Validation of an Alcohol Misuse Classifier in Hospitalized Patients

Daniel To[1], **Brihat Sharma**[2], **Niranjan Karnik**[3], **Cara Joyce**[1,4,5], **Dmitriy Dligach**[2,4,5], **Majid Afshar**[1,4,5]

[1]Stritch School of Medicine, Loyola University Chicago, Maywood, IL

[2]Department of Computer Science, Loyola University Chicago, Chicago, IL

[3]Department of Psychiatry & Behavioral Sciences, Rush Medical College, Rush University, Chicago, IL

[4]Department of Public Health Sciences, Loyola University Chicago, Maywood, IL

[5]Center for Health Outcomes and Informatics Research, Loyola University Chicago, Maywood, IL

Abstract

BACKGROUND: Current modes of identifying alcohol misuse in hospitalized patients rely on self-report questionnaires and diagnostic codes which have limitations including low sensitivity. Information in the clinical notes of the electronic health record (EHR) may further augment the identification of alcohol misuse. Natural language processing (NLP) with supervised machine learning has been successful at analyzing clinical notes and identifying cases of alcohol misuse in trauma patients.

METHODS: An alcohol misuse NLP classifier, previously developed on trauma patients who completed the Alcohol Use Disorders Identification Test, was validated in a cohort of 1,000 hospitalized patients at a large, tertiary health system between January 1, 2007 and September 1, 2017. The clinical notes were processed using the clinical Text Analysis and Knowledge Extraction System. The National Institute on Alcohol Abuse and Alcoholism (NIAAA) guidelines for alcohol misuse were used during annotation of the medical records in our validation dataset.

RESULTS: The alcohol misuse classifier had an area under the receiver operating characteristic curve of 0.91 (95% CI 0.90–0.93) in the cohort of hospitalized patients. The sensitivity, specificity, positive predictive value, and negative predictive value were 0.88 (95% CI 0.85–0.90), 0.78 (95%

Corresponding Author: Majid Afshar, MD, MSCR, Center for Translation Research and Education, 2160 S First Avenue, Building 115, Room 445, Maywood, IL 60153, USA (majid.afshar@lumc.edu).

HUMAN SUBJECTS PROTECTIONS

The study was performed in compliance with the World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects, and was reviewed by Loyola University Chicago Health Science Division Institutional Review Board

CONFLICT OF INTEREST

No conflict of interest to disclose amongst the authors

CI 0.74–0.82), 0.85 (95% CI 0.82–0.87), and 0.82 (95% CI 0.78–0.86), respectively. The Hosmer-Lemeshow Test (P=0.13) demonstrates good model fit. Additionally, there was a dose-dependent response in alcohol consumption behaviors across increasing strata of predicted probabilities for alcohol misuse.

CONCLUSION—The alcohol misuse NLP classifier had good discrimination and test characteristics in hospitalized patients. An approach using the clinical notes with NLP and supervised machine learning may better identify alcohol misuse cases than conventional methods solely relying on billing diagnostic codes.

BACKGROUND AND SIGNIFICANCE

Between 2006 and 2014, the rate of alcohol-related emergency department (ED) visits increased 47%, with over 40% of these patients requiring hospital admission[1]. Alcohol-related disorders rank second in 7-day unplanned hospital readmissions, greater than heart disease and respiratory failure[2]. Multiple observational studies have demonstrated that the prevalence of alcohol misuse is higher in hospitalized patients than in the general population or outpatient settings[3,4]. However, much of the existing evidence to identify patients with alcohol misuse relies on billing diagnosis codes which have very poor sensitivity[5].

Methods to reliably and efficiently identify patients with alcohol misuse remains a challenge in both clinical surveillance and for research purposes. Manual screening with interviewer-administered questionnaires require hospital staff who are not at point of care with the provider and do not build the same level of trust, which contributes to suboptimal screening and missed opportunities for intervention. Further, building questionnaires into the electronic health record (EHR) and hiring staff to administer them require additional effort and investment. When surveyed, over 90% of patients stated that they would accurately report their alcohol consumption to their provider in order to get the best medical care[6]

Using information from the clinical notes, specifically the documentation of the patient's social and behavioral history during intake, may be more effective than conventional methods of using billing codes. In addition, the notes may contain other details outside explicit mentions about alcohol consumption that may be informative. Clinical notes in the electronic health record (EHR) are a rich source of data, but their unstructured format renders them complex and difficult to analyze. An estimated 80% of the data in an EHR system resides in an unstructured format and much of the behavioral information such as substance use are embedded in the notes[7–9]. Traditional statistical modeling approaches may not be optimal for examining alcohol misuse in the notes and more advanced machine learning methods may produce better classifier from text features. We previously developed an NLP and machine learning algorithm for identifying alcohol misuse[10]; however, the classifier was trained and tested in a cohort of trauma patients. Scaling the tool to all hospitalized patients has not been validated and may serve as an important computable phenotype for epidemiology and surveillance studies.

We aim to validate the test characteristics of our previously published Alcohol NLP classifier in a cohort of hospitalized adult patients at a tertiary health system to further elucidate its generalizability in the hospitalized patient population. We hypothesize that the

classifier will have a sensitivity and specificity above 80% when applied in non-trauma patients.

METHODS

PATIENT SELECTION AND ENVIRONMENT

We tested the alcohol NLP classifier in a case-control approach with 1,000 non-trauma inpatient hospitalizations at Loyola University Medical Center between January 1, 2007 and September 1, 2017. An oversampling of encounters with positive International Classification of Diseases (ICD) codes for alcohol use disorders, testing for blood alcohol concentration (BAC), and orders for the Clinical Institute Withdrawal Assessment (CIWA) was performed to provide a sizable sample of at-risk patients for alcohol misuse. The case-control design was to provide a better validation set for discrimination (AUC ROC) between alcohol misuse and no misuse.

VALIDATION DATA SET OF CASES AND NON-CASES FOR ALCOHOL MISUSE

Cases and non-cases were determined by the annotator through chart review of the entire electronic health record for the 1,000 patients selected for the validation dataset. The annotator received substance use training through Loyola's Institute for Transformative Interprofessional Education and completed online didactics for Screening, Brief Intervention, and Referral to Treatment (SBIRT). Prior to independent review, the annotator met an inter-rater reliability with Cohen's kappa coefficient of >0.75 with an attending critical care physician (MA).

The following criteria were applied to identify cases of alcohol misuse: (1) National Institute of Alcoholism and Alcohol Abuse (NIAAA) quantity limits for alcohol misuse and unhealthy behaviors associated with alcohol consumption[11]. The quantity limits were calculated from clinical documentation of drinking behaviors such as "patient drank a pint of vodka daily"; (2) alcohol-related injuries including emergency department encounters with a blood alcohol concentration (BAC) level 80mg/dL; (3) CIWA 8 at any point during the hospitalization with alcohol-related symptoms[12], (4) physician diagnosis of alcohol misuse. Patients with documented quantity and frequency of alcohol use that did not meet NIAAA or no clear documentation of alcohol misuse were classified as no alcohol misuse.

PROCESSING OF CLINICAL TEXT AND FEATURE EXTRACTION

Linguistic processing of the clinical notes during hospitalization for the machine learning model was performed using the clinical Text Analysis and Knowledge Extraction System (cTAKES; http://ctakes.apache.org)[13]. Named entity mentions (i.e. anatomy, symptoms, disease, procedures, etc.) were identified from the notes and mapped to a concept unique identifier (CUI) from the National Library of Medicine's metathesaurus in the Unified Medical Language System (UMLS) database. For example, the named entity mention for 'drinking problems' was assigned C0085762, and the named entity mention for 'alcohol abuse' from the notes was mapped to a separate CUI than 'history of alcohol abuse', which was C0221628. Each named entity mention was also analyzed for its negation status (i.e. 'no alcohol abuse' or 'without drinking problems'). The CUIs were subsequently normalized

using a term-frequency, inverse document-frequency (TF-IDF) transformation to weigh the CUIs to account for frequently used words across all notes. The weighted CUIs were fed as independent variables into the machine learning model which was tuned to the highest area under the receiver operating characteristic curve (AUC ROC). The final machine learning model was based on the original development paper.[10] The model was a logistic regression with a LASSO (Least Absolute Shrinkage and Selection Operator) regularization.

ANALYSIS WITH SUPERVISED MACHINE LEARNING

The trained CUI-based model used on our validation dataset is accessible at https://github.com/brihat9135/AlcoholNLP_Classifier. In the original study, the model was divided into 80% (n=1137) for training and 20% (n=285) for testing against a reference standard of the Alcohol Use Disorders Identification test, a well-validated reference dataset recommended by the World Health Organization for identifying alcohol misuse using sex-specific cutoffs of 5 and 8 for females and males, respectively. In examining learning curves for our model, we noted peak effects for AUC ROC were achieved once sample size exceeded 1,200. Therefore, the model was retrained on the entire cohort (n= 1,422) of the original trauma patients to provide better power for the machine learning model. The same hyperparameters from the original model were applied with a logistic regression LASSO regularization. The final, trained model selected 25 CUIs out of approximately 10,000 that were predictive of alcohol misuse (Table 1). Our updated model selected more CUIs than the original 16-CUI model; however, the top features were highly correlated with representations for ethanol, drinking, and intoxication. The correlation matrix between the original 16-CUI model trained on 80% of AUDIT data and the updated 25-CUI version from the full 100% of available AUDITs are shown in Supplemental 1.

The model was subsequently validated on our annotated cohort of 1,000 hospitalized non-trauma patients. Discrimination of the prediction models was evaluated using the AUC ROC. Goodness-of-fit was formally assessed by the Hosmer-Lemeshow test and verified visually with a calibration plot. The following test characteristics were examined: sensitivity, specificity, negative predictive value (NPV), and positive predictive value (PPV). Analysis was performed using Python Version 3.6.5 (Python Software Foundation) and RStudio Version 1.1.463 (RStudio Team, Boston, MA). The Institutional Review Board of Loyola University Chicago approved this study.

RESULTS

PATIENT AND DATA CHARACTERISTICS

Patient characteristics and drinking patterns of the hospital encounters are displayed in Table 2. Patients with alcohol misuse had greater frequencies in weekly alcohol consumptions, number of drinks per occasion, number of binge episodes, higher mean level of blood alcohol concentration (BAC), and greater frequency of ICD codes for alcohol use disorders (p<0.01 for all comparisons). During chart review, 14 patients had CIWA scores 8 recorded for non-alcohol related agitation and treatment so they were classified as no alcohol misuse.

PERFORMANCE OF ALCOHOL NLP CLASSIFIER IN HOSPITALIZED PATIENTS

The validation dataset of patients was comprised of 58.5% (n=585) with any level of alcohol misuse. The alcohol NLP classifier had excellent discrimination with an AUC ROC of 0.91 (95% CI 0.90–0.93) (Figure 1). The sensitivity/recall, specificity, PPV/precision, and NPV of the classifier were 0.88 (95% CI 0.85–0.90), 0.78 (95% CI 0.74–0.82), 0.85 (95% CI 0.82–0.87), and 0.82 (95% CI 0.78–0.86), respectively.

The model fit the data well by Hosmer-Lemeshow Test (P=0.13). Recalibration measures did not further improve model fit (Figure 2). Alcohol consumption behaviors across increasing strata of predicted risk for alcohol misuse demonstrated a dose-dependent response in the following categories: (1) quantity of alcohol consumption; (2) frequency of alcohol consumption; (3) proportion with binge drinking; (4) BAC levels (Figure 3). Face validity was also noted with increasing proportions of physician orders for alcohol withdrawal monitoring (Clinical Institute Withdrawal Assessment for Alcohol) and diagnostic codes for alcohol-related conditions with increasing predicted risk for alcohol misuse (Table 3).

APPLICATION OF THE ALCOHOL NLP CLASSIFIER

We applied the alcohol NLP classifier to the entirety of our health system's clinical data warehouse comprised of 229,884 hospitalized adult encounters between January 1, 2007 and September 30, 2017. The data corpus was comprised of over 31 million notes. The NLP Alcohol Classifier identified 8.7% of patient encounters with alcohol misuse compared to 4.9% (p<0.001) from diagnostic codes for alcohol-related diagnoses (ICD-9/10: 291.02–291.9, 303.00–303.93, 305.00–305.03, F10.1x, F10.2x, F10.9x).

DISCUSSION

We previously showed our alcohol NLP classifier carried good discrimination and calibration to identify trauma patients with alcohol misuse[10]. In our adult non-trauma hospitalized cohort, we continue to show good discrimination and calibration to identify patients with alcohol misuse with a dose-dependent response when examined across strata of predicted probabilities. Our computable phenotype for alcohol misuse is a useful approach for epidemiology and surveillance studies that may identify more cases from hospitalizations than conventional methods using billing diagnostic codes.

Clinical NLP has already demonstrated a major positive effect on research and practice[14–16]. Health systems including the Veterans Affairs have shown the benefit of NLP over manual review and used it for identifying depression and self-harm[17], identifying cases of cirrhosis[18], and identification of reportable cancer[19]. Modern NLP has fused with machine learning to better learn from data with the most powerful NLP methods relying on supervised machine learning, taking advantage of reference standards (i.e. AUDIT) to learn and predict for unseen cases. Few studies have examined the role for NLP in building computable phenotypes for behavioral disorders like alcohol misuse.

One study by Wang et al. uses regular expressions to extract specific queries about alcohol use[20]. This approach does not capitalize on the strengths of standardized feature creation

from all text using medical ontologies such as our use of CUIs. A more modern approach with NLP treats each CUI as a feature and requires no domain knowledge. Although using regular expression-based queries may be useful, a data-driven approach using machine learning allows for the discovery of new, predictive features. Our approach used a list of medical concepts as variables that were fed into a model for predicting alcohol misuse. In this approach, discovery was not limited by domain knowledge or expertise, and other entity mentions outside the alcohol domain may prove predictive (i.e. medical history or social/behavioral determinants of health). In addition, our classifier leveraged not only the provider's documentation but also ancillary notes, diagnostics reports, and embedded medication notes. In our list of the CUIs selected during training of the model, concepts such as lorazepam (medication commonly used to treat alcohol withdrawal), sedated state, assault, injury wounds, and fracture (possible physical consequences of alcohol misuse) were predictive for alcohol misuse. Our classifier leveraged not only the provider's documentation but also ancillary notes, diagnostics reports, and embedded medication notes.

Although the NLP classifier was not designed to determine the degree of misuse, we also showed the predicted probabilities from our classifier were correlated to the degree of misuse. A dose dependent response was notable in measurements of alcohol misuse. The positive correlation between BAC order and CIWA order suggested that the classifier was representative of physician behavior since physicians were more likely to order BAC or CIWA in patients suspected for alcohol misuse. In addition, the increasing percentage of CIWA scores greater than 8 further suggested that higher predicted probability correlated to alcohol dependence, which is more commonplace in hospitalized patients than the general population[3]. Our classifier also identified those without a billing diagnosis for alcohol misuse. We highlighted this by showing a nearly two-fold increase in detection of any alcohol misuse when using our alcohol NLP classifier versus structured diagnostic codes.

To date, innovations in health informatics for alcohol misuse have been focused on new data capture tools and applications. Little evidence exists demonstrating the application of NLP and machine learning in EHRs for detection of alcohol misuse. Expansion of our tool into the hospital setting may enable a standardized method to perform surveillance on all patient encounters and is the first step for a more automated and comprehensive approach. By deriving the tool from existing data and using CUIs, which better account for lexical variation and semantic ambiguities between providers and health systems, our approach may have interoperability. Large volume health systems may benefit from an automated and comprehensive algorithm, and our trained model is available for application: https://github.com/brihat9135/AlcoholNLP_Classifier.

There are several limitations present in the study. First, the NLP classifier was validated in a single-center health system and external validation in other health systems is needed. Our sample was enriched for cases and may not be generalizable to cohorts with varying prevalence and may lead to variations in the positive predictive value. False positive may have occurred in our application of the classifier to the entire cohort of hospitalized patients. The lack of a gold standard for our validation testing is another limitation despite our best efforts to use chart reviews for identifying cases and non-cases. Variations in practice for capture of social and behavioral determinants of health in other health systems may exist and

affect the performance of the classifier. The addition of structured data such as nursing flowsheet items, laboratory data, and medication data may improve the net reclassification of our models but we found adequate performance from just using the notes alone. Lastly, processing of the notes into CUIs requires local expertise and may pose a barrier for interoperability across health systems but we have previously developed and applied a large-scale NLP architecture for health systems to benchmark[21].

CONCLUSION

The NLP classifier that was developed on trauma patients has shown to be generalizable to hospitalized patients. Further external validation at other hospitals is necessary to test the generalizability before widely implementing the NLP classifier at other institutions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

REFERENCES

1. White AM, Slater ME, Ng G, Hingson R, Breslow R. Trends in alcohol-related emergency department visits in the united states: Results from the nationwide emergency department sample, 2006 to 2014. Alcohol Clin Exp Res. 2018;42(2):352–359. doi: 10.1111/acer.13559 [doi] [PubMed: 29293274]

2. Fingar KR, Barrett ML, Jiang JH. Comparison of all-cause 7-day and 30-day readmissions, 2014. the HCUP report : Healthcare cost and utilization project (HCUP): Statistical briefs;2017 ASI 4186–20.230;statistical brief no. 230. . 2017 https://statistical.proquest.com/statistica1insight/result/pqpresultpage.previewtitle?docType=PQSI&titleUri=/content/2017/4186-20.230.xml. Accessed 2/23/19

3. Roson B, Monte R, Gamallo R, et al. Prevalence and routine assessment of unhealthy alcohol use in hospitalized patients. Eur J Intern Med. 2010;21(5):458–464. doi: 10.1016/j.ejim.2010.04.006 [doi] [PubMed: 20816605]

4. Doering-Silveira J, Fidalgo TM, Nascimento CL, et al. Assessing alcohol dependence in hospitalized patients. Int J Environ Res Public Health. 2014; 11(6):5783–5791. doi: 10.3390/ijerph110605783 [doi] [PubMed: 24879488]

5. Lid TG, Eide GE, Dalen I, Meland E. Can routine information from electronic patient records predict a future diagnosis of alcohol use disorder? Scand J Prim Health Care. 2016;34(3):215–223. https://www.ncbi.nlm.nih.gov/pubmed/27404326 [PubMed: 27404326]

6. Smothers BA, Yahr HT. Alcohol use disorder and illicit drug use in admissions to general hospitals in the united states. Am J Addict. 2005;14(3):256–267. doi: N43378242V397407 [pii] [PubMed: 16019976]

7. Ford E, Carroll JA, Smith HE, Scott D, Cassell JA. Extracting information from the text of electronic medical records to improve case detection: A systematic review. J Am Med Inform Assoc. 2016;23(5): 1007–1015. doi: 10.1093/jamia/ocv180 [doi] [PubMed: 26911811]

8. Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF. Extracting information from textual documents in the electronic health record: A review of recent research. Yearb Med Inform. 2008:128–144. doi: me08010128 [pii] [PubMed: 18660887]

9. Shivade C, Raghavan P, Fosler-Lussier E, et al. A review of approaches to identifying patient phenotype cohorts using electronic health records. J Am Med Inform Assoc. 2014;21(2):221–230. doi: 10.1136/amiajnl-2013-001935 [doi] [PubMed: 24201027]

10. Afshar M, Phillips A, Karnik N, et al. Natural language processing and machine learning to identify alcohol misuse from the electronic health record in trauma patients: Development and internal validation. J Am Med Inform Assoc. 2019;26(3):254–261. doi: 10.1093/jamia/ocy166 [doi]. [PubMed: 30602031]

11. National Insitute of Alcohol Abuse and Alcoholism. Drinking levels defined. https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/moderate-binge-drinking. Accessed July, 2019

12. Sullivan JT, Sykora K, Schneiderman J, Naranjo CA, Sellers EM. Assessment of alcohol withdrawal: The revised clinical institute withdrawal assessment for alcohol scale (CIWA-ar). Br J Addict. 1989;84(11): 1353–1357 [PubMed: 2597811]

13. Savova GK, Masanz JJ, Ogren PV, et al. Mayo clinical text analysis and knowledge extraction system (cTAKES): Architecture, component evaluation and applications. J Am Med Inform Assoc. 2010;17(5):507–513. doi: 10.1136/jamia.2009.001560 [doi] [PubMed: 20819853]

14. Jones BE, South BR, Shao Y, et al. Development and validation of a natural language processing tool to identify patients treated for pneumonia across VA emergency departments. Appl Clin Inform. 2018;9(1): 122–128. doi: 10.1055/s-0038-1626725 [doi] [PubMed: 29466818]

15. Castro VM, Dligach D, Finan S, et al. Large-scale identification of patients with cerebral aneurysms using natural language processing. Neurology. 2017;88(2): 164–168. doi: 10.1212/WNL.0000000000003490 [doi] [PubMed: 27927935]

16. Carrell DS, Cronkite D, Palmer RE, et al. Using natural language processing to identify problem usage of prescription opioids. Int J Med Inform. 2015;84(12): 1057–1064. doi: 10.1016/j.ijmedinf.2015.09.002 [doi] [PubMed: 26456569]

17. Haerian K, Salmasian H, Friedman C. Methods for identifying suicide or suicidal ideation in EHRs. AMIA Annu Symp Proc. 2012;2012:1244–1253 [PubMed: 23304402]

18. Chang EK, Yu CY, Clarke R, et al. Defining a patient population with cirrhosis: An automated algorithm with natural language processing. J Clin Gastroenterol. 2016;50(10):889–894. doi: 10.1097/MCG.0000000000000583 [doi] [PubMed: 27348317]

19. Osborne JD, Wyatt M, Westfall AO, Willig J, Bethard S, Gordon G. Efficient identification of nationally mandated reportable cancer cases using natural language processing and machine learning. J Am Med Inform Assoc. 2016;23(6): 1077–1084. doi: 10.1093/jamia/ocw006 [doi] [PubMed: 27026618]

20. Wang Y, Chen ES, Pakhomov S, et al. Automated extraction of substance use information from clinical texts. AMIA Annu Symp Proc. 2015;2015:2121–2130 [PubMed: 26958312]

21. Afshar M, Dligach D, Sharma B, et al. Development and application of a high throughput natural language processing architecture to convert all clinical documents in a clinical data warehouse into standardized medical vocabularies. J Am Med Inform Assoc. 2019. doi: ocz068 [pii]

- Clinical notes were mapped to standardized terminology in the Metathesaurus

- NLP classifier performed well with hospitalized patients

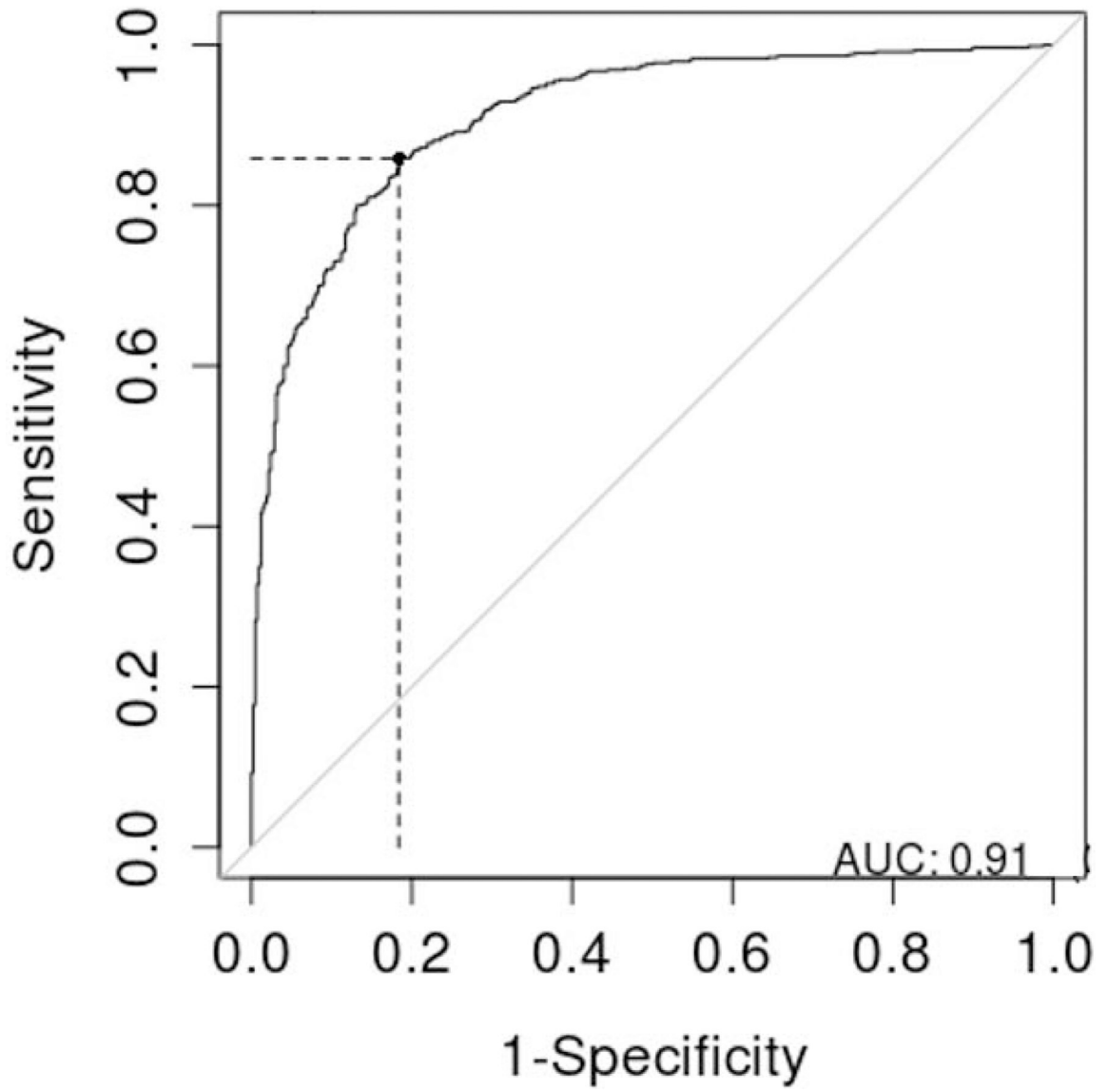- Alcohol use behavior is correlated with increasing probabilities for alcohol misuse
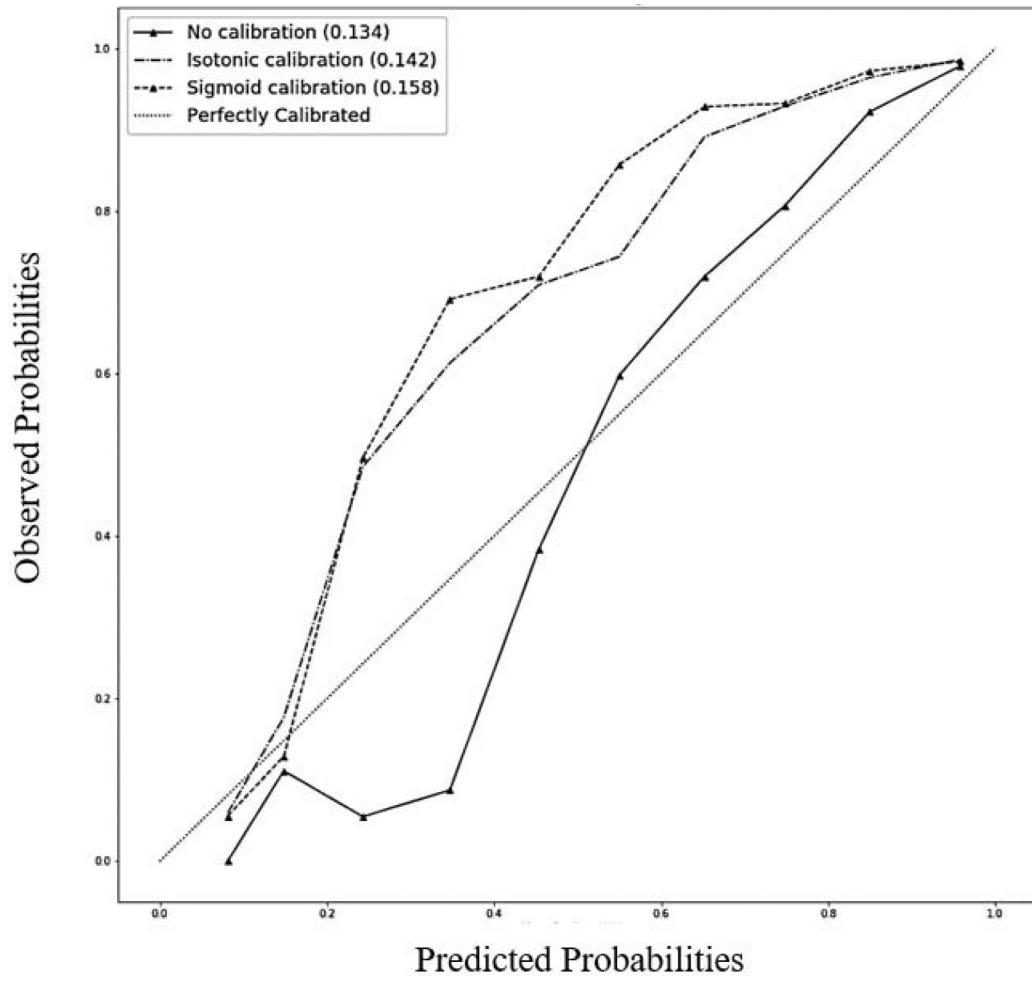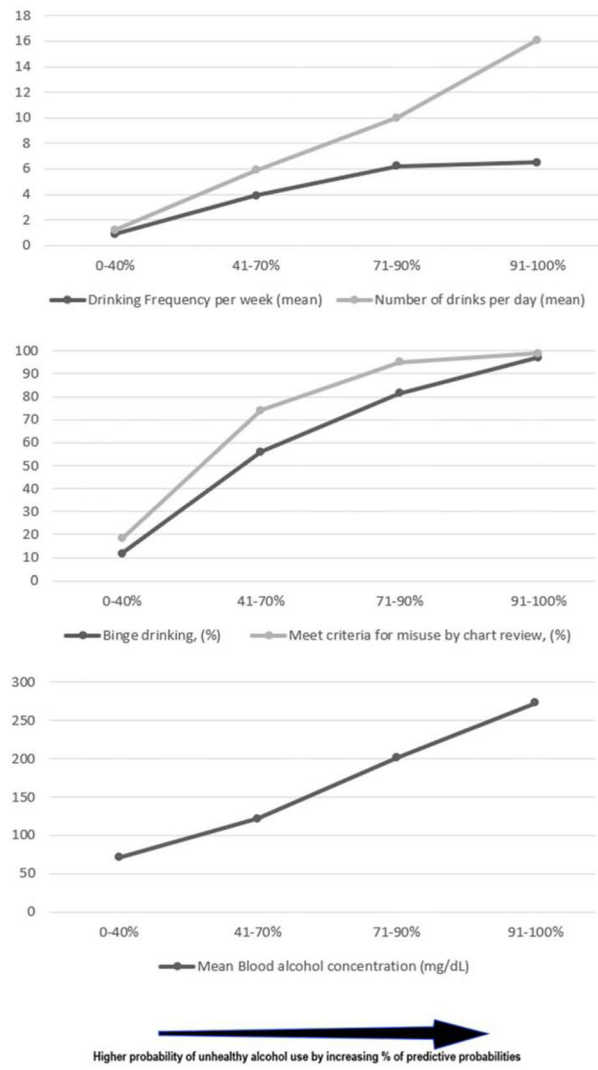
Figure 1.

Figure 2.

Figure 3.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

25 Concept unique identifiers selected during training of machine learning model to identify cases of alcohol misuse

Positive CUI features (β coefficients from logistic regression classifier	c0001962	Ethanol, 22.05
	c0024002	Lorazepam, 5.89
	c0562381	Victim of abuse finding, 3.59
	c0001973	Alcoholic Intoxication, Chronic, 2.70
	c0149531	Fracture of pelvis, 2.33
	c0004063	Assault, 1.81
	c1273870	Management procedure, 1.76
	c0034606	Radionuclide Imaging, 1.67
	c0043250	Injury wounds, 1.37
	c1272883	Injection, 0.98
	c1299583	Independently able, 0.91
	c0003086	Ankle, 0.42
	c0034929	Reflex action, 0.37
	c3263723	Traumatic injury, 0.34
	c0235195	Sedated state, 0.012
	c0016658	Fracture, 0.0024
Negative CUI features (β coefficients from logistic regression classifier	c0024687	Mandible, −0.054
	c0277814	Sitting position, −0.22
	c0558145	Skin appearance normal (finding), −0.27
	c0039225	Tablet Dosage Form, −0.30
	c0030193	Pain, −0.51
	c0231683	Gait normal, −0.68
	c1513302	Mild Adverse Event, −1.09
	c1292890	Procedure on hip, −2.50
	c0020538	Hypertensive disease, −5.89

Table 2.

Demographics and Alcohol Consumption Behaviors (Data collected during chart review of 1,000 hospitalized patients for the validation dataset)

Characteristic	Total (n=1000)	Misuse (n=585)	No Misuse (n=415)	P-value
Age, mean (SD)	49 (13.6)	47 (13.6)	51 (13.4)	<0.001
Race (White), n (%)	578 (57.8)	338 (57.8)	240 (57.8)	0.99
Drinking frequency per week, mean (SD) (n=509)	3.4 (3.3)	6.3 (1.8)	0.4 (1.3)	<0.001
Number of drinks per day, mean (SD) (n=519)	5.9 (8.7)	11.12 (9.54)	0.32 (0.83)	<0.001
Binge [*], n (%)	475 (47.5)	475 (81.2)	0 (0)	<0.001
BAC [**] order, n (%)	526 (52.6)	470 (80.3)	56 (13.5)	<0.001
BAC [**], mean (SD)	169 (144)	188 (141)	11.7 (21.4)	<0.001
CIWA [***] order, n (%)	296 (29.6)	266 (45.5)	30 (7.2)	<0.001
CIWA [***] 8, n (%)	208 (70.3)	194 (73.0)	14 (46.7)	0.006
History of misuse but not current, n (%)	121 (12.1)	11 (19)	110 (26.5)	<0.001
ICD for alcohol use disorder [****], n (%)	302 (30.2)	265 (42.3)	37 (8.9)	<0.001

[*] Binge = BAC > 0.08g/dL which is about 4 drinks for women and 5 drinks for men in 2 hours[11].

[**] BAC: Blood alcohol concentration

[***] CIWA: Clinical Institute Withdrawal Assessment for Alcohol

[****] ICD: International Classification of Diseases (ICD-9/10: 291.0–291.9, 303.00–303.93, 305.00–305.03, F10.1x, F10.2x, F10.9x)

Table 3.

Demographics and Alcohol Consumption Behaviors Stratified by Predicted Probability (Data collected during chart review of 1,000 hospitalized patients for the validation dataset)

Characteristic	0–40th Percentile N=399	41–70th Percentile N=300	71–90th Percentile N=200	91–100th Percentile N=100	P-value
Predicted Probability of Alcohol Misuse	0.499 – 0.075	0.844 – 0.500	0.981 – 0.845	0.999 – 0.982	
Age, median (IQR)	51 (41–60)	50 (38–59.3)	50 (42–58)	44 (36.8–54)	0.004
Race (White), n (%)	226 (56.6)	171 (57)	117 (58.5)	63 (63)	0.70
Drinking frequency per week, mean (SD) (n=509)	0.9 (2.0)	3.9 (3.3)	6.2 (2.1)	6.5 (1.5)	<0.001
Number of drinks per day, mean (SD) (n=519)	1.2 (2.8)	5.9 (8.8)	10.0 (8.8)	16.1 (10.4)	<0.001
Binge *, n (%)	47 (11.8)	168 (56)	163 (81.5)	97 (97)	<0.001
BAC ** order, n (%)	68 (17)	188 (62.7)	174 (87)	96 (96)	<0.001
BAC **, mean (SD)	71.4 (84.8)	122 (115)	201(146)	273 (143)	<0.001
CIWA *** order, n (%)	26 (6.5)	115 (38.3)	91 (45.5)	54 (54)	<0.001
CIWA *** 8, n (%)	14 (53.8)	74 (64.3)	76 (83.5)	44 (84.6)	0.021
History of misuse but not current, n (%)	59 (14.8)	55 (18.3)	7 (3.5)	0 (0)	<0.001
ICD for alcohol misuse ****, n (%)	31 (7.8)	79 (26.3)	124 (62)	68 (68)	<0.001
Annotator Classification of Misuse, n (%)	73 (18.3)	222 (74)	190 (95)	99 (99)	<0.001

*
Binge = BAC > 0.08g/dL which is about 4 drinks for women and 5 drinks for men in 2 hours[11].

**
BAC: Blood alcohol concentration

***
CIWA: Clinical Institute Withdrawal Assessment for Alcohol

****
ICD: International Classification of Diseases (ICD-9/10: 291.0–291.9, 303.00–303.93, 305.00–305.03, F10.1x, F10.2x, F10.9x)