Storage and Travel	'fridge' (0.0291), 'pen' (0.0194), 'refrigerated' (0.0168), 'pens' (0.0140), 'temperature' (0.0139), 'room' (0.0133), 'traveling' (0.0133), 'use' (0.0124), '56 days' (0.0111), 'cold' (0.0110)	592
8	Switching from Saxenda to Wegovy	'saxenda' (0.0584), 'wegovy' (0.0114), 'switched' (0.0078),	483

(continued on next page)

Table 1 (*continued*)

Topic number	Topic label generated by ChatGPT 3.5	Representative words (c-TF-IDF score)	Number of representative posts in topic
		'switch' (0.0075), 'saxenda ozempic' (0.0071), 'switching' (0.0071), 'saxenda wegovy' (0.0066), 'shortage' (0.0062), 'ozempic' (0.0057), 'switching saxenda' (0.0056)	
9	Managing constipation on Ozempic	'constipation' (0.0342), 'fiber' (0.0148), 'constipated' (0.0131), 'poop' (0.0125), 'miralax' (0.0110), 'bowel' (0.0093), 'stool' (0.0091), 'laxatives' (0.0087), 'laxative' (0.0073), 'magnesium' (0.0071)	476

experiences with constipation as a side effect of Ozempic, featuring keywords such as 'constipation,' 'fiber,' 'laxatives,' and 'bowel.'

3.2. Dynamic topic modelling

We subsequently conduct DTM to investigate the temporal variation in the frequency of discussions about these topics.

Dynamic topic modeling was conducted to analyze and gain insight into the temporal variation in the frequency of the predominant topics of discussion that emerged in our dataset of Ozempic-related Reddit posts (Blei and Lafferty 2006; Grootendorst, M., 2022; Jung et al., 2023). Fig. 2 shows how the frequency of the top 5 topics changed over time from the years 2021 to 2023.

Between October 2022 to January 2023, the frequency of discussions related to "Semaglutide and Weight Loss Experiences" (Topic 0) increased significantly, making it the most discussed topic from January 2023 onward. Before this period, the discussion topics fluctuated between "Semaglutide Dosage Considerations" (Topic 1), "Insurance Denial for Ozempic Treatment" (Topic 3), and "Pen Usage and Dosage Administration" (Topic 4) being more prominent. This trend indicates a growing public interest in semaglutide for weight loss and reflects the shifting focus of online discussions about this medication.

4. Discussion

This study implemented a topic modeling approach (BERTopic) to extract and analyze topics from discussions in Ozempic-related subreddits between April 2019 and December 2023. The dataset (n = 46,491) contained posts from three subreddits: r/ozempic, r/ozempic-forweightloss, and r/semaglutide. ChatGPT 3.5 was used to label the extracted topics according to the resulting keywords and representative documents. DTM analysis was employed to visualize the evolution of the frequency in discussions about the extracted topics over time.

The results show that the predominant topic of discussion amongst the posts was the use of semaglutide/Ozempic for weight loss. Users would often discuss their successful or unsuccessful weight loss progress after Ozempic usage and what factors may impact the results (e.g., medical/health history). The frequent discussions about Ozempic for weight loss are in line with societal trends that emphasize body image and fitness, which are significantly shaped by pervasive cultural norms in Western societies (Faw et al., 2021). In the literature, there is increasing evidence that suggests women's perceptions of and satisfaction with their bodies are heavily impacted by their social networks and discussions regarding body image (Faw et al., 2021; Mills and Fuller-Tyszkiewicz, 2017). Studies, such as those by Han et al. (2024), have shown a dramatic surge in online searches for the off-label use of Ozempic for cosmetic weight loss in the United States over the past five years, reflecting the societal focus on weight management. Consistent with the overarching theme of weight loss and engaging in body-related co-rumination (Faw et al., 2021), topic 6 was represented by discussions about tracking weight loss progress after Ozempic intake. Individuals would post about their starting weights, current weights, and goal weights, as well as the time that it took to lose certain amounts of pounds or kilograms. As platforms such as Reddit provide individuals a space to share personal experiences and seek advice, it may also inadvertently amplify the perceived effectiveness of Ozempic as a weight loss solution. This is in line with Basch et al. (2023), who highlighted the role of social media in shaping public perceptions of Ozempic over evidence-based publications. The emphasis on off-label use of Ozempic for weight loss, despite it not being FDA-approved for such purposes, underscores a disconnect between regulatory guidance and public use. This discrepancy highlights a potential risk of misinformation and misuse, warranting attention from healthcare professionals and policymakers.

The second most frequent topic (topic 1) pertained to the dosage of Ozempic while topic 4 was related to concerns about dosage amounts via pen/injection administration. In particular, users would inquire as to whether certain dosages will induce unwanted side effects, which is in line with the tenth most discussed topic. The tenth most discussed topic

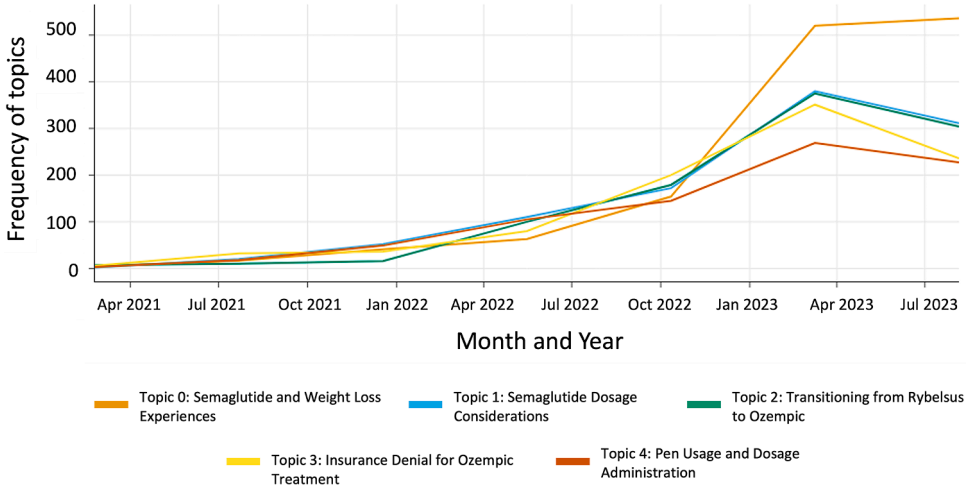

Fig. 2. Dynamic Topic Modeling of the frequency of the top 5 topics in the dataset over time.

(topic 9) was about the side effects associated with Ozempic usage. Users would discuss the occurrence of constipation and request for advice on how to ameliorate it. Aside from constipation and other gastrointestinal issues (e.g., nausea, vomiting, diarrhea), other known side effects of Ozempic and other GLP-1 agonists include pancreatitis, gallbladder disease, hypoglycemia, kidney failure, vision changes, tachycardia, depression, and suicidal ideation (Han et al., 2024; Suran, 2023). However, the general public does not seem to commonly discuss these risks as they did not emerge in the top 10 most discussed topics from the dataset. The findings resonate with Bartal et al. (2024), who explored the gap between user-reported experiences with GLP-1 Receptor Agonists on social media (X (formerly Twitter) and Reddit), and medically documented side effects, such as bezoar, atheroma, and sinus tachycardia. This emphasizes the need for enhanced communication strategies from healthcare providers to bridge this knowledge gap, ensuring users have access to comprehensive information about the risks associated with Ozempic use.

Switching between the various GLP-1 agonists emerged as a commonly discussed topic. Changing from Rybelsus to Ozempic was the third most discussed topic, while the ninth most discussed topic was switching from Saxenda to Wegovy. For instance, users would discuss their lack of weight loss with Rybelsus, inquire about the differences between the different brands, or discuss what would be the equivalent dose to taking Ozempic while taking Rybelsus. The brands differ in terms of dosage, administration, and indication (Suran, 2023). The overall theme of these topics reflects users' exploration of different medications within the same class to optimize weight loss outcomes, indicating an adoption of a trial-and-error approach. However, there is limited guidance on how to safely and effectively manage transitioning between various GLP-1 agonists, even in cases where Ozempic is appropriately used for diabetes management (Almandoz et al., 2020).

Insurance companies not covering the costs of the prescription or pharmacies denying the provision of Ozempic without a type 2 diabetes diagnosis were also commonly discussed. This is likely because Ozempic is not approved for weight loss in the UK, the USA, and the European Union (Han et al., 2024; Powell and Taylor, 2024). However, although both Ozempic and Wegovy contain the same active ingredient (semaglutide), only Wegovy is approved by relevant medicines authorities (e.g., FDA, European Medicines Agency) for weight loss and management in people with obesity (Powell and Taylor, 2024).

Overall, the topic modeling analysis demonstrates that while a varied range of topics emerged, the overarching theme of discussion in Ozempic-related subreddits was related to the off-shelf use of Ozempic and its counterparts for weight loss purposes. These results along with the DTM analysis reveal Ozempic-user's experiences and queries when using Ozempic, and support the growing evidence that the drug class is exponentially gaining popularity and public interest in weight management and weight loss (Han et al., 2024; Powell and Taylor, 2024; Basch et al., 2023; Suran, 2023). The results additionally support Han et al. (2024) and Basch et al. (2023), who note that social media sources (particularly from the non-professional general public) may have a greater impact and influence on public perception and usage of Ozempic over scientific evidence-based publications. Moreover, discussions from community members mainly focused on weight loss, administration methods, and overcoming insurance denial rather than awareness of severe health risks following Ozempic use (see Bartal et al. (2024) for examples of unreported potential adverse side effects).

The off-label use of Ozempic for body image enhancement has seen exponential growth recently due to its prevalence on social media, news outlets, and celebrity influence (Han et al., 2024). The findings of this study highlight the potential value of social media platforms collaborating with healthcare professionals and regulatory agencies to create interventions to prevent the misuse of Ozempic for weight loss. Such partnerships might include implementing prevention campaigns on social media to raise awareness among the general public about the dangers of off-label Ozempic use (Basch et al., 2023).

4.1. Limitations

The present study has some limitations to mention. First, it focuses solely on a single social media platform (Reddit) and could be enhanced by including discussions from other forums or platforms (e.g., Twitter, TikTok). Second, the analysis only considered submitted posts for topic modeling and did not include the accompanying comments. Insights from sub-discussions within the comments could provide additional understanding of Ozempic-related discussions on Reddit. Moreover, as Reddit is an anonymous platform, the demographic information of individuals who wrote the posts is not known, which may be useful to further inform our understanding of kratom usage. Lastly, as noted by Valdez & Patterson (2022), caution is needed when interpreting our findings, as the anonymity of Reddit users may also affect the authenticity of the posts.

4.2. Conclusion

Following the methodology of Fong et al. (2024), this study employed BERTopic, a natural language processing topic modeling approach, to efficiently extract the main topics of discussion from 46,491 Reddit posts from three subreddits (r/ozempic, r/ozempicforweightloss, r/semaglutide) that were posted between April 2019 and December 2023. The analysis revealed a prominent focus on the off-label use of Ozempic for weight loss, highlighting the significant influence of social media in promoting this practice without adequate information on potential health risks. The predominant discussion topics included weight loss, dosage concerns, insurance denial, tracking progress, and managing side effects. Notably, the study found that discussions often centered on switching between different GLP-1 agonists, with users prioritizing weight loss as their main goal rather than using these medications for their intended purpose of diabetes management.

Our findings reinforce existing evidence of the growing popularity and public interest in using Ozempic for weight management (Han et al., 2024; Powell and Taylor, 2024; Basch et al., 2023; Suran, 2023), and highlight a lack of discussion about the adverse consequences of Ozempic misuse. This trend underscores a critical need for enhanced public education regarding the risks associated with off-label Ozempic use. Given the role of social media in promoting the off-label use of Ozempic, these platforms could also be leveraged for prevention campaigns. Healthcare professionals and regulatory agencies might benefit from using social media platforms to inform consumers about the health risks and discourage unregulated prescriptions and usage (Basch et al., 2023). Addressing the gaps in public knowledge and ensuring accurate information dissemination could help mitigate the risks associated with unregulated use of Ozempic, ultimately promoting safer practices and informed decision-making among users.

CRediT authorship contribution statement

Seraphina Fong: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Alessandro Carollo:** Writing – review & editing, Writing – original draft, Formal analysis, Conceptualization. **Lambros Lazuras:** Writing – review & editing. **Ornella Corazza:** Writing – review & editing, Conceptualization. **Gianluca Esposito:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors (Ornella Corazza, Gianluca Esposito, Lambros Lazuras) are Editor-in-Chief, Editorial Board Member, and Guest Editor, respectively, for Emerging Trends in Drugs, Addictions, and Health and were not involved in the editorial review or the decision to publish this article.

Funding

The authors received no funding for this work.

Data availability statement

The raw data used for the present study was obtained from a public repository accessed here in light of Reddit's application programming interface (API) privacy policy changes in June 2023: https://www.reddit.com/r/pushshift/comments/1akrhg3/separate_dump_files_for_the_top_40k_subreddits/.

The repository is a large public dataset containing posts and comments from the top 40,000 subreddits obtained and stored publicly by an approved Reddit moderator.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.etdah.2024.100157.

References

Almandoz, J.P., Lingvay, I., Morales, J., Campos, C., 2020. Switching between glucagon-like peptide-1 receptor agonists: rationale and practical guidance. Clin. Diabetes 38 (4), 390–402.

Arillotta, D., Floresta, G., Guirguis, A., Corkery, J.M., Catalani, V., Martinotti, G., Sensi, S.L., Schifano, F., 2023. GLP-1 receptor agonists and related mental health issues; insights from a range of social media platforms using a mixed-methods approach. Brain Sci. 13 (11), 1503.

Arillotta, D., Floresta, G., Papanti Pelletier, G.D., Guirguis, A., Corkery, J.M., Martinotti, G., Schifano, F., 2024. Exploring the potential impact of GLP-1 receptor agonists on substance use, compulsive behavior, and libido: insights from social media using a mixed-methods approach. Brain Sci. 14 (6), 617.

Bartal, A., Jagodnik, K.M., Pliskin, N., & Seidmann, A. (2024). Utilizing AI and social media analytics to discover adverse side effects of GLP-1 receptor agonists. arXiv preprint arXiv:2404.01358.

Basch, C.H., Narayanan, S., Tang, H., Fera, J., Basch, C.E., 2023. Descriptive analysis of TikTok videos posted under the hashtag #Ozempic. J. Med. Surg. Public Health 1, 100013.

Baumgartner, J., Zannettou, S., Keegan, B., Squire, M., Blackburn, J., 2020. The pushshift reddit dataset. Proc. Int. AAAI Conf. Web Social Media 14, 830–839.

Blei, D.M., Lafferty, J.D., 2006. Dynamic topic models. In: Proceedings of the 23rd international conference on Machine learning, pp. 113–120.

Carollo, A., Corazza, O., Mantovani, M., Silvestrini, N., Rabin, O., Esposito, G., 2024. Performance-enhancing substances in sport: a scientometric review of 75 years of research. Drug Test. Anal. 1–12.

Choi, J., Jang, H., 2023. Topic and sentiment analysis of Reddit's vaping-cessation community. In: 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE, pp. 4878–4880.

Devlin, J., Chang, M.W., Lee, K., & Toutanova, K. (2018). Bert: pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv: 1810.04805*.

Egger, R., Yu, J., 2022. A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. Front. Sociol. 7, 886498.

Faw, M.H., Davidson, K., Hogan, L., Thomas, K., 2021. Corumination, diet culture, intuitive eating, and body dissatisfaction among young adult women. Pers. Relatsh. 28 (2), 406–426.

Fong, S., Carollo, A., Prevete, E., Corazza, O., Esposito, G., 2024. Dynamic Topic Modeling of Kratom Use and Experiences: Insights on 13 years of Reddit Discussions. https://doi.org/10.31219/osf.io/5hyjx.

Grootendorst, M. (2022). BERTopic: neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.

Grootendorst, M.P. (n.d.-c). FAQ - BERTopic. https://maartengr.github.io/BERTopic/faq.html#how-do-i-remove-stop-words.

Grootendorst, M.P. (n.d.-b). Best Practices - BERTopic. https://maartengr.github.io/BERTopic/getting_started/best_practices/best_practices.html.

Grootendorst, M.P. (n.d.-a). 6B LLM & Generative AI - BERTopic. https://maartengr.github.io/BERTopic/getting_started/representation/llm.html#chatgpt.

Han, S.H., Safeek, R., Ockerman, K., Trieu, N., Mars, P., Klenke, A., Furnas, H., Sorice-Virk, S., 2024. Public interest in the off-label use of glucagon-like peptide 1 agonists (Ozempic) for cosmetic weight loss: a Google trends analysis. Aesthet. Surg. J. 44 (1), 60–67.

Jung, H.S., Lee, H., Kim, J.H., 2023. Unveiling gryptocurrency conversations: Insights from data mining and unsupervised learning across multiple platforms. IEEE Access 11, 130573–130583.

Khawagi, W.Y., Alotabi, J., Shamlan, W.B., Abdullah, S., Bawazir, M., Alqaysi, L., Alsharif, R., Bajood, M., Aldoboke, M., Bayounis, M., Alkhammash, T., Alzahrani, K., 2023. KSA patients with obesity: their knowledge and awareness level of the side effects of ozempic. Arch. Pharm. Pract. 14 (S), A06231510, 2023.

Lopes, A.C., Roque, M.D.F.D.S.M., Lourenço, O., Herdeiro, M.T., Morgado, M.A.N.V.P., 2023. Gastrointestinal disorders potentially associated with Semaglutide: an analysis from the Eudravigilance Database. Expert. Opin. Drug Saf. 1–7.

Melton, C.A., Olusanya, O.A., Ammar, N., Shaban-Nejad, A., 2021. Public sentiment analysis and topic modeling regarding COVID-19 vaccines on the Reddit social media platform: a call to action for strengthening vaccine confidence. J. Infect. Public Health 14 (10), 1505–1512.

Mills, J., Fuller-Tyszkiewicz, M., 2017. Fat talk and body image disturbance: a systematic review and meta-analysis. Psychol. Women. Q. 41 (1), 114–129.

Myerson, M., Paparodis, R.D., 2024. Pharmacotherapy of weight-loss and obesity with a focus on GLP 1-receptor agonists. J. Clin. Pharmacol.

Ng, Q.X., Lee, D.Y.X., Yau, C.E., Lim, Y.L., Ng, C.X., Liew, T.M., 2023. Examining the public messaging on 'loneliness' over social media: an unsupervised machine learning analysis of twitter posts over the past decade. In: Healthcare, 11. MDPI, p. 1485.

Pleasants, E., Roy, A., Weidert, K., Marshall, C., Upadhyay, U.D., Cheshire, C., Prata, N., 2023. Understanding r/abortion through natural language processing: using a dynamic topic modeling approach to explore use of an online community for abortion during 2022. Contraception 127, 110187.

Pont-Fernandez, S., Kheyfets, M., Rogers, J.M., Smith, K.E., Epstein, D.H., 2023. Kava (Piper methysticum) in the United States: the quiet rise of a substance with often subtle effects. Am. J. Drug Alcohol Abuse 49 (1), 85–96.

Powell, J., Taylor, J., 2024. Use of dulaglutide, semaglutide, and tirzepatide in diabetes and weight management. Clin. Ther.

Quddos, F., Hubshman, Z., Tegge, A., Sane, D., Marti, E., Kablinger, A.S., Gatchalian, K. M., Kelly, A.L., DiFeliceantonio, A.G., Bickel, W.K., 2023. Semaglutide and Tirzepatide reduce alcohol consumption in individuals with obesity. Sci. Rep. 13 (1), 20998. https://doi.org/10.1038/s41598-023-48267-2.

Anon, *Reddit Homepage*, 2024. Reddit. February 1. https://www.redditinc.com/.

Simonato, P., Scoppettone, A., Bersani, F.S., Rossato, S., 2022. The performance enhanced: clinical relationships between performance-and image-enhancing drugs, physical exercise and mental disorders. Emerg. Drugs Sport 151–159.

Smith, K.E., Rogers, J.M., Schriefer, D., Grundmann, O., 2021. Therapeutic benefit with caveats?: analyzing social media data to understand the complexities of kratom use. Drug Alcohol Depend. 226, 108879.

Suran, M., 2023. As Ozempic's popularity soars, Here's what to know about semaglutide and weight loss. JAMa 329 (19), 1627–1629.

Valdez, D., Patterson, M.S., 2022. Computational analyzes identify addiction help-seeking behaviors on the social networking website Reddit: insights into online social interactions and addiction support communities. PLOS. Digit. Health 1 (11), e0000143.

Wanchoo, K., Abrams, M., Merchant, R.M., Ungar, L., Guntuku, S.C., 2023. Reddit language indicates changes associated with diet, physical activity, substance use, and smoking during COVID-19. PLoS. One 18 (2), e0280337.

Anon. Watchful1. (n.d.). Subreddit comments/submissions 2005-06 to 2023-12. https://www.reddit.com/r/pushshift/comments/1akrhg3/separate_dump_files_for_the_top_40k_subreddits/.

Xu, Z., Fang, Q., Huang, Y., Xie, M., 2024. The public attitude towards ChatGPT on reddit: a study based on unsupervised learning from sentiment analysis and topic modeling. PLoS. One 19 (5), e0302502.

Yao, L.F., Ferawati, K., Liew, K., Wakamiya, S., Aramaki, E., 2023. Disruptions in the cystic fibrosis community's experiences and concerns during the COVID-19 pandemic: topic modeling and time series analysis of Reddit comments. J. Med. Internet Res. 25, e45249.