Social Media for Exploring Adverse Drug Events Associated with Multiple Sclerosis

Nevine Nawar  
Alexandria University  
nawrh@yahoo.com

Loknath Sai Ambati  
Dakota State University  
LoknathSai.Ambati@trojans.dsu.edu

Omar El-Gayar  
Dakota State University  
Omar.El-Gayar@dsu.edu

Gridhar Bojja  
Dakota State University  
Gridhar.Bojja@trojans.dsu.edu

Abstract

Multiple Sclerosis (MS) affects 400,000 people in the USA and almost 2.5 million people worldwide. There is no cure for MS. A variety of disease-modifying therapies are currently available. They aim to reduce disease activity that ultimately leads to disability. However, such drugs have adverse effects that vary widely among patients making the choice of a suitable drug particularly challenging. With the proliferation of social media, this research aims to understand the perspective of people with MS on social media (Twitter) in regard to Adverse Drug Events (ADEs) and to analyze ADEs as perceived by MS patients. This study helps in understanding ADEs associated with MS drugs and can further inform future medical research by highlighting and prioritizing additional clinical trials needed to better assess such adverse drug effects.

1. Introduction

Multiple Sclerosis (MS) is considered to be an immune-mediated disease that occurs when the immune system attacks the central nervous system (CNS). The immune system’s attack damages myelin, oligodendrocytes, and underlying nerve fibers, impairing the ability of the CNS to send or receive messages [1]. The most unsettling aspect of MS is that its cause is still not clear. Based on the clinical criteria, frequency of clinical relapses, time taken for disease progression, and lesion development as shown by MRI findings, MS is classified into four types, namely Relapsing-remitting MS (RRMS), Secondary Progressive MS (SPMS), Primary progressive MS (PPMS) and Clinically Isolated Syndrome (CIS). Almost 85% of people having MS fall under RRMS, which is characterized by relapses followed by periods of complete or partial recovery [2]. Some early signs and symptoms of MS include tingling and numbness, optic neuritis, fatigue, pain, muscle spasm and weakness, balance problems, bladder issues, cognitive difficulties, depression, and anxiety. MS patients are prescribed disease-modifying drugs to delay neurological disability progression [3], [4]. ADEs are unfavorable and unintentional signs or symptoms associated with the use of a medication [5]. The Center for Disease Control (CDC) advises adults to follow precautions to reduce the risk of ADEs by keeping a list of their medications, following directions, asking questions, keeping up with any blood testing recommended by the doctor, and taking all the medicines strictly as directed. The burden of ADEs to the nation’s economy is estimated at $30.1 billion annually [6]. Prescription drugs for MS are known to result in a variety of adverse effects, some of which are very serious. While many of these adverse effects are reported in the literature, their extent varies substantially among MS patients making it particularly challenging to identify a suitable drug for a particular patient. This is also a recurring problem as patients often need to switch medications during the course of treatment.

With the proliferation of social media, health became a prevalent topic that people discuss on these platforms [7]. Today, social media has evolved as a fertile platform for patients to exchange information regarding their experiences with a life-changing condition such as MS. This includes sharing information about symptoms, disease progression, and adverse drug effects. Such wealth of information creates opportunities for healthcare professionals and researchers to gain further insights into various disease conditions as they relate to symptoms, adverse drug effects and experiences.

A few studies have proposed ADE detection methods utilizing social media data [8]. These social media platforms can be patient discussion forums or social networking platforms such as Facebook and Twitter. To the best of our knowledge, there are only two studies that utilized social media data for assessing patterns and sentiments regarding MS treatments. One exploratory study investigated opinions regarding MS
treatments and found that oral MS treatments tend to show higher sentiment scores than injectable treatments [9], while another study explored patterns of treatment switching by MS patients [10].

Accordingly, in this research, we complement prior research by extracting prevalent MS ADEs as reported on social media (Twitter). Further, we conduct a sentiment analysis that provides a holistic perspective of patients’ perception of various drugs given their side effects. The analysis covers 18 different MS drugs and leverages text mining to quantify the volume and sentiments of the ensuing discussion, and to capture the dominant ADEs categories for each drug. Overall, this sheds light on the safety and effectiveness of drugs used for MS treatment. The outcome can help patients and clinicians identify appropriate courses of action. Further, the results can inform future medical research by highlighting and prioritizing additional clinical trials needed to better understand such adverse drug effects.

The remainder of the paper is organized as follows: Section 2 provides a brief review of related work while section 3 describes the research methodology. Section 4 presents the results obtained while section 5 includes an interpretation and discussion of our findings with reference to extant research. Section 6 concludes by summarizing the research and discussing the limitations and research scope for the future.

2. Related work

ADEs post-marketing surveillance can rely on different types of data obtained from various sources [11] such as Electronic Health Records (EHR), and voluntarily reported data on such systems as the Food and Drug Administration’s (FDA) Adverse Event Reporting System (FAERS) [12]. However, these sources could suffer from low reporting rates and limited scope. Accordingly, a number of studies utilized social media platforms as a source potentially offering a relatively broad and timely data to understand and address health-related issues [13], [14]. Of particular interest to the medical informatics research community is leveraging social media to monitor ADEs [8]. For example, a few studies [15]–[18] used Twitter to detect and track ADEs while other studies relied on healthcare social media forums [19], [20]. Some studies on ADEs proposed and implemented near real-time pipelines and models for ADE detection and prediction [18], [21]. ADEs studies utilized a variety of techniques such as but not limited to; association rule mining [20], natural language processing (NLP) and deep learning [22], extracting deep linguistic features [23], ensemble classification [24], statistical modeling [25], and topic modeling.

With respect to MS, one study used an exploratory approach to analyze sentiments pertaining to MS treatments using twitter data [9]. The study identified tweets mentioning MS treatments and used them to analyze patients’ sentiments. The study conducted in 2014 did not analyze adverse effects nor did it include a number of now available newer medications. Another study used automated listening with filtering and analysis of data to examine the patterns of treatment switching between infectable, oral, and infusion therapies for MS patients [10]. Accordingly, this study, complement prior research by focusing on the ADEs of MS treatments from user contributed contents, namely Twitter, as well as further exploring patients’ sentiments towards MS medications on a wider selection of these drugs.

3. Methodology

We followed the pipeline in Figure 1 for performing this research. The pipeline included four stages: data discovery, data collection, data preparation, and data analysis. We employed Crimson Hexagon (CH) (now Brandwatch) to understand the public’s perspective on MS treatments and ADEs. CH is an AI-powered consumer insights tool used to scrape data from social media and analyze it for insights. It is a social media analytics tool that utilizes machine learning and text mining approaches developed by the founder of CH [26].

![Figure 1. ADE detection pipeline](image-url)
The Data discovery stage consisted of creating a Boolean search query for CH using the drug names. The search query was limited to tweets in the English language, from 1 Jan 2010 to 1 Jan 2020. The MS treatments considered for this study were Avonex, Betaseron, Copaxone, Glatiramer Acetate, Extavia, Glatopa, Plerigridy, Rebif, Aubagio, Gilenya, Mavenclad, Mayzent, Tecfidera, Vumerity, Lemtrada, Novantrone, Ocrevus, and Tysabri. We standardized the search based on the brand names of the drugs, as some of the active ingredients (generic name) are indicated for other conditions. CH search query employed case insensitive approach for searching the tweets.

In the data collection stage, the search query retrieved the required data from CH for manual examination for relevant posts. The retrieved data for all the drugs was examined particularly for unusually high volume of tweets on a particular day or time period to help us redefine the inclusion and exclusion criteria in the data discovery stage. This examination of tweets was manually done. Overall, we found that there was quite an unusual activity due to retweets. Further, tweets containing hyperlinks were considered to be irrelevant as they often link to advertisement or news articles about a particular medication rather than the patients’ experience. Tweets with a stock market symbols of pharmaceutical companies were excluded. Ultimately, we developed the search query template shown below:

("Drug Brand Name 1" OR "Drug Brand Name 2" OR "Drug Brand Name 3" OR ..........)
AND – (http* OR RT)
AND – (Merck OR Sereno OR Pfizer OR Biogen OR Bayer OR Novartis OR Teva OR Momenta OR Sandoz OR Sanofi OR Genentech OR "Roche Group")

Search queries for some drugs had to be customized based on examining the tweets related to the drugs respectively. For example, we found a Twitter handle named ‘@aubagio’. However, there was no relation to MS for tweets coming from this handle.

Once the finalized query was run in CH, we proceeded to the data analysis stage. We analyzed the results for each drug tweets by excluding references to other drugs to make sure each tweet reflects only one medication. Although some tweets comparing different medications would be lost, it was the best way to make sure we performed a valid analysis on each drug. We analyzed the sentiments and emotions towards MS treatment using the pre-trained classifier in CH. We employed the ReadMe algorithm [26] to analyze the posts that fall into predefined categories representing various adverse drug effects. The MS ADEs predefined categories follow the list of common side effects defined by the National MS Society for each drug. The ReadMe algorithm is appropriate when the objective of the study is to categorize the posts into a pre-defined number of categories and calculate the proportion of posts in each category [26]. This provides unbiased categorization and proportions even when the customized classifier performs poorly.

4. Results

A total of 51,362 unique tweets mentioned MS treatments after removing the tweets containing manufacturer or company names and hyperlinks. Overall, 16% of the tweets included age identification, while 65% identified gender. The available demographics of tweets are shown in Figure 2.

![Figure 2a. Distribution of tweets by age group](image1)

![Figure 2b. Distribution of tweets by gender](image2)
approximately very similar. Extavia and Tysabri had the highest positive sentiment proportion. Rebif and Avonex had the highest negative sentiment proportion. Copaxone, Ocrevus, Tysabri, Lemtrada, Plegridy, Gilenya, Mayzent and Extavia had slightly more positive sentiments than negative.

![Figure 3. Sentiment and emotion distribution of tweets](image)

Table 1. Distribution of Tweets Data for Each Drug

<table>
<thead>
<tr>
<th>Drug</th>
<th># of Tweets (with Exclusion Criteria)</th>
<th># of Tweets (without Exclusion Criteria)</th>
<th>Proportions of tweet sentiments (with exclusion criteria)</th>
<th>Proportions of tweet sentiments (without exclusion criteria)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Aubagio</td>
<td>1,156</td>
<td>1,526</td>
<td>0.282</td>
<td>0.358</td>
</tr>
<tr>
<td>Avonex</td>
<td>3,945</td>
<td>5,099</td>
<td>0.340</td>
<td>0.361</td>
</tr>
<tr>
<td>Betaseron</td>
<td>910</td>
<td>1,303</td>
<td>0.245</td>
<td>0.325</td>
</tr>
<tr>
<td>Copaxone</td>
<td>8,733</td>
<td>10,421</td>
<td>0.340</td>
<td>0.330</td>
</tr>
<tr>
<td>Extavia</td>
<td>265</td>
<td>324</td>
<td>0.668</td>
<td>0.112</td>
</tr>
<tr>
<td>Gilenya</td>
<td>5,007</td>
<td>5,969</td>
<td>0.315</td>
<td>0.255</td>
</tr>
<tr>
<td>Lemtrada</td>
<td>3,512</td>
<td>4,009</td>
<td>0.374</td>
<td>0.306</td>
</tr>
<tr>
<td>Mavenclad</td>
<td>234</td>
<td>293</td>
<td>0.345</td>
<td>0.345</td>
</tr>
<tr>
<td>Mayzent</td>
<td>41</td>
<td>53</td>
<td>0.300</td>
<td>0.250</td>
</tr>
<tr>
<td>Novantrone</td>
<td>79</td>
<td>105</td>
<td>0.105</td>
<td>0.235</td>
</tr>
<tr>
<td>Ocrevus</td>
<td>2,430</td>
<td>2,912</td>
<td>0.364</td>
<td>0.326</td>
</tr>
<tr>
<td>Plegridy</td>
<td>333</td>
<td>450</td>
<td>0.333</td>
<td>0.297</td>
</tr>
<tr>
<td>Rebif</td>
<td>4,376</td>
<td>5,383</td>
<td>0.318</td>
<td>0.372</td>
</tr>
<tr>
<td>Tecfidera</td>
<td>4,881</td>
<td>5,958</td>
<td>0.338</td>
<td>0.352</td>
</tr>
<tr>
<td>Tysabri</td>
<td>10,760</td>
<td>12,365</td>
<td>0.430</td>
<td>0.310</td>
</tr>
<tr>
<td>Vumerity</td>
<td>15</td>
<td>24</td>
<td>0.200</td>
<td>0.200</td>
</tr>
<tr>
<td>Total</td>
<td>46,677</td>
<td>56,194</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4 depicts the distribution of tweets over time for MS treatments. The distribution results are similar to the study by Ramagopalan et al. [9] for that specific time period. The relative proportions of the adverse effects for each drug is shown in Figure 5. Adverse events for Avonex, Betaseron, Extavia, Rebif, and Plegridy were quite similar, as these medications have the same active ingredient, “Interferon Beta”. More than 50% of Aubagio’s side effects were attributed to diarrhea, hair loss/thinning, and liver problems. More than 61% of Copaxone side effects were attributed to injection site problems and shortness of breath.

Liver problems alone contributed to 45% of Gilenya’s side effects. Liver problems, low white blood cells, and flu/fever/body aches accounted for 68% of interferon-beta’s side effects. Headache, rashes, and kidney problems contributed to 64% of the Lemtrada’s side effects. Mavenclad’s side effects were mostly attributed to hair loss/thinning (30%), low white blood cell count (25%), mouth sores and/or shingles (25%), and rashes (21%). Headache and low heart rate constituted 58% of Mayzent’s side effects. Novantrone side effects were caused by skin rash/reactions (75%), and infections (25%). Ocrevus’s side effects were mainly itchy skin (40%), rash (33%), and throat irritation (27%). Hot flushes and nausea contributed to 65% of Tecfidera’s side effects. Headache and dizziness contributed to 56% of Tysabri’s side effects.
Figure 4. Distribution of tweets over time for each treatment
Figure 5. Tweet classification of ADEs for each treatment (with exclusion criteria)
5. Discussion

Based on the available demographics shown in Figure 2, women tweeted more about MS treatments compared to men. People who were older than 35 years tweeted more compared to other age groups. According to the National Multiple Sclerosis Society, MS is three time more common among women and most patients are diagnosed between the ages of 20 and 50 years [27].

The sentiment analysis in this study is consistent with Ramagopalan et al. [9]. The findings of the discovered sentiment scores from Table 1 show that treatments such as Extavia, Gilenya, Lemtrada, Mayzent, Ocrevus, and Tysabri have higher positive than negative tweet proportions. Extavia is the only injectable treatment in this category. Lemtrada, Tysabri, and Ocrevus are infusion treatments, while Gilenya and Mayzent are oral treatments. Mavenclad and Tecfidera, both oral treatments, appear to have a neutral sentiment among the public. Vumerity is new to the market and limited data is available for it. Based on the tweets and sentiment analysis, Aubagio, an oral drug, has a substantial negative sentiment due to the side effects associated with it.

With respect to the distribution of tweets over time (figure 4), the tweets volume for each treatment shows spikes at various time points. Further investigation into the Twitter corpus revealed that spikes in the tweet volume coincide with news regarding the FDA approval of new MS drugs. Each treatment’s sentiment scores clearly show a positive intent and excitement for newer treatments that hit the market. Most of the treatments have word clusters related to social support and symptoms associated with the treatments. For instance, the word clusters for Extavia have “Cost” as the frequent word, and the other words in the clusters are “iPhone” and related terms. When the tweet corpus for Extavia was further investigated, it revealed an interesting story: people compare the cost of Extavia treatment to that of iPhone revealing the public discourse on the soaring cost of MS drugs. The finding is in accordance with an existing concern over the high pricing of MS specialty medications [28], [29].

To the best of our knowledge, we present here the first analysis on the proportion of ADEs for each MS treatment. The discovered symptoms and ADEs are mapped to the valence of tweets for each treatment as summarized in Figure 5. Adverse events for Avonex, Betaseron, Extavia, Rebif, and Plegridy were quite similar, as these medications have the same active ingredient, “Interferon Beta”. Liver problems, low white blood cells, and flu/fever/body aches contributed to 68% of interferon-beta’s side effects. Nearly two-thirds (61%) of Copaxone’s side effects were caused by injection-site reactions and shortness of breath. The findings are generally consistent with the listed side effects [30] for this group of medications. Although flu-like symptoms and injection-site reactions are listed among the most common side effects [31] for interferon-beta, they made up only 17% and 9% respectively of discussed ADEs. One explanation is that patients focused on the more serious side effect of liver problems (34%) in their tweets.

In the oral medication group, slightly more than half (51%) of Aubagio’s side effects were attributed to diarrhea, hair loss/thinning, and liver problems. Liver problems alone accounted for nearly half (45%) of Gilenya’s side effects. Hot flushes and nausea contributed to 65% of Tecfidera’s side effects. Mavenclad side effects were primarily attributed to hair loss/thinning (30%), low white blood cell count (25%), mouth sores and/or shingles (25%), and rashes (21%). Headache and low heart rate constituted 58% of Mayzent’s side effects. The findings for the oral group were mostly consistent with the listed side effects of the included medications [30]. However, Mavenclad, an oral immunosuppressant that is generally indicated to patients who have had an inadequate response or could not tolerate one or more MS therapies, had markedly fewer discussed ADEs as compared to the listed side effects [32], [33]. This can be explained by the recent FDA approval for the medication as well as its being a less commonly prescribed one, which could have affected data availability.

In the infusion medications group, headache, rashes, and kidney problems contributed to 64% of Lemtrada’s side effects. Novantrone side effects were caused by skin rash/reactions (75%), and infections (25%). Ocrevus’s side effects were mainly itchy skin (40%), rash (33%), and throat irritation (27%). Headache and dizziness contributed to 56% of Tysabri’s side effects. The adverse events such as Progressive multifocal leukoencephalopathy (PML) are discussed in tweets related to Tysabri. Most of the tweets that talk about PML usually discuss the risk of a dangerous brain infection, contributing to fear and panic. The risk of PML with Tysabri can be addressed by the implementation of a risk management plan (RMP) [34]. The findings for the infusion medications were mostly consistent with the listed side effects except for Novantrone. Its most common listed side effects include nausea and hair thinning, yet our findings show skin rash/reactions to account for 75% of discussed ADEs. It is important to note that the risk of cardiac toxicity posed by this medication shadows...
The results and recommendations of this study are relatively broader and timely. User volunteering reporting systems by relying on surveillance conducted through centralized MS treatments and the patient’s tolerance of the side effects may warrant a switch to other treatments. Based on the course of MS and individual patient’s response, second-line or third-line medications may be selected. The ADEs identified in this study for each treatment is consistent with the National Multiple Sclerosis Society’s comparisons of different drugs and their ADEs.

This study emphasizes the significance of further adoption of text mining techniques for detecting adverse drug events. Multiple sclerosis is not common but is a serious cause of neurological disability throughout adult life. Its prevalence has seen a marked increase since 1990. From a practical standpoint, automatically analyzing social media users posting with machine learning techniques will provide the medical community with valuable insights on patients’ behavior and reaction towards ADEs, symptoms, and treatments associated with MS. These may aid in recognizing the current state of MS treatments and the ADEs associated with it and can improve healthcare providers' understanding of the commonness of a particular ADE associated with an MS treatment as reported by patients.

6. Conclusion

This research focused on the ADEs of MS treatments from user-contributed contents on social media. The sentiment analysis and categorization of ADEs for various MS treatments indicate that the discussed side effects for injectable medications (Interferon-beta (Avonex, Betaseron, Extavia, Plegridy and Rebif), and Copaxone) are slightly higher when compared to infusion or oral medications. The underlying reasons for patients’ dislike of injectable medications are the side effects caused by the injection itself (frequent use of needles, injection-site reactions, and rashes). Furthermore, there is a noticeable excitement for new treatments such as Vumerity, Mavenclad, and Mayzent. Overall, the results shed light on the safety and effectiveness of drugs used for MS treatment and complement other post-marketing surveillance conducted through centralized volunteering reporting systems by relying on a relatively broader and timely user-contributed data. The results and recommendations of this study can help patients and clinicians identify appropriate courses of action and can inform future medical research by highlighting and prioritizing additional clinical trials needed to better understand such adverse drug effects.

Some limitations of this study include constraints related to the data, namely, language (English), data source (Twitter), and analytics techniques. Twitter data might not reflect the actual patient population. According to Wojcik and Hughes, Twitter users tend to be a much younger audience, with 10% of users creating 80% of all tweets published. Future research may consider expanding to other languages and social media platforms such as health-related forums, Reddit, and Facebook. While CH provides extensive analytic and visualization capabilities, it may be limiting with respect to the diversity and complexity of the available techniques. For example, other machine learning such as topic mining coupled with thematic content analysis based on demographics and ADE type may shed further insights regarding the ADEs of various MS treatments.

MS, a progressive degenerative disease of the central nervous system, has yet no cure. Current treatment options aim to reduce the contributory immune system’s attacks thereby reducing the number of relapses and slowing the progression of disability. Different oral, injectable, and infusion drugs for the disease are available with more drugs being introduced to the market. As MS treatment is long-term, ADEs present a challenge that affects the choice of drugs and the decision to remain on a given drug or change to another. As a result, the thorough understanding of these ADEs will be an ongoing endeavor.

7. References


