

Interrogating social media sources to gain insight on opioid-induced constipation symptoms as reported by patients

Steven Brewer¹ and Tessa Kennedy-Martin²

¹Text Mining Solutions, York, United Kingdom; ²Kennedy Martin Health Outcomes, Brighton, United Kingdom

Objective

This exploratory project aimed to employ text mining techniques to gather information from a range of social media sources, including patient forums, to better understand patient-reported symptoms associated with opioid-induced constipation (OIC).

Social media data are increasingly used for pharmacovigilance, as highlighted in a review by Sarker et al (2015), but have not yet been widely used to gather insight on patient-reported outcomes (PRO). The importance of gathering PRO data – where the status of an individual's health condition is reported directly by the patient, without interpretation by a clinician or anyone else – is now well accepted. Indeed, as far back as 2010 the Pew Research Center's survey, *The Social Life of Health Information* (Fox, 2011), reported that 2% of patients and 6% of caregivers shared their health experiences online.

OIC is the most common adverse effect associated with prolonged use of opioids; it has been estimated that between 40% and 95% of patients develop OIC (Kumar et al, 2014). Its presence can undermine the clinical utility of opioid treatment. In a multinational, Internet-based survey that was designed to assess impact of opioid-induced bowel dysfunction in patients receiving opioid therapy for chronic pain and taking laxatives, it was found that one-third of the patients had skipped doses, reduced or stopped their use of opioid therapy in order to resolve their gastrointestinal side effects (Bell et al, 2009).

What is text mining?

Text mining is the term used to describe the process by which important information can be found within documents in an electronic format using computer technology. In much the same way that someone can read a newspaper and pick out salient facts, opinions, and inferences, we emulate this process using computer coding based on rules and algorithms. Text mining may also be referred to as 'natural language processing' or 'computational linguistics'. It encompasses a multitude of specialist disciplines, skills, and components such as machine learning, parsing, and ontologies and dovetails with wider themes such as text analytics, big data, deep learning, and artificial intelligence.

Method

This was an exploratory project that used text mining techniques. We used web crawlers (Visual Web Spider/Win Web Crawler) to visit a list of 174 links to websites likely to contain information relevant to OIC. Any page containing a mention of terms related to OIC (e.g. OIC, opiate induced constipation, opiate related constipation, opiate constipation, opiate-induced constipation, opioid-related constipation, opioid constipation) was saved for further processing. The crawl depth¹ was set to level 4 and this yielded over 40,000 candidate pages.

The identified pages were text mined to extract mentions of symptoms associated with OIC. To do this, we built a machine learning-based application using GATE Learning Framework (<https://github.com/GateNLP/gateplugin-LearningFramework/wiki>) and a training corpus created specifically for this application. A model was built using a training corpus containing positive and negative examples, evaluated against a gold standard corpus and ultimately deployed in the application pipeline to process hitherto unseen records (web pages). This involved the use of a Library for Support Vector Machines (LIBSVM) classifier at the sentence level to identify documents containing sentences with phrases of significance.

The narrative content of the posts was anonymised and analysed manually. All relevant results were exported to an SQL database. We generated a feature-rich CRUD² web application to enable direct access to the results with full search/filter and export functionality. Doing this facilitated the selection of records that mentioned a specific symptom. These were then exported from the database and the results presented using data visualisation packages (QlikView® and Microsoft Power BI).

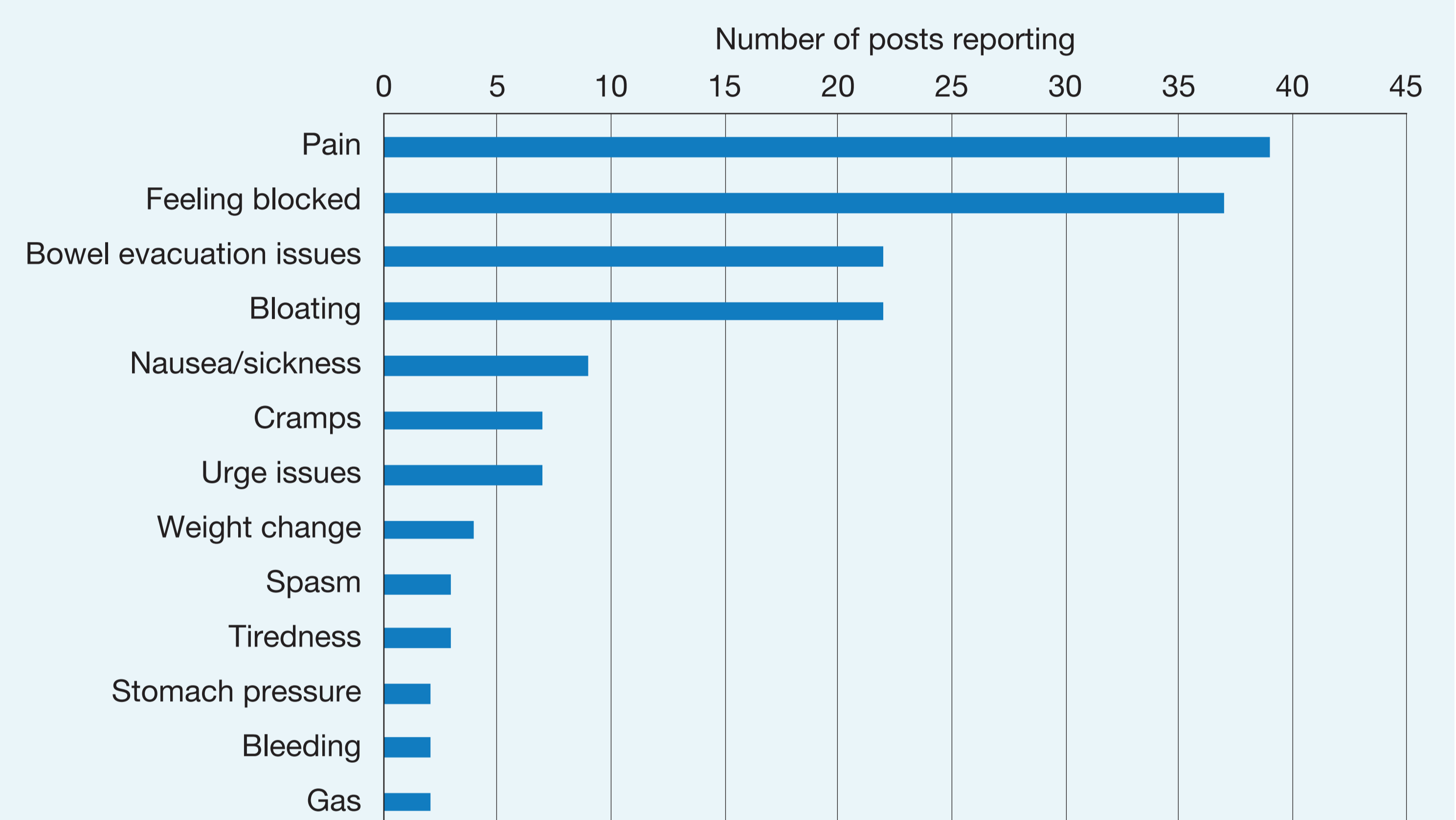
¹The crawl depth is the number of levels of referred links the crawler will follow: a setting of 1 will follow links only from the first page, level 2 follows links from the first page plus links on referred pages, etc.

²Create, Read, Update, and Delete operations are the basic manipulations of database records.

Results

Among the 42,000 web pages retrieved, we found 128 posts that described symptoms of OIC as reported by patients. Some posts mentioned more than one symptom. The most common symptoms were having pain (39 posts), feeling blocked (37 posts), evacuation issues (22 posts), and bloating (22 posts). Other symptoms included anal spasm, weight change, cramps, nausea/sickness, tenderness, stomach pressure, headache, urge, stitch, leakage, anus not closing, gas, brain fog, bleeding, and feeling tired. **Figure 1** shows symptoms reported by more than one post. Patients also reported fear (of rupture and of piles).

Figure 1: Symptoms (>1 post) of OIC as reported in social media posts



Conclusions

Using data from social media posts, we identified a wide range of symptoms being reported by patients in relation to their OIC. These symptoms were much broader than not being able to have a bowel movement: they included other problems such as pain, bloating, and spasm.

More generally, this exploratory research project shows that social media can be an important source of insight on patient symptoms. Researchers can quickly gain access to very large amounts of data and the insights could be invaluable when designing further qualitative or quantitative research or in better understanding the burden of a condition.

The use of social media data for healthcare research is in its infancy, and – as highlighted by Swirsky et al (2014) – there is little guidance in this area. Studying social media posts might be seen simply as a direct observation of public behaviour, but as this area of research evolves it will become increasingly crucial to consider confidentiality and ethical issues.

References

- Bell T, Panchal SJ, Miaskowski C, et al. The prevalence, severity, and impact of opioid-induced bowel dysfunction: Results of a US and European Patient Survey (PROBE 1). *Pain Med* 2009;10:35–42.
- Fox S. *The Social Life of Health Information*, 2011. Pew Research Center. Available at <http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011> (accessed Oct 2017).
- Kumar L, Barker C, Emmanuel A. Opioid-induced constipation: Pathophysiology, clinical consequences, and management. *Gastroenterol Res Pract* 2014;2014:141737.
- Sarker A, Ginn R, Nikfarjam A, et al. Utilizing social media data for pharmacovigilance: A review. *J Biomed Inform* 2015;54:202–212.
- Swirsky ES, Hoop JG, Labott S. Using social media in research: New ethics for a new meme? *Am J Bioeth* 2014;14:60–61.



www.kmho.co.uk



www.textminingsolutions.co.uk