Everything Everywhere AI At Once: A conversation about the rising use of artificial intelligence in health research and clinical care.

Chair: Amanda Roxburgh

Chair's email: Amanda.Roxburgh@burnet.edu.au

Discussants:

Community member [to be confirmed] with lived/living experience of drug use.

Clinician researcher Dr David Goodman-Meza, Kirby Institute, Sydney UNSW.

Email: dgoodman@kirby.unsw.edu.au

Authors:
Ting Xia[1], Tina Lam[1], Joanna Dipnall[2,3], Jane Hayman[4], Crystal Leung[1], Bishaal Tej Gurung[1,5], Angus Skeen[1,5], Nazifa Rahman[1], Amy McNeilage[1,6], Richard Beare[7,8], Nadine E Andrew[7,8], Amanda Roxburgh[9], Paul Dietze[9], Suzanne Nielsen[1], Dan Anderson-Luxford[10] Benjamin Riordan[10], Jack Delminico[10], Emmanuel Kuntsche[10]

[1]Monash Addiction Research Centre, Eastern Clinical School, Monash University, Melbourne, Australia, [2]School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia, [3]School of Medicine, Deakin University, Geelong, Australia, [4]Victorian Injury Surveillance Unit, Monash University Accident Research Centre, Melbourne, Australia, [5] School of Medicine, Monash University, Melbourne, Australia, [6] Sydney Medical School, Faculty of Medicine and Health, The University of Sydney, Sydney, New South Wales, Australia, [7]Peninsula Clinical School, School of Translational Medicine, Monash University, Melbourne, Australia, [8]National Centre for Healthy Ageing, Monash University, Melbourne, Australia, [9]Burnet Institute, Melbourne, Australia, [10]Centre for Alcohol Policy Research, La Trobe University

Presenter's emails: ting.xia@monash.edu, Crystal.Leung@monash.edu, B.Riordan@latrobe.edu.au, D.Anderson-Luxford@latrobe.edu.au

Aim: Artificial intelligence (AI), and the sophistication of large language models, has evolved at a rapid rate over the past few years and is increasingly being used in public health research to better capture data on drug-related harms and other important clinical indicators to inform clinical care. There are both benefits and risks to the use of AI within public health research. Benefits include the ability to more accurately capture comorbid conditions and clinical indicators from healthcare data to ensure the best clinical care is provided. Risks include concerns about privacy, increasing the stigma for people who use drugs, and that sometimes AI just gets it wrong. Now is a critical time to have these conversations about how we can best use AI in research in ways that minimise any associated risks while harnessing the value of healthcare data and improving health outcomes. This symposium will provide findings on a few key projects using artificial intelligence (specifically natural language processing (NLP) approaches) to better capture the prevalence of alcohol and drug-related harms from a range of different sources (including electronic health records and

social media platforms). Following the presentations, a community member with lived/living experience [to be confirmed] and clinician researcher Dr David Goodman-Meza will explore the tensions between the benefits and risks of AI use within the AOD sector. We hope to generate an insightful conversation about ways forward with respect to AI use in health research and clinical care.

Disclosure of Interest Statement: This work was supported by the Australian National Health and Medical Research Council (NHMRC, GNT2037997). The authors thank the Department of Health for providing the data and the Victorian Injury Surveillance Unit (VISU) for data extraction. Suzanne Nielsen is the recipient of an NHMRC Investigator Grant (#2025894).

PRESENTATION 1: Machine Learning and Natural Language Approaches for Identifying Psychotropic Medicine Overdose in Electronic Health Records: A Scoping Review

Presenting Authors: Ting Xia

Introduction: Globally, psychotropic drug overdose mortality has increased significantly over the past decade. Existing drug-related harm reporting systems rely on structured coding system, yet much relevant information is recorded in free-text clinic documents, leading to underreporting or misclassification. Machine Learning (ML) and Natural Language Processing (NLP) techniques have the potential to synthesise unstructured free-text clinical data from electronic health records (EHRs) to detect psychotropic drug overdose and enhance surveillance of harms. This study aims to evaluate the extent to which NLP/ML techniques detect drug overdoses from unstructured EHR data, examine methodological approaches and limitations, and identify research gaps to inform future directions.

Methodology: A scoping review was conducted across four databases in March 2026. Studies applying NLP and ML to detect analgesic, psychotropic and substance-related overdose events from EHR data were included. Title and abstract were screened by two independent reviewers, with full text by three. Data extraction was completed by one reviewer and verified by another for quality and consistency. This review summarised the techniques used, model performance metrics and their application to different drug classes.

Results: The search identified 2111 unique records. After title and abstract screening, 31 studies underwent full-text review, with 11 meeting the inclusion criteria. All 11 included studies were conducted in the United States (US), focusing on opioid overdose. Common techniques included support vector machine (SVM), Naïve Bayes classifiers, and XGBoost algorithms. Advanced NLP models such as the Pooled Flair BERT model achieved high performance (PPV of 99.16%, sensitivity of 99.10% and F1-score 99.13%). While unstructured NLP data enhanced overdose detection, comparability was limited by dataset variability, incomplete clinical notes and drug-class.

Conclusion: NLP and ML models show strong potential in detecting drug overdoses from unstructured data, outperforming traditional models. However, incomplete clinical documentation and variability across institutions remain key barriers to broader clinical adoption.

PRESENTATION 2: Beyond Diagnostic Codes: Enhancing Opioid Overdose Identification in Emergency Department Data via Natural Language Processing of Clinical Text

Presenting Authors: Crystal Leung, Ting Xia

Introduction: Emergency departments (EDs) are a critical point of contact for opioid overdose-related presentations. However, reliance on structured coding for surveillance can lead to misclassification or missed cases, as much of the clinically relevant information is captured in unstructured free-text fields. We aimed to develop and validate a pipeline that applies natural language processing (NLP) and Machine Learning (ML) classifiers to identify opioid overdoses in routinely collected ED data, incorporating both structured and unstructured free-text fields.

Methods: An NLP-enhanced ML pipeline was trained and developed on a stratified sample of the Victorian Emergency Minimum Dataset (VEMD) records. Medical concepts were extracted from free text using the Medical Concept Annotation Tool (MedCAT) to generate a Bag-of-Words (BOW) representation. Multiple clinical ontologies, including SNOMED CT and UMLS, were tested for Named Entity Recognition (NER) while concept-filtering parameters and model configurations were adjusted. Subsequently, the BOW representation and structured data were combined to build ML models for binary classification of opioid overdoses. Several model architectures, including Logistic Regression, Random Forest (RF) and XGBoost, were assessed against a manually coded gold standard using standard performance metrics.

Key Findings and Next Steps: The pretrained UMLS model and XGBoost classifier achieved the best performance, with a macro F1 score of 0.772 and a precision of 0.86, outperforming the customised SNOMED-CT model (precision = 0.16, F1 = 0.26) and the existing administrative ICD-10 diagnosis codes (precision 0.12, F1 = 0.20). Currently, the pipeline is limited in classifying short, ambiguous texts, and text with drug brands and ED abbreviations. Further customisation and fine-tuning will be tested by training on customised concepts and context tagging.

Implications on communities, practice, policy: The project provides a foundation for future integration into national monitoring systems and linkage with mortality or toxicology data, enabling improved monitoring and informing evidence-based policy and public health interventions.

PRESENTATION 3: It is on everything, everywhere, all at once: Using AI to estimate how common alcohol is in the modern media environment.

Presenting Authors: Benjamin Riordan, Dan Anderson-Luxford, Samatha Salim

Introduction: Smart devices, streaming services, and ubiquitous internet have provided us with access to more media than ever before. Unfortunately, alcohol references and advertising are extremely common in modern media and can impact alcohol use. With traditional methods like content analysis, it is too burdensome to estimate just how much alcohol people see, however, using AI models for zero-shot learning (where no training is needed) researchers can easily analyse large datasets. We will overview how accurate these models are at detecting alcohol in different media (music, social media, movies) and examples of how we are using these models in our research.

Approach: We tested a range of AI models that can be used for zero-shot learning to identify alcohol in text (social media posts, music) and images (movies, social media posts) and compared these to human annotation to estimate accuracy.

Key Findings: For text, LLMs like Llama3 accurately identified alcohol in social media posts and lyrics (~96%) and researchers can include examples in their prompts to improve accuracy. For images, using older smaller models like CLIP can be accurate when alcohol is front and centre, but more recent models like LaVA are accurate even when alcohol is in the background or blurry. Tutorials can be found on CAPR's github (https://github.com/ltu-capr).

Conclusions and Next Steps: AI models can be applied to extremely large amounts of media data and be used to estimate how common alcohol is and when incorporated with data donations help understand the impact exposure has on alcohol use.

Implications: Recent AI models can be used to identify alcohol in media content and could be used to assist with film classifications, identify advertising, identifying lyrics with explicit content, and included with plugins to block alcohol images online.

PRESENTATION 4: The Alcohol Industry Watchdog: Can Large Language Models Monitor Industry Influence on Policy Making?

Presenting Authors: Dan Anderson-Luxford, Benjamin Riordan, Jack Delminico, Emmanuel Kuntsche

Introduction: Monitoring the alcohol industry's influence on policymaking is vital to understanding how its commercial activities (e.g., lobbying, campaign funding) shape public health regulation. Traditional policy surveillance relies on manual coding limiting scalability. Recent advances in large language models (LLMs) offer promise for automating the identification and classification of alcohol policy discourse. This study assesses LLMs' utility in monitoring alcohol policy through: (a) identifying references, (b) classifying stance across texts, and (c) comparing positions over time and domains.

Approach: Using a few-shot learning, an automated pipeline was developed to detect alcohol references, policy issues, and stance (pro, anti, neutral). The model was applied to two corpora: 1) Australian federal parliamentary transcripts from 1901–2025 and 2) industry

materials including policy submissions, annual reports, and self-regulatory documents. The model's performance was evaluated using precision, recall, and F1 scores. Descriptive trend analyses of the prevalence and characteristics of alcohol policy references were conducted, alongside a comparison of political party and industry agreement on key policy issues.

Results: Preliminary validation revealed an accuracy above 75% across classification tasks, with higher accuracy for policy identification, and lower accuracy on policy stance, especially when ambiguous or vague language was used. Trend analyses are also presented and discussed.

Conclusions: These findings affirm the potential of AI-enabled surveillance to monitor alcohol policies and industry influence effectively. Integrating debates, submissions, and corporate documents within a unified framework offers a scalable, transparent method to trace policy narratives and stakeholder positions.

Implications: Such LLM-based approaches have broad applications for public health policy monitoring, research on the commercial determinants of health, and the analysis of large regulatory texts. Opportunities for integrating this approach into a living tool for ongoing monitoring of alcohol industry influence are discussed.