Identifying Alcohol Use Disorder Phenotypes Using Regular Expressions

Michaela Robinson

Dr. Alvin Jeffery

BEST DS-2 Intern, Fisk University

Department of Biomedical Informatics, Vanderbilt University Medical Center

VANDERBILT UNIVERSITY
MEDICAL CENTER

Alcohol use disorder is common, costly, & problematic

Around **28.8 million adults** ages 18 and older were reported with Alcohol Use Disorder (AUD) in 2021. Additionally, **753,000 youth** (12-17 years) were diagnosed with AUD (NIH, 2020).

It can cause disruption to daily life, relationships, and health. The average cost of alcohol use in the US averages around **$2,000 a year** and more.

Unfortunately, it may go undetected by clinicians who may not recognise the signs of AUD; stigma

https://www.apa.org/topics/substance-use-abuse-addiction/alcohol-disorders

https://landmarkrecovery.com/the-cost-of-alcohol-use-disorder/

Alcohol and Substance Use Disorder Phenotyping: ICD Codes

- [International Statistical Classification of Diseases and Related Health Problems](#) (ICD) is a medical classification list that contains diseases, diagnosis, and symptoms.

- These codes are used to help identify cases like substance use disorders (SUD) which extends to alcohol use disorders (AUD).

- However, ICD codes may not be the most accurate way to detect cases due to complex electronic health record data (E.H.R).

3

Natural Language Processing (NLP) and AUD Identification

- Studies have shown that using natural language processing (NLP) to identify SUD concepts in clinical notes have

  promising results.

- Automation of Addiction Behaviors Checklist for opioid use disorder
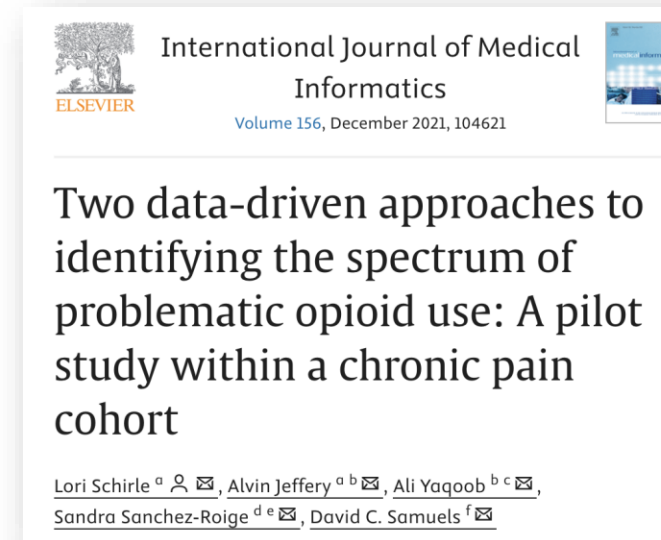
Schirle, L., Jeffery, A., Yaqoob, A., Sanchez-Roige, S., & Samuels, D. C. (2021). Two data-driven approaches to identifying the spectrum of problematic opioid use: A pilot study within a chronic pain cohort. *International Journal of Medical Informatics*, 156, 104621. https://doi.org/10.1016/j.ijmedinf.2021.104621

Chatham, A. H., Bradley, E. D., Schirle, L., Sanchez-Roige, S., Samuels, D. C., & Jeffery, A. D. (2023). Detecting Problematic Opioid Use in the Electronic Health Record: Automation of the Addiction Behaviors Checklist in a Chronic Pain Population. *medRxiv : the preprint server for health sciences*, 2023.06.08.23290894. https://doi.org/10.1101/2023.06.08.23290894

How can we expand and use this knowledge for AUD?

- This project aims to expand the phenotyping of SUD by focusing on AUD within a chronic pain cohort.

- Using natural language processing techniques, the primary objective is to use **regular expressions (regex)** to search clinical notes for key terms aligned with the **Alcohol Use Disorders Identification Test (AUDIT)**.

Why is this important?

Automation and Application:

If automation of AUD phenotyping with regex proves to be accurate and reliable, this method can enhance diagnosis, identification, and prevention within multiple clinical and genetic contexts.

Methods: Chronic Pain Cohort

Total Patients (VUMC Synthetic Derivative)	Sex Demographics	Race Demographics
8,063	Female - 63% Male – 37%	Asian – 1% Black - 16% White - 80.6% Other - 1.5%

- Chronic pain defined by ICD codes (relating to chronic pain) occurring 2 different days in patient E.H.R
- 3,485,348 clinical notes compiled into 68 note files

Chatham, A. H., Bradley, E. D., Schirle, L., Sanchez-Roige, S., Samuels, D. C., & Jeffery, A. D. (2023). Detecting Problematic Opioid Use in the Electronic Health Record: Automation of the Addiction Behaviors Checklist in a Chronic Pain Population. *medRxiv : the preprint server for health sciences*, 2023.06.08.23290894. https://doi.org/10.1101/2023.06.08.23290894

Methods: AUDIT-C

- The Alcohol Use Disorders Identification Test (AUDIT) is a screening tool used to assess drinking behaviors and alcohol consumption.

- The AUDIT-C is a shortened version of this checklist that includes the first three questions of the AUDIT. Our team previously developed regex patterns for the first three questions.

AUDIT-C Questionnaire

Patient Name _____ Date of Visit _____

1. **How often do you have a drink containing alcohol?**
   - ☐ a. Never
   - ☐ b. Monthly or less
   - ☐ c. 2-4 times a month
   - ☐ d. 2-3 times a week
   - ☐ e. 4 or more times a week

2. **How many standard drinks containing alcohol do you have on a typical day?**
   - ☐ a. 1 or 2
   - ☐ b. 3 or 4
   - ☐ c. 5 or 6
   - ☐ d. 7 to 9
   - ☐ e. 10 or more

3. **How often do you have six or more drinks on one occasion?**
   - ☐ a. Never
   - ☐ b. Less than monthly
   - ☐ c. Monthly
   - ☐ d. Weekly
   - ☐ e. Daily or almost daily

Methods:
AUDIT and RegEx

- Of the 10 questions, this project focused on developing questions #4, 8, and 9.

- These questions were converted into **regular expressions (regex)** which are tools that can search for patterns and matches to specific key words and phrases within text.

4. How often during the last year have you found that you were not able to stop drinking once you had started?

(0) Never
(1) Less than monthly
(2) Monthly
(3) Weekly
(4) Daily or almost daily

8. How often during the last year have you been unable to remember what happened the night before because you had been drinking?

(0) Never
(1) Less than monthly
(2) Monthly
(3) Weekly
(4) Daily or almost daily

9. Have you or someone else been injured as a result of your drinking?

(0) No
(2) Yes, but not in the last year
(4) Yes, during the last year

Example for question 8

```
/\b(.?\b(alcohol|etoh|alchol|drinking\salcohol|drinking\sheavily|drinking)\b.*?\b(blackout(s)?|black\sout(s)?|memory\sloss|forg(?:et(s)?|etting|ot|)|brownout)\b.?)\b
```

Text **Tests** NEW   1 match (0.0ms)

John Doe, a 45-year-old male, presents with complaints of significant memory loss following an episode of heavy alcohol consumption. The patient reports that he drank six beers within a few hours during a social event last night. Upon waking this morning, he discovered that he could not recall substantial portions of the previous evening, including conversations and activities that occurred while he was intoxicated. The patient describes this memory loss as a ""blackout,"" where he was conscious and engaging in activities but unable to form lasting memories.

Example for question 9

Results - Chronic Pain Cohort

- After testing regex patterns on 15 of the compiled note files, we found:

  - 2 matches for **question 9**

 - 1 match for **question 8**

- No matches for **question 4**
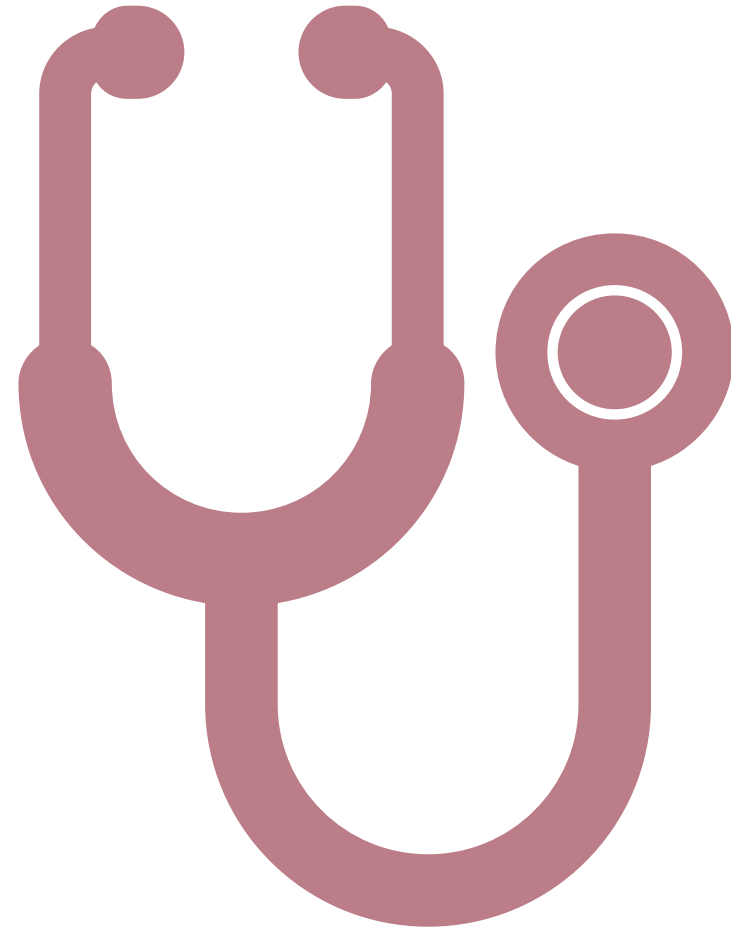
Challenges and Future Directions

Cohort Selection:

The chronic pain cohort for this project was initially created for a study on opioid use disorders.

Future Studies:

Potentially explore using a different cohort, such as one specifically selected for alcohol use disorder or including patients with liver conditions, to enhance generalizability and robustness of findings.
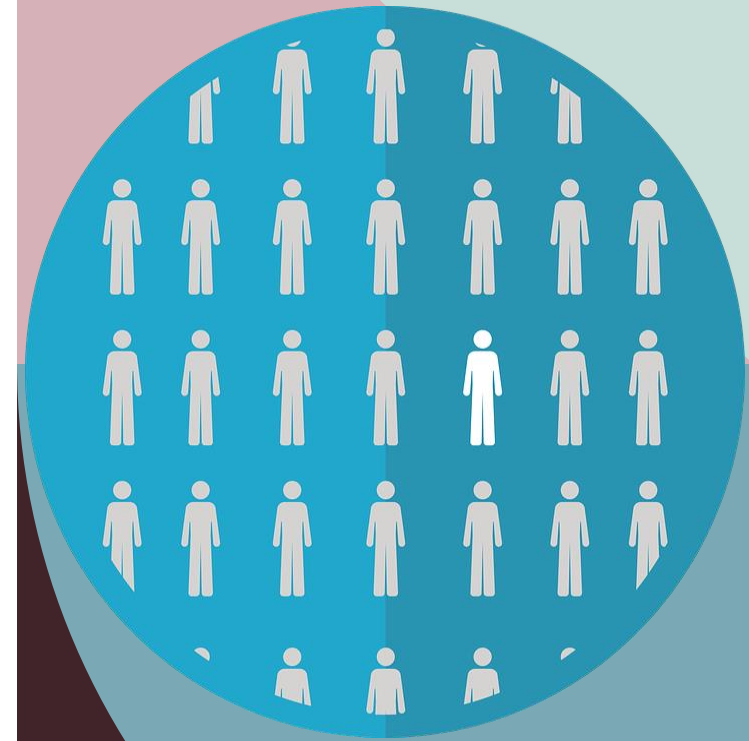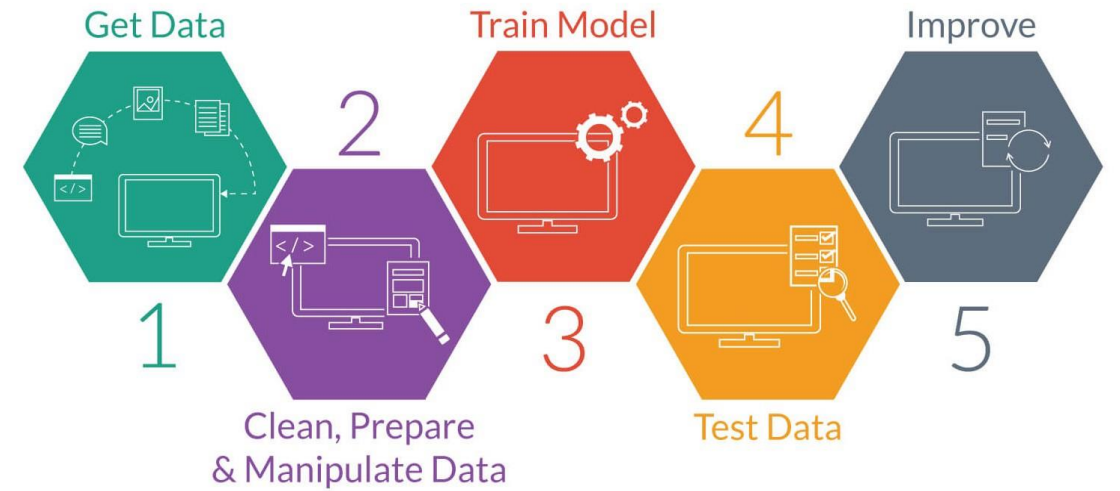
Challenges and Future Directions
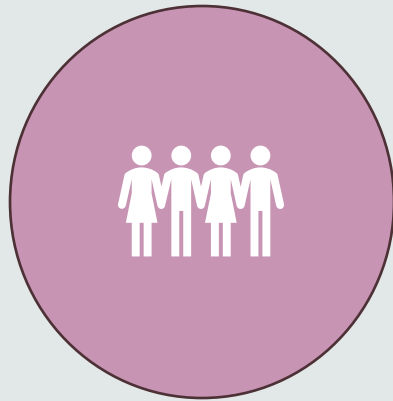

Regular Expressions:

- Searching for simple phrases or words serves as an easy way to identify patterns within text.

- However, searching for more nuanced terms may prove to be a more complex process.
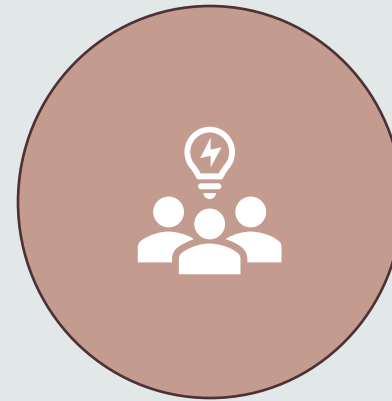
Advanced NLP Techniques:

- Incorporate advanced LLMs to capture the nuances of the AUDIT questions more effectively.

14

Future Steps

New cohort focused on AUD

Further testing and eventual evaluation by experts

Acknowledgements & Funding

Thank you!

Any Questions?

Michaela.robinson@vumc.org
Mprobinson07@my.fisk.edu

AUDIT Checklist

The Alcohol Use Disorders Identification Test: Interview Version

Read questions as written. Record answers carefully. Begin the AUDIT by saying "Now I am going to ask you some questions about your use of alcoholic beverages during this past year." Explain what is meant by "alcoholic beverages" by using local examples of beer, wine, vodka, etc. Code answers in terms of "standard drinks". Place the correct answer number in the box at the right.

1. How often do you have a drink containing alcohol?

   (0) Never [Skip to Qs 9-10]
   (1) Monthly or less
   (2) 2 to 4 times a month
   (3) 2 to 3 times a week
   (4) 4 or more times a week

2. How many drinks containing alcohol do you have on a typical day when you are drinking?

   (0) 1 or 2
   (1) 3 or 4
   (2) 5 or 6
   (3) 7, 8, or 9
   (4) 10 or more

3. How often do you have six or more drinks on one occasion?

   (0) Never
   (1) Less than monthly
   (2) Monthly
   (3) Weekly
   (4) Daily or almost daily

   Skip to Questions 9 and 10 if Total Score for Questions 2 and 3 = 0

4. How often during the last year have you found that you were not able to stop drinking once you had started?

   (0) Never
   (1) Less than monthly
   (2) Monthly
   (3) Weekly
   (4) Daily or almost daily

5. How often during the last year have you failed to do what was normally expected from you because of drinking?

   (0) Never
   (1) Less than monthly
   (2) Monthly
   (3) Weekly
   (4) Daily or almost daily

6. How often during the last year have you needed a first drink in the morning to get yourself going after a heavy drinking session?

   (0) Never
   (1) Less than monthly
   (2) Monthly
   (3) Weekly
   (4) Daily or almost daily

7. How often during the last year have you had a feeling of guilt or remorse after drinking?

   (0) Never
   (1) Less than monthly
   (2) Monthly
   (3) Weekly
   (4) Daily or almost daily

8. How often during the last year have you been unable to remember what happened the night before because you had been drinking?

   (0) Never
   (1) Less than monthly
   (2) Monthly
   (3) Weekly
   (4) Daily or almost daily

9. Have you or someone else been injured as a result of your drinking?

   (0)   No
   (2)   Yes, but not in the last year
   (4)   Yes, during the last year

10. Has a relative or friend or a doctor or another health worker been concerned about your drinking or suggested you cut down?

   (0) No
   (2) Yes, but not in the last year
   (4) Yes, during the last year

Record total of specific items here

If total is greater than recommended cut-off, consult User's Manual.

18