User Feedback on the Use of a Natural Language Processing Application to Screen for Suicide Risk in the Emergency Department

James L. Pease, PhD
Devyn Thompson, MSW
Jennifer Wright-Berryman, PhD
Marci Campbell, BA

Abstract

Suicide is the 10th leading cause of death in the USA and globally. Despite decades of research, the ability to predict who will die by suicide is still no better than 50%. Traditional screening instruments have helped identify risk factors for suicide, but they have not provided accurate predictive power for reducing death rates. Over the past decade, natural language processing (NLP), a form of machine learning (ML), has been used to identify suicide risk by analyzing language data. Recent work has demonstrated the successful integration of a suicide risk screening interview to collect language data for NLP analysis from patients in two emergency departments (ED) of a large healthcare system. Results indicated that ML/NLP models performed well identifying patients that came to the ED for suicide risk. However, little is known about the clinician's perspective of how a qualitative brief interview suicide risk screening tool to collect language data for NLP integrates into an ED workflow. This report highlights the feedback and observations of patient experiences obtained from clinicians using brief suicide screening interviews. The investigator used an open-ended, narrative interview approach to inquire about the qualitative interview process. Three overarching

Address correspondence to James L. Pease, School of Social Work, College of Allied Health Sciences, University of Cincinnati, Cincinnati, OH, USA. peasejs@ucmail.uc.edu.

Devyn Thompson, School of Social Work, College of Allied Health Sciences, University of Cincinnati, Cincinnati, OH, USA. peasejs@ucmail.uc.edu.

Jennifer Wright-Berryman, School of Social Work, College of Allied Health Sciences, University of Cincinnati, Cincinnati, OH, USA. peasejs@ucmail.uc.edu.

Marci Campbell, Clarigent Health, Mason, OH, USA.

themes were identified: behavioral health workflow, clinical implications of interview probes, and integration of an application into provider patient experience. Results suggest a brief, qualitative interview method was feasible, person-centered, and useful as a suicide risk detection approach.

Introduction

Suicide is the 10th leading cause of death in the USA and globally. Over 47,000 people nationally[1] and 700,000 worldwide[2] died by suicide in the last year. Despite decades of research, rates continue to rise as the ability to predict who will die by suicide remains elusive. A recent meta-analysis suggests that current methods of predicting risk for suicide death are no better than 50% or random chance.[3] Theories abound about why people die by suicide — from sociological[4] to biological[5] to psychological.[6–10] These theories aid in understanding risk factors for suicide but have not delivered adequate predictive models for reducing death rates, nor have they resulted in screening instruments that have adequate predictive value.[3,11]

Suicide risk predictive models are more complex than can be measured using traditional instruments with a constrained number of risk factors. To accommodate this complexity, future suicide screening will require more sophisticated modeling techniques, accomplished with machine learning (ML). In the last decade, ML has been identified as the next necessary step in modeling complex risk factors,[3] and different ML methods have been used to develop better screening techniques.[3,12]

Natural language processing (NLP), a form of machine learning, is a method used to identify suicide risk by analyzing language data.[12] Studies analyzing language data collected from existing medical records or assessment transcripts have been successful in improving the prediction of future suicide attempts over traditional clinical care.[13,14] Pioneering researchers in the field[15–17] developed the first NLP technology to accurately identify suicide risk from a corpus of suicide notes, then expanded to collecting speech data using real-time patient interviews in an emergency department. Obtaining a language sample in a clinical setting requires a brief interview; therefore, it must be integrated into the clinical workflow. Recent work has demonstrated the successful integration of a suicide risk screening interview to collect language data for NLP analysis from patients in two emergency departments (ED) of a large healthcare system.[18] Results from this study suggested that ML/NLP models performed well in identifying patients that came to the ED for suicide risk in an area of the country where speech dialects vary from language samples used in the original development of the technology.[16] However, little is known about the clinician's perspective of how a qualitative brief interview suicide risk screening tool to collect language data for NLP integrates into an ED workflow.

This report highlights user feedback and observations of patient experiences obtained from clinicians following a study using brief suicide screening interviews. This type of interview differs from standard care, which typically employs short, self-report, standardized scales to screen for suicide risk. Due to the differences in time and attention involved in administering the different screening tools, it is critical to understand if a qualitative screener can be successfully implemented into an ED workflow and if there are barriers or benefits to a qualitative suicide risk screening process.

Method

Data collection

This study was nested in a larger study conducted in two emergency departments in a major healthcare system in the Southeastern USA. Patients were randomly assigned to either a treatment or control group ($N = 70$) based on results from a standardized suicide risk screener. Probes developed in previous research[16] about hopes, secrets, anger, fear, and emotional pain (called MHSAFE) were used by clinical staff to elicit emotional responses for NLP analysis in a 4–10-min interview. The patient consent and probes were gathered through an application installed on computer tablets, and the tablets were used to record the entire patient interview. After the study was completed, three clinicians who administered the MHSAFE probes (suicidal patients only, $N = 37$) were interviewed to further assess feasibility, user experience, and perceptions of patient experience with the screening process. All three clinicians were masters-level trained mental health staff, with one being a team supervisor. The interviews lasted approximately 1 h each. The interviews were conducted and recorded using Zoom (a virtual interface), and audio files were transcribed using an online transcription service.

The investigator used an open-ended, narrative interview approach[19] prompting all three participants with similar questions; however, the participant directed the content and flow of the response. The investigator began with a single question based on the topic area, such as behavioral health workflow, use of the MHSAFE probes in the ED environment, and benefits and barriers of the qualitative screening technique. The participants responded without any interruption, so they had control over what information they shared. The investigator then asked follow-up questions triggered by the participant's response.

Data analysis

Verbatim transcripts were independently hand-coded by two investigators. Braun and Clarke's[19] inductive, thematic approach was used for the following steps of data analysis: (1) transcribe the recordings, (2) generate initial codes, (3) search for initial themes, (4) review themes against the data, (5) define and name themes, and (6) produce the report. Through iterative theme-data checking and discussion, the investigators reached consensus on three overarching themes and subthemes. Microsoft Excel was used to organize data, codes, and themes.

Results

Overarching theme 1: Behavioral health workflow

The behavioral health clinicians were asked to describe their usual emergency department workflow to understand routine suicide risk assessment and triage process before implementing the MHSAFE probes. Overlapping content and themes from the team members were used to create a summary. According to the behavioral health clinician participants:

Upon arrival to the ED, all patients are first taken through the registration process. After registration and being medically cleared, patients are then given an abbreviated version of the Columbia-Suicide Severity Rating Scale (C-SSRS),[20] in which a high score triggers a consult from the behavioral health department. Once the consult request is received, an assigned staff member reviews the patient's medical history, lab work, and any notes regarding the patient's current mental state.

Whether the patient is COVID-19 positive or in an outlying emergency room determines if they are seen through telehealth or in-person.

During the behavioral health consultation, the professional inquires into the patient's symptoms and administers the full version of the C-SSRS. In addition to the suicide screening, the behavioral health professional conducts a brief psychosocial assessment. Following approval from the patient, the behavioral health professional will also seek out any potential collateral information from family members and friends. They then report their care recommendations to the attending doctor, who makes the final decision for care. The possible outcomes of the assessment include the possibility of discharge, voluntary commitment, or involuntary commitment to the psychiatric unit.

Overarching theme 2: Clinical implications of MHSAFE probes

Subtheme 1: Qualitative suicide screening elicits more clinically relevant information than the C-SSRS   The three clinicians reported that using the MHSAFE probes helped the patients be more "forthcoming" about their emotions and mental health status. They all suggested that more information was obtained via use of the probes than standard screening. One clinician stated, "the probes are 'interesting', not what patients are used to being asked." Another clinician reported, "the probes provided space for patients to open up more about their life and who they are." Finally, another clinician said the patients were more likely to "pour their hearts out" and "shared way more than expected."

Subtheme 2: Person-centered suicide screening   All three participants reported that the use of open-ended probes to perform a suicide screening provided a more person-centered method than traditional standardized scales. One participant mentioned the patient "felt like what they were saying was important; felt seen and heard." Another clinician mentioned that the "probes are not as 'accusatory' as standardized scales may feel to patients, and the probes provide validation of what patients are experiencing and feeling." Additionally, one clinician stated, "the probes allowed for both the patient and the interviewer to slow down and pace themselves" and the other reported, "the probes emphasize the patient's point of view and honor the patient's 'core feelings'". Other examples of patient centered screening include observations that there was "space for patients to open up more about their life and who they are" and "share more than they expected".

Subtheme 3: Emotional pain probe triggers vulnerability   Each of the clinicians acknowledged that the final probe about emotional pain was often difficult for the patient to discuss. The clinicians reported that when asking about emotional pain, they observed the patients' vulnerabilities the most. One clinician mentioned that this probe "brought up strong emotions" and another stated that they needed to "find a way to de-escalate patients" after asking this probe.

Overarching theme 3: Integration of app into provider-patient experience

Subtheme 1: Use of app   The clinicians reported the app loaded onto the tablet was "intuitive and user-friendly," although they also mentioned there were a few times that users had difficulty "signing in and had to sign in twice." One clinician stated using the app "did not feel cumbersome or time-consuming" and that it was "smoothly integrated into the clinical process."

Subtheme 2: MHSAFE probes as a non-traditional suicide screening method   The MHSAFE probes do not ask directly about suicide (do not use the word "suicide" or "suicidal"). The clinicians stated that this method is helpful because it may identify folks who come in with needs not related to

suicide but are, in fact, suicidal. One clinician stated that the app could help identify patients "in the gray area, or those who may be malingering." Another clinician asserted that this method of screening could "help catch signs of suicide that the professional is not already detecting."

Discussion

Screening for suicide risk has become a standard of care,[21] and the pace of the environment requires a brief and user-friendly tool. This study explored the feedback of mental health clinicians who participated in a larger research study investigating the use of a brief, qualitative interview to collect language data for NLP analysis to identify suicide risk. The feasibility of using such an approach is critical to understand given the time constraints of the ED workflow.

Currently, emergency departments typically use standardized instruments to identify suicide risk. These instruments conserve time, however, may not produce information concerning the drivers of a patient's suicide risk. Additionally, recent literature suggests these instruments may have inadequate predictive power in identifying who will later die by suicide.[22,23] Therefore, additional approaches should be developed and tested to expand the options for medical and mental health clinics to improve risk response decision-making.

A qualitative screening technique emulates a brief assessment and will provide more information for clinical decision-making. Feedback suggests that the brief interview used in this study was not only acceptable to clinicians, but also feasible when integrated into the ED workflow. Lastly, the clinicians reported that more useful information was gleaned during the interview process, which may increase the value of the time spent over a standardized instrument.

Some limitations of this study are noted. Feedback from the patient population would be useful in understanding the experience of this method of risk screening. It would also be helpful to glean insights into the differences between the use of a standardized questionnaire and a qualitative approach from the patient's perspective. In addition, this study describes interviews with three behavioral health clinicians, which limits the generalizability of the findings. A larger number of user interviews could provide more breadth and depth of feedback. The study took place in one setting, and feedback from users in a variety of settings could also provide helpful information.

Implications for Behavioral Health

User feedback obtained by interviewers on the use of a brief qualitative interview to assess suicide risk yielded some encouraging results in terms of its potential usability in the ED. Three overarching themes emerged: behavioral health workflow, clinical implications of MHSAFE probes, and integration of app into provider-patient experience.

Regarding the workflow, clinicians reported that the application was feasible for integration in the emergency department. This feedback is important to consider, as most emergency room physicians use behavioral health consultants to inform them on outcomes of suicide screening, and the screen should fit into the pathway of care.[24] Additionally, questions loom about the effectiveness of employing standard suicide screeners in emergency departments.[25] If a change is made to a qualitative suicide screen interview approach, it is critical that it can be easily integrated.

The feedback from providers about use of MHSAFE probes suggests the prompts are more conducive to sharing one's emotions and being more forthcoming with information than standardized screeners. This could elicit important clinically relevant information not obtained from closed-ended, dichotomous, survey questions. In addition, reports of the probes providing opening up and sharing about their life highlight the differences between this open-ended interview style and

traditional methods of conducting standardized scales, which are often self-report and ask directly about suicide. In a study of patients who did not disclose suicidal ideation to their healthcare provider and went on to make a suicide attempt within the next 60 days, listening and open expressions of caring were identified as qualities that would allow patients to be more forthcoming about their suicidal ideation.[26] In addition, the feedback on increased emotional vulnerability suggests the probes are getting at a novel and possibly unique theme in suicide risk assessment. The emotional pain probe from the interview brought about "strong emotions" and is consistent with the Suicide Crisis Inventory (SCI), a recent suicide prediction measure with strong psychometrics, which has an emotional pain subscale.[27] Additionally, Klonsky and May identified emotional pain as a key contributor to suicide risk in their Three-Step Theory (3ST) of suicide.[28] Other themes suggest the use of an application is intuitive and does not distract from the interview process. It is timely, given the normality of the use of technology in medicine in the wake of the pandemic, to include technology in suicide prevention efforts.[29]

Conclusion

Although an emergency department relies on quick and concise diagnostic processes, feedback from mental health clinicians using a brief qualitative screener and tablet technology is feasible for this setting. Future directions will include expanding the use of MHSAFE to other settings and different populations and obtaining patient-level feedback for continuous improvement of the patient experience.

Data Availability The data collected were not part of a research study using human subjects. Contact author for data availability.

Declarations

Ethical Approval This manuscript was not part of a research study using human subjects.

Guarantor: James Pease (JP) is the guarantor of this manuscript.

Conflict of Interest The authors declare no competing interests.

References

1. Stone DM, Jones CM, Mack KA. Changes in suicide rates — United States, 2018–2019. *MMWR Morbidity and Mortality Weekly Report* 2021; 70(8):261–268. Available at https://doi.org/10.15585/mmwr.mm7008a1. Accessed 19 December, 2022.
2. *Suicide*. Geneva, SW: World Health Organization, 2021. Available at https://www.who.int/news-room/fact-sheets/detail/suicide (2021). Accessed 16 December, 2022.
3. Franklin JC, Ribeiro JD, Fox KR, et al. Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychological Bulletin* 2017; 143(2):187–232. Available at https://doi.org/10.1037/bul0000084. Accessed 19 December, 2022.
4. Durkheim E. *Suicide: A Study in Sociology*. London, UK: Taylor and Francis, 2005. Available at https://doi.org/10.4324/9780203994 320. Accessed 19 December, 2022.
5. Oquendo MA, Sullivan GM, Sudol K, et al. Toward a biosignature for suicide. *The American Journal of Psychiatry* 2014; 171(12):1259–1277. Available at https://doi.org/10.1176/appi.ajp.2014.14020194. Accessed 19 December, 2022.
6. Baumeister RF. Suicide as escape from self. *Psychological Review* 1990; 97(1):90–113. Available at https://doi-org.uc.idm.oclc.org/10.1037/0033-295X.97.1.90. Accessed 19 December, 2022.

7. Beck AT, Steer RA, Kovacs M, et al. Hopelessness and eventual suicide: a 10-year prospective study of patients hospitalized with suicidal ideation. *The American journal of psychiatry* 1985; 142: 559–563. Available at https://doi:https://doi.org/10.1176/ajp.142.5.559. Accessed 19 December, 2022.

8. Joiner TE. *Why People Die by Suicide*. Cambridge, MA: First Harvard University Press, 2005.

9. Van Orden KA, Witte TK, Cukrowicz KC, et al. The interpersonal theory of suicide. *Psychological Review* 2010; 117(2):575–600. Available at https://doi-org.uc.idm.oclc.org/10.1037/a0018697. Accessed 19 December, 2022.

10. Shneidman ES. *Suicide as Psychache: A clinical approach to self-destructive behavior*. Jason Aronson, https://psycnet.apa.org/record/1993-98267-000 (1993). Accessed 19 December, 2022.

11. Linthicum KP, Schafer KM, Ribeiro JD. Machine learning in suicide science: Applications and ethics. *Behavioral Sciences and the Law* 2019; 37:214–222. Available at https://doi.org/10.1002/bsl.2392. Accessed 19 December, 2022.

12. Bernert RA, Hilberg AM, Melia R, et al. Artificial intelligence and suicide prevention: A systematic review of machine learning investigations. *International Journal of Environmental Research and Public Health* 2020; 17(16): 1–25. Available at https://doi.org/10.3390/ijerph17165929. Accessed 19 December, 2022.

13. Barak-Corren Y, Castro VM, Javitt S, et al. Predicting suicidal behavior from longitudinal electronic health records. *American Journal of Psychiatry* 2017; 174(2): 154–162. Available at https://doi-org.uc.idm.oclc.org/10.1176/appi.ajp.2016.16010077. Accessed 19 December, 2022.

14. Walsh CG, Ribeiro JD, Franklin JC. Predicting risk of suicide attempts over time through machine learning. *Clinical Psychological Science* 2017; 5(3): 457–469. Available at https://doi-org.uc.idm.oclc.org/10.1177/2167702617691560. Accessed 19 December, 2022.

15. Pestian J, Nasrallah H, Matykiewicz P, et al. Suicide note classification using natural language processing: A content analysis. *Biomedical Informatics Insights* 2010; 3: BII.S4706. Available at https://doi.org/10.4137/BII.S4706. Accessed 19 December, 2022.

16. Pestian JP, Grupp-Phelan J, Bretonnel Cohen K, et al. A Controlled Trial Using Natural Language Processing to Examine the Language of Suicidal Adolescents in the Emergency Department. *Suicide and Life-Threatening Behavior* 2016; 46(2):154–159. Available at https://doi.org/10.1111/sltb.12180. Accessed 19 December, 2022.

17. Pestian JP, Sorter M, Connolly B, et al. A machine learning approach to identifying the thought markers of suicidal subjects: A prospective multicenter trial. *Suicide and Life-Threatening Behavior* 2017; 47(1):112–121. Available at https://doi.org/10.1111/sltb.12312. Accessed 19 December, 2022.

18. Cohen J, Wright-Berryman J, Rohlfs L, et al. Integration and validation of a natural language processing machine learning suicide risk prediction model based on open-ended interview language in the emergency department. *Frontiers of Digital Health* 2022; 4:818705. Available at https://doi.org/10.3389/fdgth.2022.818705. Accessed 16 December, 2022.

19. Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006; 3(2):77–101. Available at https://doi.org/10.1191/1478088706qp063oa. Accessed 19 December, 2022.

20. Posner K, Brown GK, Stanley B, et al. The Columbia-suicide severity rating scale: Initial validity and internal consistency findings from three multisite studies with adolescents and adults. *American Journal of Psychiatry* 2011; 168(12):1266–1277. Available at https://doi-org.uc.idm.oclc.org/10.1176/appi.ajp.2011.10111704. Accessed 19 December, 2022.

21. Heyland M, Delaney KR, Shattell M. Steps to achieve universal suicide screening in emergency departments: a call to action. *Journal of psychosocial nursing and mental health services* 2018; 56(10):21–26. Available at https://doi.org/10.3928/02793695-20180503-03. Accessed 19 December, 2022.

22. Chung TH, Hanley K, Le YC, et al. A validation study of PHQ-9 suicide item with the Columbia Suicide Severity Rating Scale in outpatients with mood disorders at National Network of Depression Centers. *Journal of affective disorders* 2023;320(1):590–594. Available at https://doi.org/10.1016/j.jad.2022.09.131. Accessed 19 December, 2022.

23. Simpson SA, Goans C, Loh R, et al. Suicidal ideation is insensitive to suicide risk after emergency department discharge: Performance characteristics of the Columbia-Suicide Severity Rating Scale Screener. *Academic Emergency Medicine* 2021; 28(6):621–629. Available at https://doi-org.uc.idm.oclc.org/10.1111/acem.14198. Accessed 19 December, 2022.

24. Betz ME, Wintersteen M, Boudreaux ED, et al. Reducing suicide risk: Challenges and opportunities in the emergency department. *Annals of Emergency Medicine* 2016; 68(6): 758–765. Available at https://doi.org/10.1016/j.annemergmed.2016.05.030. Accessed 19 December, 2022.

25. Simpson SA, Goans C, Loh R, et al. Suicidal ideation is insensitive to suicide risk after emergency department discharge: Performance characteristics of the Columbia-Suicide Severity Rating Scale Screener. *Academic Emergency Medicine*. Epub ahead of print 2020. Available at https://doi.org/10.1111/acem.14198. Accessed 12 December, 2022.

26. Richards JE, Whiteside U, Ludman EJ, et al. Understanding why patients may not report suicidal ideation at a health care visit prior to a suicide attempt: A qualitative study. *Psychiatric Services* 2018; 70(1):40–45. Available at https://doi-org.uc.idm.oclc.org/10.1176/appi.ps.201800342. Accessed 19 December, 2022.

27. Galynker I, Yaseen ZS, Cohen A, et al. Prediction of suicidal behavior in high-risk psychiatric patients using an assessment of acute suicidal state: The suicide crisis inventory. *Depression and Anxiety* 2017; 34(2):147–158. Available at https://doi.org/10.1002/da.22559. Accessed 19 December, 2022.

28. Klonsky ED, May AM. The Three Step Theory (3ST): A new theory of suicide rooted in "Ideation-to-Action" framework. *International Journal of Cognitive Psychotherapy*. 2015; 8:114–129. Available at https://doi.org/10.1521/ijct.2015.8.2.114. Accessed 19 December, 2022.

29. Larsen ME, Cummins N, Boonstra TW, et al. The use of technology in suicide prevention. *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC* 2015; 2015-November: 7316–7319. Available at https://doi.org/10.1109/EMBC.2015.7320081. Accessed 19 December, 2022.