

VU Research Portal

Using structural topic modelling to reveal patterns in reports on opioid drugs in a pharmacovigilance database

Lösch, Lea; Brown, Patrick; van Hunsel, Florence

published in

Pharmacoepidemiology and Drug Safety
2022

DOI (link to publisher)

[10.1002/pds.5502](https://doi.org/10.1002/pds.5502)

document version

Publisher's PDF, also known as Version of record

document license

Article 25fa Dutch Copyright Act

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Lösch, L., Brown, P., & van Hunsel, F. (2022). Using structural topic modelling to reveal patterns in reports on opioid drugs in a pharmacovigilance database. *Pharmacoepidemiology and Drug Safety*, 31(9), 1003-1006. <https://doi.org/10.1002/pds.5502>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

BRIEF REPORT

WILEY

Using structural topic modelling to reveal patterns in reports on opioid drugs in a pharmacovigilance database

Lea Lösch¹ | Patrick Brown² | Florence van Hunsel³ 

¹Athena Institute, Vrije Universiteit Amsterdam, The Netherlands

²AISSR, University of Amsterdam, Amsterdam, The Netherlands

³Netherlands Pharmacovigilance Centre Lareb, 's-Hertogenbosch, The Netherlands

Correspondence

Florence van Hunsel, Netherlands Pharmacovigilance Centre Lareb, 's-Hertogenbosch, The Netherlands.
Email: f.vanhunsel@lareb.nl

Abstract

Background: Adverse drug reaction (ADR) reports in pharmacovigilance databases often contain coded information and large amounts of unstructured or semi-structured information in plain text format. The unstructured format and sheer volume of these data often render them neglected. Structural topic modelling (STM) represents a potentially insightful way of harnessing these valuable data and to detect grouping or themes in spontaneous reports to aid signal detection.

Purpose: This was an explorative study of the potential for structural topic modelling to identify useful patterns in ADR reports involving opioid drugs in a pharmacovigilance database.

Methods: A dataset of ADR reports on opioid drugs reported to the Netherlands Pharmacovigilance Centre Lareb from 1991 to December 2020 was used, comprising a total of 3069 unique reports. Qualitative text analysis was combined with STM, an automated text analysis method, to examine these data.

Results: In reports submitted directly by patients and healthcare professionals, 11 meaningful topics were identified, whereby patient experience reports, particularly in relation to pain (relief), and the timing of intake and ADRs of tramadol and paracetamol, were the most common. Of the 12 topics identified in reports received via marketing authorization holders, patch and skin-related side effects, addiction and constipation were the most prevalent.

Conclusions: The STM-based analysis identified information that cannot always be captured by coding with the Medical Dictionary for Regulatory Activities (MedDRA[®]). The identified topics reflect findings in the literature on opioids.

KEYWORDS

adverse drug reaction, opioid drugs, pharmacovigilance, structural topic modelling

Key Points

- Data from pharmacovigilance centres can provide important insights into the risks associated with opioid products in daily practice.
- Reports in pharmacovigilance databases often contain coded information and a large amount of unstructured or semi-structured information in plain text format.
- Structural topic modelling represents a potentially insightful way of harnessing these valuable data and to detect grouping or themes in spontaneous reports to aid signal detection.
- The topics found in this study on opioid drugs include information that cannot always be captured in Standardized MedDRA[®] terms.

Plain Language Summary

Pharmacovigilance centres receive reports on adverse drug reactions (ADRs) from healthcare professionals and patients with the purpose of early identification of potential drug safety issues. Certain elements of information in these reports are coded, for instance with the medical dictionary for regulatory activities (MedDRA®). However, there is a large amount of unstructured or semi-structured information in these reports available. The unstructured format and sheer volume of these data often render them neglected. Structural topic modelling (STM), an automated text analysis method, represents a potentially insightful way of harnessing these valuable data and to detect grouping or themes in these reports to aid the detection of potential drug safety issues. This was an explorative study of the potential for STM to identify useful patterns in ADR reports in a pharmacovigilance database. This method was tested on a dataset of opioid reports, received by the Netherlands Pharmacovigilance Centre Lareb from 1991 to December 2020. A total of 3069 unique reports were included. The STM-based analysis identified information that cannot always be captured by coding with MedDRA®. The identified topics reflect findings in the literature on opioids.

1 | BACKGROUND

The use of highly potent opioids can lead to serious health risks among users and the widespread (long-term) use and abuse of these products is concerning.¹ Data from pharmacovigilance centres can provide important insights into the risks associated with these products in daily practice.²⁻⁴ Schifano et al.³ for instance analyzed the spontaneous reports relating to fentanyl misuse, abuse, dependence and withdrawal-related adverse drug reactions (ADRs) as identified in several databases for spontaneous reporting used worldwide. For that analysis, ADRs that were recorded according to the Medical Dictionary for Regulatory Activities (MedDRA) were analysed⁵ and trends in the number of cases on misuse/abuse/dependence/withdrawal issues and very high-dosage intake were identified.³ This pilot study focuses on opioids due to their societal and medical relevance as well as the increase in ADR reports reaching the Lareb database, especially since 2014/2015.

ADR reports in pharmacovigilance databases often contain coded information and a large amount of unstructured or semi-structured information in plain text format.⁶ For instance patient reports generally involve a large amount of unstructured information in the narrative about the impact of the ADR and circumstances of use.⁷ The unstructured format and sheer volume of these data often render them neglected.⁶ The application of natural language processing/text mining of reports in a pharmacovigilance database is a potential way to make use of the rich insights from patient ADR narratives.^{6,8} Structural topic modelling (STM)⁹ represents a potentially insightful way of harnessing these valuable data and to detect grouping or themes in spontaneous reports to aid signal detection.

2 | PURPOSE

This was an explorative study of the potential for structural topic modelling to identify useful patterns in ADR reports involving opioid drugs in a pharmacovigilance database.

3 | METHODS

A dataset of ADR reports on opioid drugs (ATC code¹⁰ 'N02A') reported to the Netherlands Pharmacovigilance Centre Lareb from 1991 to December 2020 was used, comprising a total of 3069 unique reports. ADR reports were recorded according to the MedDRA®. The reports had either been received directly by the pharmacovigilance centre from patients, via healthcare professionals who fill in the reporting form on the Lareb website, or indirectly via the Marketing Authorization Holder (MAH).

Qualitative text analysis was combined with STM, an automated text analysis method, to examine this existing but untapped data source. Topic modelling algorithms are a range of machine learning methods for discovering manifest and latent topics in a large collection of text documents.¹¹ "Topics" are formed based on the co-occurrence of certain words.¹² This is rooted in the linguistic assumption that words which tend to systematically appear together (co-occur) across multiple texts are also associated thematically. A topic is thus a collection of words representing interpretable themes within documents, which in our dataset are reports submitted to Lareb.⁹ STM is a specific variant of topic models, which allows the inclusion of document-specific metadata, such as 'date' or 'author'. STM was thus employed to (a) identify various topics in the reports and (b) investigate topic prevalence as a function of covariates such as date of receipt, source of report (MAH/Lareb) and type of drug. A detailed qualitative analysis of each topic resulting from the topic model enabled an evaluation of the quality of the model and the detection of more subtle nuances and sentiments that might otherwise have been overlooked.

4 | RESULTS

Eleven meaningful topics were identified in reports submitted directly by patients and healthcare professionals and 12 topics in reports

TABLE 1 Topics found in the opioid reports with structural topic modelling

Topics in both MAH and direct reports	Topic proportion MAH/direct reports	Topics in MAH reports only	Topic proportion	Topics in direct reports only	Topic proportion
Problems with patches and skin-related ADRs, especially for buprenorphine and fentanyl; problems described with the patch include skin irritation, itching, the patch coming off and reduced efficacy of the patch	0.186/0.027	Accounts of addiction; related to three oxycodone products	0.13	Patients' accounts of their experiences or problems and their course of treatment with different medications	0.059
Pain experienced by patients using opioids, either through lack of efficacy or painful ADRs	0.08/0.041	Constipation; associated with buprenorphine and oxycodone in cancer patients	0.106	Timing of intake and occurrence of ADRs for tramadol and paracetamol, for example, "This side effect also occurs at night when a tablet is taken for the night"	0.041
ADRs related to the use of tramadol with paracetamol, specifically changes in the international normalized ratio (INR) and drug interactions are mentioned in the MAH reports and a dry mouth and sweating in the direct reports	0.076/0.013	Overdose and death; often in patients with a long medical history sometimes receiving palliative care; overdose sometimes intentional	0.075	Impact of the ADR on patients' lives; topic linked to a standard question on the reporting form asking for this information	0.038
		Dosage adjustment; adjustment of concomitant medications, changes of opioid dosages, drug interactions and drug addictions	0.07	Hospital admissions in patients using oxycodone	0.032
		Respiratory depression often caused by an opioid intoxication, especially for morphine	0.064	Dosage descriptions; descriptions of the exact frequency and dosage of the intake of various medicines	0.03
		Drug exposure during pregnancy; either during delivery but also use during pregnancy. Reports concern the treatment of mothers and the characteristics/well-being of the neonate, often with specific regard to (the risk of) neonatal withdrawal/abstinence symptoms	0.056	Medical history appeared in two topics with detailed descriptions as an answer to a standard question on the ADR reporting form	0.013, 0.010
		Twenty Individual Case Safety Reports (ICSRs); identified in Facebook conversations in a group for chronic pain. Often nothing is known except the side effects (e.g., nausea, addiction) and the suspected drug. All reports link to the case numbers of all other 20 reports, suggesting that this is how the topic was formed	0.055	Nausea; sometimes described in combination with other ADRs such as migraines, dizziness and vomiting. Mostly occurs with tramadol, for pain relief diclofenac was sometimes given	0.019
		Blood and urine screenings; toxicological analyses, urine and blood test/screenings in general and the documentation of laboratory results dominate this topic	0.055		
		Product quality issues for a certain product (adhesion, color, quality)	0.047		

Note: The indicated topic proportions for direct reports do not add up to 1, as topics that are not meaningful in terms of content, for example, those that were formed based on standard text phrases from report submission forms, are not listed.

received via MAHs. Reports relating to “patches and skin”, “pain and various symptoms” and the “combination of tramadol and paracetamol”, for example, were clearly present in both. In contrast, reports addressing “pregnancy and birth” and “overdose and death” featured only in MAH reports, with many more insights into patient experiences in the Lareb reports. Table 1 provides an overview of the topics identified through the different reporting routes.

STM also illuminated the relationship between the prevalence of a topic and specific opioids: for example, reports dealing with respiratory depression have a significantly ($p < 0.001$) higher prevalence for morphine. Among the reports attributed to morphine, the topic “respiratory depression” has a proportion of 18.86% (compared to 6.4% when all reports are considered).

Finally, an important challenge to STM involving this type of data emerged whereby a notable proportion of the data were found to derive from the standard wording of questions in the Lareb reporting system. These recurrent patterns of words were then automatically reproduced in the narratives aggregated in the database, thus posing a hindrance to grasping the underlying content. Future research would be aided through techniques for filtering out such wordings generated by the system rather than those reporting.

5 | CONCLUSIONS

(Structural) Topic Modelling has been used to identify drug safety issues in large quantities of unstructured data such as publications from the literature¹³ and social media.¹⁴ Chen et al.¹⁴ used named entity recognition combined with signal detection and topic modelling for an in-depth analysis, focused on methylphenidate through mining patients' narratives on social media. They found that this approach was able to detect potential signals and to provide a better understanding of patients' behaviors regarding drugs, including misuse. STM, especially in iterative combination with qualitative text analysis,¹⁵ represents an innovative and promising approach to analyzing and utilizing large amounts of unstructured ADR text data in a pharmacovigilance database, in a comprehensive and nuanced way. The topics include information that cannot always be captured in standardized MedDRA[®] terms. The identified topics reflect findings in the literature on opioids, helping validate our findings and the rationale for using STM for the analysis of pharmacovigilance data. Future research using this approach should start with developing more insight into the many standard text formulations in the narratives of direct reports, as these are changed over time. Removing these formulations will aid the validity of the analysis and pertinence of the findings. Given that the identified topics were found to vary in frequency over time, STM could be performed as an addition signal detection screening tool, for instance by looking at new identified topics as these emerge either monthly or yearly.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Florence van Hunsel  <https://orcid.org/0000-0001-8965-3224>

REFERENCES

- Harned M, Sloan P. Safety concerns with long-term opioid use. *Expert Opin Drug Saf.* 2016;15(7):955-962.
- Saucier R, Wolfe D, Dasgupta N. Review of case narratives from fatal overdoses associated with injectable naltrexone for opioid dependence. *Drug Saf.* 2018;41(10):981-988.
- Schifano F, Chiappini S, Corkery JM, Guirguis A. Assessing the 2004-2018 fentanyl misusing issues reported to an international range of adverse reporting systems. *Front Pharmacol.* 2019;10(46). <https://doi.org/10.3389/fphar.2019.00046>
- Dasgupta N, Schnoll SH. Signal detection in post-marketing surveillance for controlled substances. *Drug Alcohol Depend.* 2009;105:S33-S41.
- The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use. Wiley Encyclopedia of Clinical Trials 2021 Accessed 16 June, 2017. <https://www.meddra.org/>
- van Hunsel F, Harmark L, Rolfes L. Fifteen years of patient reporting—what have we learned and where are we heading to? *Expert Opin Drug Saf.* 2019;18(6):477-484.
- Rolfes L, van Hunsel F, Wilkes S, van Grootheest K, van Puijenbroek E. Adverse drug reaction reports of patients and healthcare professionals—differences in reported information. *Pharmacoepidemiol Drug Saf.* 2015;24(2):152-158.
- Harpaz R, Callahan A, Tamang S, et al. Text mining for adverse drug events: the promise, challenges, and state of the art. *Drug Saf.* 2014; 37(10):777-790.
- Roberts ME, Stewart BM, Tingley D, Airolidi EM, eds. The structural topic model and applied social science. *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation.* Vol 4. Neural Information Processing Society; 2013:1-20.
- Norwegian Institute of Public Health. ATC structure and principles 2011. Accessed 1 June, 2017. https://www.whocc.no/atc/structure_and_principles/
- DiMaggio P, Nag M, Blei D. Exploiting affinities between topic modeling and the sociological perspective on culture: application to newspaper coverage of U.S. government arts funding. *Poetics.* 2013;41(6):570-606.
- Blei DM. Probabilistic topic models. *Commun ACM.* 2012;55(4):77-84.
- Zou C. Analyzing research trends on drug safety using topic modeling. *Expert Opin Drug Saf.* 2018;17(6):629-636.
- Chen X, Faviez C, Schuck S, et al. Mining patients' narratives in social media for pharmacovigilance: adverse effects and misuse of methylphenidate. *Front Pharmacol.* 2018;9:541.
- Eickhoff M, Wieneke R, (Eds) Understanding topic models in context: A mixed-methods approach to the meaningful analysis of large document collections AIS eLibrary; 2018.

How to cite this article: Lösche L, Brown P, van Hunsel F.

Using structural topic modelling to reveal patterns in reports on opioid drugs in a pharmacovigilance database.

Pharmacoepidemiol Drug Saf. 2022;31(9):1003-1006. doi:10.1002/pds.5502