Utilizing AI and Social Media Analytics to Discover Unreported Adverse Side Effects of GLP-1 Receptor Agonists Used for Obesity Treatment

Alon Bartal
Bar-Ilan University
alon.bartal@biu.ac.il

Kathleen M. Jagodnik
Bar-Ilan University
kathleen.jagodnik@biu.ac.il

Nava Pliskin
Ben-Gurion University
pliskinn@bgu.ac.il

Abraham Seidmann
Boston University
avis@bu.edu

Abstract

Adverse side effects (ASEs) of drugs, revealed after FDA approval, pose a threat to patient safety. To promptly detect overlooked ASEs, we developed a digital health methodology capable of analyzing massive public data from social media, published clinical research, manufacturers' reports, and ChatGPT. We uncovered ASEs associated with the glucagon-like peptide 1 receptor agonist (GLP-1 RA) medications used to treat diabetes and obesity, a market expected to grow exponentially to $133.5 billion USD by 2030. Using a named entity recognition model, our method successfully detected 15 potential ASEs of GLP-1 RAs, overlooked upon FDA approval. Our data-analytic approach revolutionizes the detection of unreported ASEs associated with newly deployed medications, leveraging cutting-edge AI-driven social media analytics. This ongoing research can increase the safety of new medications in the marketplace by unlocking the power of social media to support regulators and manufacturers in the rapid discovery of hidden ASE risks.

Keywords: Adverse Side Effect (ASE), Artificial Intelligence (AI), Glucagon-Like Peptide 1 Receptor Agonist (GLP-1 RA), Social Media Analytics

1. Introduction

Over recent decades, obesity has surged as a global health crisis (World Health Organization, 2021), with numbers nearly tripling since 1975, surpassing 650 million individuals globally, among 1.9 billion adults who are overweight (World Health Organization, 2021). Obesity significantly elevates the risk of developing an array of health disorders, including type 2 diabetes (T2D), high blood pressure, heart disease, respiratory problems, joint problems, and gallbladder disease (Pearson-Stuttard et al., 2022). The consequences of obesity are projected to consume more than 3% of the gross domestic product in the United States (U.S.) in the coming decades (Okunogbe et al., 2022).

Glucagon-like peptide 1 receptor agonists (GLP-1 RAs), used to treat obesity and T2D, effectively reduce weight, improve comorbidities, and lower blood sugar levels (Moore et al., 2023). Since the U.S. Food and Drug Administration (FDA) approved the first drug in this class, exenatide, in 2005, the use of GLP-1 RA such as dulaglutide (Trulicity), liraglutide (Victoza), and semaglutide (Ozempic, Rybelsus) has surged (Trujillo et al., 2021). In 2023, prescriptions spiked, driven by both T2D and obesity treatments (Pritchett, 2023).

Semaglutide's popularity, fueled by celebrity uses for weight loss, led to supply shortages (Han et al., 2023). However, GLP-1 RA also have documented Adverse Side Effects (ASEs), including gastrointestinal disorders, metabolic problems, and eye, renal, urinary, and nutritional disorders (Zhang et al., 2023). Since some drugs in this class were approved in recent years, their complete list of ASEs profiles remain under-characterized. Moreover, clinical trials, typically involving only a few thousand human subjects, often cannot detect rare ASEs or those with significant latent development (Berlin et al., 2008). For example, non-obese individuals who use GLP-1 RA for weight loss are not commonly represented in clinical trials. Limited knowledge about ASEs lead physicians to prescribe medications with incomplete information, risking patient safety.

Furthermore, some FDA-approved drugs have been withdrawn from the market when their ASEs became known only after widespread use (Ridings, 2013).

URI: https://hdl.handle.net/10125/109558
978-0-9981331-8-8

Thalidomide, for example, was prescribed to women during pregnancy to treat nausea, but led to severe birth defects in infants whose mothers had taken the drug, ultimately resulting in its withdrawal from the market (Ridings, 2013). Vioxx, which was approved by the FDA for the treatment of pain and inflammation, was withdrawn due to its being associated with increased risk of heart attack and stroke (Ortiz, 2004).

These historical pharmacological precedents motivate the need to add another layer of safety beyond the initial FDA approval. Social media can supplement these resources by allowing researchers to preemptively detect ASEs before regulatory intervention (J.-Y. Lee et al., 2021). We propose a digital health methodology that mines data from social media, extending beyond existing pharmacovigilance insights provided by drug manufacturers and clinical trials, to comprehensively identify unreported ASEs of newly developed drugs.

This study contributes to the understanding of ASEs related to GLP-1 RA through a multifaceted knowledge-discovery methodology. Our methodology can serve as an effective 'early warning' sensor for identifying unreported ASEs. We construct an ASE-ASE network, based on co-mentioned ASEs in social media posts, representing a novel approach to revealing groups of ASEs with similar network patterns. We further contribute by applying network community detection and classification based on mention frequency, offering a nuanced perspective on ASE prevalence and interconnections. Overall, our study contributes methodologically and conceptually by offering valuable insights for drug safety assessment, public health management, and the broader understanding of GLP-1 RA related ASEs.

Organization: Section 2 covers related literature. Section 3 elaborates on the research question and hypotheses. Section 4 presents the characteristics of the analyzed data. Section 5 provides a detailed description of the methodology employed. Sections 6 presents the results of applying our methodology to the data. Finally, the discussion and conclusions in Section 7 interpret the findings in the context of the existing literature, explore their implications, and summarize key insights.

2. Literature Review

Drug ASEs are recorded in databases such as the FDA Adverse Event Reporting System (FAERS) (U.S. Food and Drug Administration, 2023), the Side Effect Resource (SIDER) (Kuhn et al., 2016), MEDLINE (Ding et al., 2001), and Embase (Elsevier, 2023). However, these databases suffer from improper indexing (Derry et al., 2001) and incomplete reporting (Nugent et al., 2016), with up to 86% of ASEs unreported (Nugent et al., 2016). Moreover, publishing bias favoring positive results, as well as practices that may conceal negative data in publications, can lead to incomplete recording of ASEs. Additionally, the diversity of resources providing data about ASEs, the heterogeneous and incomplete nature of these data and patients not reporting ASEs (Nugent et al., 2016) make it challenging to characterize ASEs of a drug (Nugent et al., 2016).

Social media provides an additional resource for researchers to identify ASEs before regulatory intervention (J.-Y. Lee et al., 2021). For example, social media analyses have been conducted to identify ASEs of medications for HIV (Godinez et al., 2023), chloroquine for COVID-19 (de França et al., 2021), and buprenorphine-naloxone treatment for opioid use (Graves et al., 2022). Specifically, Reddit (Godinez et al., 2023; Graves et al., 2022) and X (formerly Twitter) (de França et al., 2021) have been used to discover pharmacovigilance knowledge. Reddit enables anonymous accounts, encouraging users to share ASE experiences more honestly compared with other platforms, as anonymity often leads individuals to disclose information more freely. Reddit permits more substantial exchanges, compared with X, which limits users to brief posts. However, the briefer X posts can also provide significant information about ASEs reported by users and about other aspects of drug usage (Back et al., 2023).

Using social media analysis, it was found that ASEs could be detected 3 to 35 months before an FDA alert was issued, depending on the drug (Adjeroh et al., 2014). A scoping review of social media use for pharmacovigilance (J.-Y. Lee et al., 2021) reported that social media monitoring can facilitate the detection of labeling changes, black box warnings, or drug withdrawals 3 months to 9 years in advance. Additionally, recent meta-analysis (Gao et al., 2022) has shown that a higher prevalence of ASE discussions on social media is associated with a shorter time to drug recall (Gao et al., 2022).

The rising popularity of GLP-1 RAs has led to a surge in online searching (Han et al., 2023) and social media posting about them, offering valuable data to supplement existing pharmacovigilance sources (Back et al., 2023). While social media posts do provide abundant information about GLP-1 RA, previous research has mainly focused on small-scale data (Basch et al., 2023; Lennon, 2023). For instance, recent analysis of the top 100 results of TikTok videos tagged #Ozempic revealed that most videos emphasized personal experiences and promoted the drug for weight loss, potentially contributing to increased demand (Basch et al., 2023). That study stressed the importance of collaboration among healthcare professionals, regulatory bodies, and social media platforms to address challenges like drug shortages. Another small-scale study on TikTok compared 16 videos in total, highlighting how Ozempic users framed obesity as a disease, unlike discussants of traditional diet methods (Lennon, 2023). Researchers have also started to use the abundant information that social media offers about GLP-1 RA for broader analysis. For example, the study (Arillotta et al., 2023) analyzed tens of thousands of posts on Reddit (12,136), YouTube (14,515), and TikTok (17,059) and uncovered associations between GLP-1 RA and various mental health issues. We further study to capture subsequent comments after prolonged drugs use.

Training Natural Language Processing (NLP) models, including those utilizing Named Entity Recognition (NER), is integral to extracting knowledge from unstructured text sources such as social media posts, electronic medical records, and medical literature (Tarcar et al., 2019). NER plays a crucial role in identifying biomedical entities such as drugs, symptoms, and diseases mentioned in the text, with potential implications for linking them to ASEs (Tarcar et al., 2019). Advanced NER models, leveraging deep-learning techniques, and utilizing contextual information surrounding entities within sentences, benefit from Large Language Models (LLMs) that are typically pre-trained on massive publicly and freely available text on the Internet, encompassing books, articles, and webpages (Temperton, 2023). LLMs can receive and generate text queries while adapting to a variety of language-related tasks beyond their training objective (Omiye et al., 2024).

In the medical field, LLMs have been applied to tasks like responding to medical exams and addressing patient queries. Despite the risk of generating inaccurate outputs, healthcare providers are increasingly integrating LLMs into clinical applications (P. Lee, Bubeck, & Petro, 2023). Noteworthy examples include LLM use in medical resident training modules and in partnerships of healthcare providers that incorporate electronic health records (Joseph, 2023). Within the landscape of general LLMs, the Generative Pre-trained Transformers (GPT) lineage, particularly exemplified by ChatGPT, a frequently used choice, especially in chat applications (Omiye et al., 2024).

ChatGPT is an LLM known for generating informative text and mimicking human-like writing (Lancaster, 2023). It has been widely used in various scientific domains, including mental health research (Bartal et al., 2023), medical research (P. Lee, Goldberg, & Kohane, 2023), and detecting drug interactions (Juhi et al., 2023). However, ChatGPT's capabilities have not yet been used to study GLP-1 RA, presenting a valuable opportunity for knowledge discovery through the methods harnessed in the present study.

Computational-based algorithms have already facilitated understanding of drug ASEs, including identifying when they occur (McMaster et al., 2023). Drug-related ASEs persist as a significant cause of morbidity and mortality in healthcare, resulting in substantial economic costs, amounting to an estimated annual cost of $495-672 billion in the U.S. alone (Watanabe et al., 2018). Moreover, accurate estimation can help pharmaceutical companies to mitigate the risk of a drug's withdrawal from the market (Onakpoya et al., 2016) and avoid reassessing ASEs through costly clinical trials (Martin et al., 2017).

3. Research Questions

RQ1. How can we identify novel ASEs of GLP-1 RAs based on social media discussions?

Many patients seek advice and exchange experiences on social media, especially concerning ASEs (Ozurumba-Dwight et al., 2022). Thus, we hypothesize:

H1. A Named Entity Recognition (NER) model applied to textual posts can uncover potential ASEs of GLP-1 RAs unobserved by pharmacovigilance methods.

RQ2. What are the mechanisms and conditions under which social media can enhance pharmacovigilance for identifying novel ASEs?

Since the frequency of ASE mentions on social media might imply the prevalence and potential severity of a particular ASE, we hypothesize:

H2. The more frequent an ASE is mentioned in social media, the more informative it is for pharmacovigilance.

4. Data

We collected five datasets covering biomedical knowledge and social media posts involving 12 GLP-1 RAs on our *Medication List* that includes both brand names and generic drug names: (1) dulaglutide [brand name: (2) Trulicity], (3) exenatide [brand names: (4) Byetta; (5) Bydureon - extended-release], (6) liraglutide [brand name: (7) Victoza], (8) lixisenatide [brand name: (9) Adlyxin], and (10) semaglutide [brand names: (11) Ozempic - injections; (12) Rybelsus - tablets].

Dataset 1: X – Using the APIfy[1] website, we

[1]https://console.apify.com

collected the 1,000 most recent posts from 2017 to 2023 that mentioned each of the 12 items on our *Medication List*. This produced a dataset that includes 11,185 posted texts, each with associated data on the posting user, posting time, and specific GLP-1 RA mentioned.

Dataset 2: Reddit – Using the Python Reddit API Wrapper (PRAW) package (Boe, 2023), we collected 269,324 posts and comments from 10 subreddits (Table 1) related to the 12 items on our *Medication List*. The posts originated in 2022 and 2023. Each post includes: 'Post ID', 'Post Author', 'Post Content', 'Post Date', 'Comment ID', 'Comment Author', 'Comment Content', 'Parent ID', and 'Parent Author'.

Table 1. Subreddits that were used to collect social media posts involving GLP-1 RAs.

Subreddit	# Members*
r/diabetes	109K
r/GLP1	1.3K
r/liraglutide	11.7K
r/MaintenancePhase	25.6K
r/Ozempic	50.1K
r/OzempicForWeightLoss	13.7K
r/semaglutidecompounds	7.2K
r/Semaglutide	45K
r/TheMorningToastSnark	11.6K
r/trulicity	1.1K

*As of October 2, 2023.

Dataset 3: PubMed – Using the National Center for Biotechnology Information (NCBI) PubMed E-utilities API (PubMed API) (Sayers, 2023) and the Biopython Python library (Chapman & Chang, 2000), we collected 13,491 articles indexed from 2017 to 2024 in PubMed that mentioned at least one item on our *Medication List* in the title or in the abstract.

Dataset 4: SIDER – We utilized the `meddra_all_se.tsv.gz` dataset from the SIDER database version 4.1, which catalogs 5,868 documented ASEs for 1,430 commercially available drugs. SIDER compiles information from publicly accessible drug documents and package inserts, providing details on drug names, ASE frequencies, and classifications.

Dataset 5: Side Effects Reported by Manufacturers and by ChatGPT – For manufacturers, we collected from the Internet ASEs reported in clinical trials and post-marketing surveillance. To address potential gaps in the ASE data reported by manufacturers, we integrated them into Dataset 5 with commonly reported ASEs collected from ChatGPT (GPT-3.5) by using the interactive web interface. For this collection, we utilized the prompt: 'Please provide a list of adverse side effects for the glucagon-like peptide 1 receptor agonists named S,' where S corresponds to an item on our *Medication List*. ChatGPT's responses incorporate insights from social media or medical literature, that are not explicitly included in our manual search of manufacturers' reports. This ensures that Dataset 5 aligns more closely with the timeframes of Datasets 1 and 2, with information up to 2023, and Dataset 3, which covers information up to January 2024.

5. Methods

5.1. Early Identification of ASEs

We identified ASEs of FDA-approved GLP-1 RAs by analyzing the texts in Datasets 1 to 3 using the ScispaCy pre-trained NER model `en_ner_bc5cdr_md` to identify biological entities. To ascertain whether these identified biological entities are potential ASEs of GLP-1 RAs, we matched them against the ASEs catalogued in SIDER (Kuhn et al., 2016) (Dataset 4), after normalizing and converting to lowercase both the identified ASEs and those in SIDER, as well as employing NLP stemming to reduce words to their root forms (Jivani et al., 2011). Each identified biomedical entity was considered a potential ASE if it existed in SIDER (Kuhn et al., 2016). To the resulting compilation of ASEs, we added ASEs collected in Dataset 5. Then, we manually grouped similar ASEs, for example, combining 'headache' and 'migraines', and standardized ASEs into their noun forms. Consolidating and analyzing these diverse sources based on dynamic and up-to-date social media data facilitated early comprehensive identification of potential ASEs that may have been overlooked during regulatory approval processes.

5.2. Analysis of ASE Mention Frequency

We analyze the frequency of ASE mentions on social media for each GLP-1 RA on our *Medication List*. Then, we examined all ASEs identified in the social media data and classified them into one of two group: (1) *Established ASEs:* ASEs reported by manufacturers and ChatGPT, and (2) *Potential ASEs:* ASEs identified by our modeling approach but not reported by manufacturers or ChatGPT.

To test *H1* and *H2*, we assessed our method's effectiveness by computing the Overlap$_{f\%}$ score across various ASE mention frequencies (Equation 1).

$$Overlap_{f\%} = |I_{f\%} \cap E| / |I_{f\%}| \tag{1}$$

$I_{f\%}$ - The set of identified established and potential

ASEs in social media data located in the top $f_{\%}$ of a ranking, based on the mentioned frequency of ASEs.

E - The set of established ASEs.

A higher Overlap$_{f_{\%}}$ score $\in [0,1]$, indicates greater success identifying established ASEs among identified ASEs. An Overlap = 1 means all identified ASEs are established ASEs, while Overlap = 0 indicates none are.

The Overlap score signifies our ability to capture ASE instances among identified ASEs (addressing *H1*); and if our approach accurately identifies established ASEs (higher Overlap score), we can confidently explore potential ASEs that our approach discovers as promising for pharmacovigilance. (addressing *H2*)

5.3. Analysis of ASE Frequency Over Time

To measure ASE mentions frequency along intervals of 14 days, we defined two frequency measures, considering only ASEs with > 10 mentions per interval: (1) Pre-Mention Average Frequency (Pre-MAF): average mentions before a selected point in time, noted by t_0; and (2) Post-Mention Average Frequency (Post-MAF): average mentions after a selected point in time (t_0).

We define a positive slope in mention frequency as Pre-MAF $<$ Post-MAF, a negative slope as Pre-MAF $>$ Post-MAF, and no slope otherwise.

5.4. Network Analysis of ASEs

Using the identified ASEs in the social media posts of Datasets 1 and 2, we built an ASE-ASE network denoted by $G = (V, E, W)$. In G, nodes (V) are identified ASEs. Two nodes mentioned in the same post are connected by an edge (E). The weights of the edges (W) represent the frequency of two ASEs being mentioned together in posts.

Identifying the co-occurrence patterns of ASEs, which is crucial in pharmacovigilance drug safety analysis (Galeano et al., 2020), was accomplished by clustering ASEs using network community detection. To identify sets of interconnected ASEs, we applied the `cluster_louvain` node-clustering algorithm from the `igraph` R library to the ASE-ASE network G. This algorithm employs a multi-level modularity optimization function, leveraging the modularity measure and adopting a hierarchical methodology. Clustering nodes into community groups enabled the identification of ASEs that frequently co-occur beyond a direct 1-hop edge in G.

Figure. 1 summarizes the steps of our methodology, used to address the two research questions (Section 3).

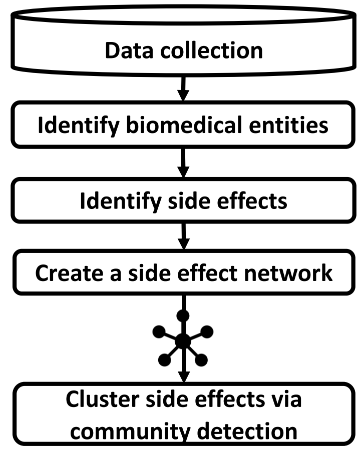

Figure 1. Overview of our methodology.

6. Results

6.1. Identification of ASEs of GLP-1 RAs

We found 120 ASEs of GLP-1 RAs and analyzed their overlaps and differences. Additionally, we looked at their mention frequencies on social media (Datasets 1, 2), in academic papers (Dataset 3), and reported percentages by drug manufacturers (Dataset 5), revealing seven subgroups.

Group 1 lists 32 ASEs found at the intersection of academic papers, social media, and manufacturers' reports or reported by ChatGPT: Cholelithiasis, Chills, Cramps, Constipation, Depression, Diarrhea, Dizziness, Dyspepsia, Fatigue, Gastritis, Headaches, Hidradenitis, Hyperglycemia, Hypersensitivity, Hypoglycemia, Hypertrophy, Increased Heart Rate, Inflammation, Insomnia, Jaundice, Microalbuminuria, Nausea, Otitis, Polyuria, Rash, Stress, Swallowing Difficulty, Thirst, Thrombosis, Thromboembolism, Thyroid Issues, and Vomiting.

Group 2 lists 39 ASEs reported in academic papers and mentioned on social media: Amnesia, Anorexia, Arthritis, Arrhythmias, Atherosclerosis, Bradycardia, Cardiomyopathy, Cholangitis, Compulsive Behavior, Dementia, Dehydration, Delusion, Encephalopathy, Erythema, Fasciitis, Fractures, Gangrene, Glycosuria, Hepatic Issues, Hypercholesterolemia, Hypotonia, Hypothyroidism, Infarction, Ketosis, Ketoacidosis, Lethargy, Malnutrition, Mania, Neuropathy, Neurotoxicity, Osteoarthritis, Osteomyelitis, Palpitations, Pneumonia, Retching, Tachycardia, Trauma, Ulcerative Wounds, and Urticaria,

Group 3 lists 14 ASEs reported exclusively in academic papers: Atheroma, Bezoar, Cardiotoxicity, Demyelination, Dyskinesia, Hyperphagia, Hyperplasia, Hypoxia, Ischemia, Nephrotoxicity, Paresthesia, Polyneuropathies, Pruritus (Itching), and Stiffness.

Group 4 lists 13 ASEs reported by ChatGPT and in manufacturers' reports: Colorectal Neoplasms, Dysgeusia, Elevated Serum Amylase, Elevated Serum Lipase, First-Degree Atrioventricular (AV) Block, Hair Loss, High Calcitonin, Increase in Amylase, Injection Site Nodule, Papillary Thyroid Cancer, Serious Allergic Reaction, Sinus Tachycardia, Thyroid C-Cell Tumors,

Group 5 lists 15 ASEs exclusively mentioned on social media, which were not reported by any other source: Anhedonia, Apathy, Aura, Bulimia, Endometriosis, Frustration, Hirsutism, Hoarseness, Infertility, Ketonemia, Narcolepsy, Onycholysis, Paralysis, Snoring, and Suicidal

Group 6 lists six ASEs mentioned on social media, by ChatGPT, and in manufacturers' reports: Appendicitis, Back Pain, Bloating, Eructation, Hepatotoxicity, and Syncope.

Group 7 lists similar ASEs of Blurred Vision, Non-Proliferative Retinopathy, and Diabetic Retinopathy that were mentioned in academic papers, by ChatGPT, and in manufacturers' reports. These ASEs were grouped into a single ASE as they describe similar ASEs.

The 15 potential ASEs in Group 5 are of great interest, as they might have been overlooked during the FDA regulatory approval process, thereby supporting *H1*. The early identification of potential ASEs through social media analytics warrants further investigation.

6.2. Analysis of ASE Mention Frequency

Established ASEs are listed in Groups 1, 4, 6, and 7; and Potential ASEs are listed in Groups 2 and 5. Figure 2 depicts the success of our model's identification of established ASEs as a function of $I_{f\%}$, where $f_{\%}$ ranges between 10% and 100% with increments of 10%.

We find that as $f\%$ increases, Overlap$_{f\%}$ decreases. This indicates that ASEs mentioned more frequently on social media ($I_{f\%}$ where $f\%$ is smaller) are more likely to be Established ASEs, supporting *H2*.

Additionally, we fitted a logarithmic curve to model the observed Overlap$_{f\%}$ pattern: $Overlap_{f\%} = -0.29\ln(x) + 0.9$, yielding an impressive R^2 value of 0.99. This curve enables accurate estimation of Overlap$_{f\%}$ scores based on the value of $f\%$, allowing us to more confidently (depending on the Overlap score) recommend potential ASEs as novel.

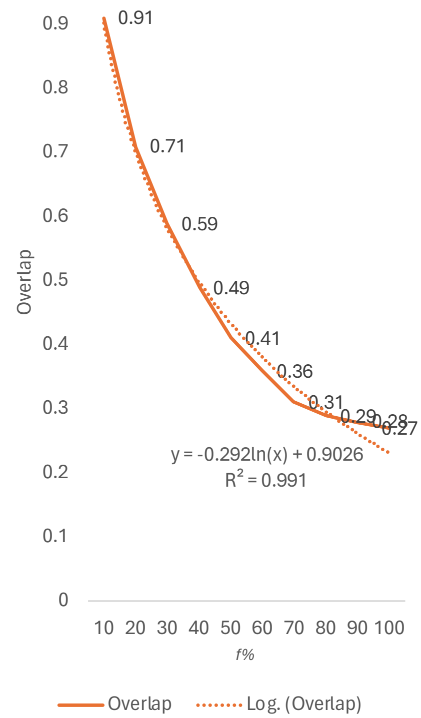

Figure 2. Overlap score as a function of $I_{f\%}$.

6.3. Analysis of ASE Frequency Over Time

Analyzing ASE mention frequency in social media with 14 days intervals reveals that Semaglutide (Ozempic), the most prescribed GLP-1 RA, has the highest number of associated ASEs mentioned in the data (Table 2).

Additionally, we found that the ASEs mentioned most frequently over time are anxiety, constipation, nausea, and vomiting. Analyzing trends in the data, we noticed a great increase of ASEs mentions starting on September 24, 2023. This date was selected as t_0 to measure ASE frequency over time. We observe an overall positive slope (Pre-MAF $<$ Post-MAF) for anxiety, constipation, nausea, and vomiting, an overall negative slope for fatigue, and no slope for depression (Table 3).

A Google search reveals that Oprah Winfrey, addressed, in her show on September 20, 2023, the trend of using weight-loss drugs. She shared her initial reluctance to use such drugs, emphasizing the societal pressure to rely on personal effort rather than medication.

Table 2. ASE mention frequency (> 3) on X and Reddit for each GLP-1 receptor agonist.

	S*	L*	D*
Anger	11	-	-
Anorexia	17	-	-
Anxiety	208	-	24
Arthritis	15	-	22
Burns	24	-	4
Chills	13	-	-
Constipation	275	-	35
Cough	13	-	-
Cramps	47	-	11
Dehydration	37	-	-
Dementia	8	-	-
Depression	168	19	27
Diarrhea	23	4	4
Dizziness	29	6	4
Fatigue	116	6	32
Frustration	6	-	-
Gastritis	59	4	13
Headaches	164	5	21
Hypotonia	20	-	-
Inflammation	141	7	11
Insomnia	11	-	3
Irritability	10	4	-
Ketosis	17	-	-
Lethargy	8	-	-
Nausea	540	52	115
Neuropathy	23	-	4
Numbness	13	4	-
Rash	21	4	12
Malnutrition	9	-	-
Stress	39	5	4
Suicidal	9	-	-
Thirst	6	-	-
Trauma	12	-	-
Ulcerative	370	32	63
Vomiting	192	11	47

*S: Semaglutide, L: Liraglutide, D: Dulaglutide

Table 3. Frequency of ASE mentions before and after the observed spike in mentions.

Adverse Side Effect	Pre-MAF	Post-MAF	Slope
Anxiety	18.0	20.0	+
Constipation	29.0	35.0	+
Depression	15.5	15.5	0
Fatigue	17.6	14.6	-
Nausea	39.1	57.8	+
Vomiting	20.5	27.0	+

6.4. Network Analysis of ASEs

Based on the unprocessed social media data (Datasets 1 and 2), we constructed an ASE-ASE weighted and non-directed network $G = (V, E, W)$. To emphasize that our method can identify ASEs in raw social media discussions (RQ1), unlike in Section 6.1, here we neither combined similar ASEs nor standardized ASEs into noun forms.

We identified 381 potential ASEs with 1,440 ASE-ASE interactions co-mentioned in the same post. To reduce noise and focus on ASEs only, we manually removed 52 falsely identified ASEs, ASE-ASE interactions with fewer than 3 co-mentions, and ASE nodes with no co-mentions in the data (having degree centrality 0). We also used the `components` function within the `igraph` R library (Csardi & Nepusz, 2006) to detect graph components. The algorithm detected three components of sizes 2, 2, and the main component of 78 nodes. We set G as the main component graph with 78 ASE nodes and 253 weighted edges. Nodes are ASEs and edges are ASE-ASE co-mentioned in X and Reddit posts.

To identify sets of interconnected ASEs, we applied clustering analysis to G using the `cluster_louvain` node-clustering algorithm (Held et al., 2016). It discovered four clusters with 34, 23, 19, and 2 ASE nodes, respectively. The node colors in Figure 3 correspond to membership in each of the four clusters, node size is proportional to its degree, and edge width is proportional to its weight (the co-mention frequency).

7. Discussion and Conclusions

Lack of comprehensive medical information on ASEs of new medications is a challenge for stakeholders in healthcare, including patients, caregivers, and healthcare providers who make decisions using incomplete knowledge. For instance, physicians may prescribe medications without complete ASE knowledge, potentially harming patients.

Our approach offers stakeholders a proactive solution by enabling the active collection of social media data to unveil ASEs associated with new medications and to validate their confidence using the Overlap score. While this study focuses on a population biased towards younger individuals, the value of early warning remains significant. A decision support system based on our approach can monitor emerging public health issues related to ASEs, facilitating timely implementation of regulatory interventions, adjustments in treatment guidelines, and enhancements of patient safety measures. Moreover, our approach can be

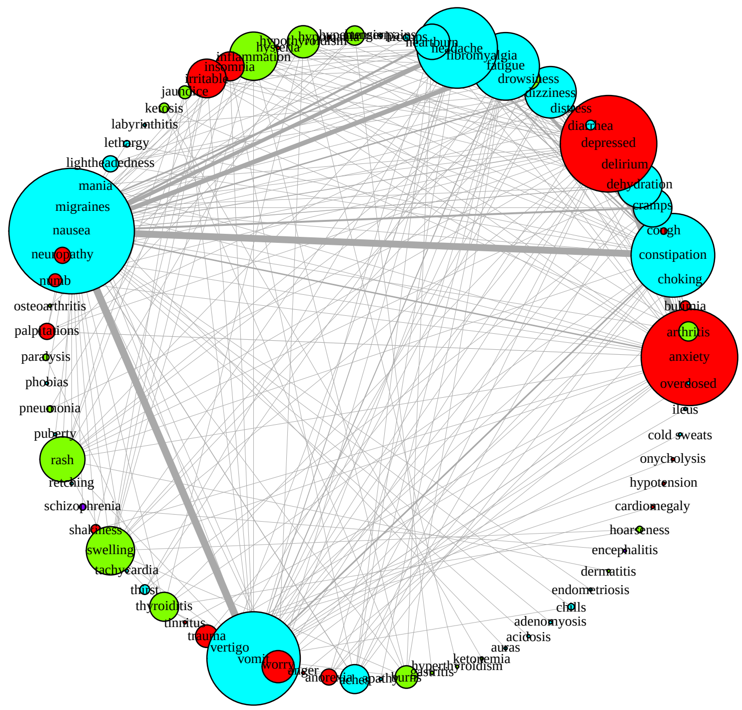

Figure 3. The ASE-ASE social media network. Node color indicate cluster affiliation.

modified to detect emerging issues or safety concerns related to any medication, not only the GLP-1 RA group, using large-scale social media data, providing substantial implications for pharmacological safety assessment and public health.

We address the pressing challenge of early warning about ASEs associated with newly released drugs, which present significant risks to public health and impose substantial costs on healthcare systems. To help pharmaceutical companies, healthcare regulators, and professionals with this vital task, we introduce a digital computational methodology designed to identify early on unreported ASEs by mining social media data, serving as an effective system that can complement existing pharmacovigilance sources.

In our approach to identifying ASEs in textual data, we deploy an NER model to analyze posts from X and Reddit discussing GLP-1 RAs, as well as academic papers studying GLP-1 RAs, ChatGPT, and manufacturers' reports. Many ASEs are included in all of our datasets, yet 14 out of 120 (11.6%) are reported exclusively by academic studies (Group 3), 13 (10.8%) solely by ChatGPT and manufacturers' reports (Group 4), and 15 (12.5%) only on social media (Group 5), thereby supporting hypothesis *H1*.

Social media captures a broader spectrum of user perspectives than other data sources, including those that may not be apparent in clinical research. In contrast with conventional approaches that merely verify whether identified ASEs are reported in biomedical datasets, we estimate their validity using the Overlap score. Our results indicate a significant correlation ($R^2 = 0.99$) between frequently mentioned ASEs on social media and established ASEs, supporting hypothesis *H2*. Consequently, our approach is able to distinguish which ASEs identified in social media are more likely novel.

Applying our methodology to the GLP-1 RA context proved that consolidating diverse data sources, especially social media, can help identify ASEs early.

References

Adjeroh, D., Beal, R., Abbasi, A., Zheng, W., Abate, M., & Ross, A. (2014). Signal fusion for social media analysis of adverse drug events. *IEEE Intelligent Systems*, *29*(2), 74–80.

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 and Insights from a range of social media platforms using a mixed-methods approach. *Brain Sciences*, *13*(11), 1503.

Back, H.-M., Choi, S.-A., & Kim, M. G. (2023). Adverse events of GLP-1 receptor agonists for weight loss: Twitter and a national pharmacovigilance database. *Drug Targets and Therapeutics*, *2*(1), 41–48.

Bartal, A., Jagodnik, K. M., Chan, S. J., Babu, M. S., & Dekel, S. (2023). Identifying women with postdelivery posttraumatic stress disorder using natural language processing of personal childbirth narratives. *American Journal of Obstetrics and Gynecology MFM*, *5*(3), 100834.

Basch, C. H., Narayanan, S., Tang, H., Fera, J., & Basch, C. E. (2023). Descriptive analysis of TikTok videos posted under the hashtag# Ozempic. *Journal of Medicine, Surgery, and Public Health*, *1*, 100013.

Berlin, J. A., Glasser, S. C., & Ellenberg, S. S. (2008). Adverse event detection in drug development: Recommendations and obligations beyond phase 3. *American Journal of Public Health*, *98*(8), 1366–1371.

Boe, B. (2023). *PRAW: The Python Reddit API Wrapper* [Accessed January 25, 2024]. https://praw.readthedocs.io/en/stable/

Chapman, B., & Chang, J. (2000). Biopython: Python tools for computational biology. *ACM Sigbio Newsletter*, *20*(2), 15–19.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, complex systems*, *1695*(5), 1–9.

de França, F. O., dos Santos, C. d. S., de Camargo Penteado, C. L., Goya, D. H., de Souza, L. M., Kamienski, C., di Genova, D. V. B., & Ramos, D. F. S. (2021). Dynamics of conflicts on the twitter social network: A case study on the use of chloroquine in brazil. *HICSS*, 1–10.

Derry, S., Kong Loke, Y., & Aronson, J. K. (2001). Incomplete evidence: The inadequacy of databases in tracing published adverse drug reactions in clinical trials. *BMC Medical Research Methodology*, *1*, 1–6.

Ding, J., Berleant, D., Nettleton, D., & Wurtele, E. (2001). Mining MEDLINE: Abstracts, sentences, or phrases? In B. 2002 (Ed.). World Scientific.

Elsevier. (2023). *Embase Database* [Accessed October 6, 2023]. https://www.embase.com/

Galeano, D., Li, S., Gerstein, M., & Paccanaro, A. (2020). Predicting the frequencies of drug side effects. *Nature Communications*, *11*(1), 4575.

Gao, Y., Duan, W., & Rui, H. (2022). Does social media accelerate product recalls? evidence from the pharmaceutical industry. *Information Systems Research*, *33*(3), 954–977.

Godinez, H., Xu, Q., McMann, T. J., Li, J., & Mackey, T. K. (2023). Analysis of online user discussions on Reddit associated with the transition of use between HIV PrEP therapy. *Frontiers in Public Health*, *11*.

Graves, R. L., Perrone, J., Al-Garadi, M. A., Yang, Y.-C., Love, J. S., O'Connor, K., Gonzalez-Hernandez, G., & Sarker, A. (2022). Thematic analysis of Reddit content about buprenorphine-naloxone using manual annotation and natural language processing techniques. *Journal of Addiction Medicine*.

Han, S. H., Safeek, R., Ockerman, K., Trieu, N., Mars, P., Klenke, A., Furnas, H., & Sorice-Virk, S. (2023). Public interest in the off-label use of glucagon-like peptide 1 agonists (Ozempic) for cosmetic weight loss: A Google Trends analysis. *Aesthetic Surgery Journal*, sjad211.

Held, P., Krause, B., & Kruse, R. (2016). Dynamic clustering in social networks using louvain and infomap method. In 2. T. E. N. I. C. (ENIC) (Ed.).

Jivani, A. G., et al. (2011). A comparative study of stemming algorithms. *International Journal of Computer Applications in Technology*, *2*(6), 1930–1938.

Joseph, A. (2023). *How ChatGPT 4 is transforming health care and medical education* [Accessed: February 5, 2024]. https://www.statnews.com/2023/07/20/chatgpt-gpt4-health-care-medical-education/

Juhi, A., Pipil, N., Santra, S., Mondal, S., Behera, J. K., Mondal, H., Pipil Sr, N., Santra Sr, S., & Behera IV, J. K. (2023). The capability of ChatGPT in predicting and explaining common drug-drug interactions. *Cureus*, *15*(3).

Kuhn, M., Letunic, I., Jensen, L. J., & Bork, P. (2016). The SIDER database of drugs and side effects. *Nucleic Acids Research*, *44*(D1), D1075–D1079.

Lancaster, T. (2023). Artificial intelligence, text generation tools and ChatGPT–Does digital watermarking offer a solution? *International Journal for Educational Integrity*, *19*(1), 10.

Lee, J.-Y., Lee, Y.-S., Kim, D. H., Lee, H. S., Yang, B. R., & Kim, M. G. (2021). The use of social media in detecting drug safety–Related new black box warnings, labeling changes, or withdrawals: Scoping review. *JMIR Public Health and Surveillance*, *7*(6), e30137.

Lee, P., Bubeck, S., & Petro, J. (2023). Benefits, limits, and risks of GPT-4 as an AI chatbot for medicine. *New England Journal of Medicine*, *388*(13), 1233–1239.

Lee, P., Goldberg, C., & Kohane, I. (2023). *The AI Revolution in Medicine: GPT-4 and Beyond*. Pearson.

Lennon, J. (2023). *Can you stomach it? A content analysis of TikTok videos and comments about weight loss with Ozempic*.

Martin, L., Hutchens, M., Hawkins, C., & Radnov, A. (2017). How much do clinical trials cost? *Nature Reviews Drug Discovery*, *16*(6), 381–382.

McMaster, C., Chan, J., Liew, D. F., Su, E., Frauman, A. G., Chapman, W. W., & Pires, D. E. (2023). Developing a deep learning natural language processing algorithm for automated reporting of adverse drug reactions. *Journal of Biomedical Informatics*, *137*, 104265.

Moore, P. W., Malone, K., VanValkenburg, D., Rando, L. L., Williams, B. C., Matejowsky, H. G., Ahmadzadeh, S., Shekoohi, S., Cornett, E. M., & Kaye, A. D. (2023). GLP-1 agonists for weight loss: Pharmacology and clinical implications. *Advances in Therapy*, *40*(3), 723–742.

Nugent, T., Plachouras, V., & Leidner, J. L. (2016). Computational drug repositioning based on side-effects mined from social media. *PeerJ Computer Science*, *2*, e46.

Okunogbe, A., Nugent, R., Spencer, G., Powis, J., Ralston, J., & Wilding, J. (2022). Economic impacts of overweight and obesity: Current and future estimates for 161 countries. *BMJ Global Health*, *7*(9), e009773.

Omiye, J. A., Gui, H., Rezaei, S. J., Zou, J., & Daneshjou, R. (2024). Large language models in medicine: The potentials and pitfalls: A narrative review. *Annals of Internal Medicine*.

Onakpoya, I. J., Heneghan, C. J., & Aronson, J. K. (2016). Worldwide withdrawal of medicinal products because of adverse drug reactions: A systematic review and analysis. *Critical Reviews in Toxicology*, *46*(6), 477–489.

Ortiz, E. (2004). Market Withdrawal of Vioxx: Is it Time to Rethink the Use of COX-2 Inhibitors? *Journal of Managed Care Pharmacy*, *10*(6), 551–554.

Ozurumba-Dwight, L., Soyinka, J., Ogbonna, C., Enwere, O., Iyevhobu, K., Iloka, O., Korie, M., & Kevin-Asogwa, L. (2022). Multi-purposely dynamic use of social media for health purposes: Scooping review of systematic reviews. *International Journal of Advanced Research in Biological Sciences*, *9*(6), 92–108.

Pearson-Stuttard, J., Holloway, S., Polya, R., Sloan, R., Zhang, L., Gregg, E. W., Harrison, K., Elvidge, J., Jonsson, P., & Porter, T. (2022). Variations in comorbidity burden in people with type 2 diabetes over disease duration: A population-based analysis of real world evidence. *EClinicalMedicine*, *52*.

Pritchett, J. (2023, May). *2022 prescription trends* [Accessed October 5, 2023]. https://www.pharmexec.com/view/2022-prescription-trends

Ridings, J. E. (2013). The thalidomide disaster, lessons from the past. *Teratogenicity Testing: Methods and Protocols*, 575–586.

Sayers, E. (2023). *E-Utilities Quick Start* [Accessed January 25, 2024]. https://www.ncbi.nlm.nih.gov/books/NBK25500/

Tarcar, A. K., Tiwari, A., Dhaimodker, V. N., Rebelo, P., Desai, R., & Rao, D. (2019). Healthcare NER models using language model pretraining. *arXiv preprint arXiv:1910.11241*.

Temperton, J. (2023). *GPT-3: How does the new language model work?* [Accessed: February 5, 2024]. https://www.sciencefocus.com/future-technology/gpt-3

Trujillo, J. M., Nuffer, W., & Smith, B. A. (2021). GLP-1 receptor agonists: An updated review of head-to-head clinical studies. *Therapeutic Advances in Endocrinology and Metabolism*, *12*, 2042018821997320.

U.S. Food and Drug Administration. (2023). *FDA Adverse Event Reporting System (FAERS) Public Dashboard* [Accessed October 5, 2023]. https://www.fda.gov/drugs/questions-and-answers-fdas-adverse-event-reporting-system-faers/fda-adverse-event-reporting-system-faers-public-dashboard

Watanabe, J. H., McInnis, T., & Hirsch, J. D. (2018). Cost of prescription drug–related morbidity and mortality. *Annals of Pharmacotherapy*, *52*(9), 829–837.

World Health Organization. (2021). *Obesity and overweight* [Accessed October 5, 2023]. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight

Zhang, Z., Zhang, Q., Tan, Y., Chen, Y., Zhou, X., Liu, S., & Yu, J. (2023). GLP-1RAs caused gastrointestinal adverse reactions of drug withdrawal: A system review and network meta-analysis. *Frontiers in Endocrinology*, *14*.