

Natural language processing markers for psychosis and other psychiatric disorders: emerging themes and research agenda from a cross-linguistic workshop

Hugo Corona Hernández¹, Cheryl Corcoran², Amélie M Achim³, Janna N de Boer^{1,4}, Tessel Boerma⁵, Sanne G Brederoo^{1,6}, Guillermo A Cecchi⁷, Silvia Ciampelli¹, Brita Elvevåg⁸, Riccardo Fusaroli⁹⁻¹¹, Silvia Giordano¹², Mathias Hauglid¹³, Arjan van Hessen^{5,14}, Wolfram Hinzen¹⁵, Philipp Homan¹⁶, Sybren F de Kloet¹⁷, Gina R Kuperberg¹⁸⁻²⁰, Kritika Maheshwari^{21,22}, Natalia B Mota^{23,24}, Alberto Parola^{9,10}, Roberta Rocca¹⁰, Sanne Schuite-Koops¹, Iris E C Sommer^{1,6}, Khiet Truong¹⁴, Alban E Voppel¹, Marieke van Vugt²⁵, Frank Wijnen⁵, and Lena Palaniyappan^{26,27*}

¹University of Groningen, University Medical Center Groningen, Department of Biomedical Sciences of Cells & Systems, Groningen, the Netherlands; ²Icahn School of Medicine at Mount Sinai, Department of Psychiatry, New York, United States of America; ³Université Laval, Département de Psychiatrie et Neurosciences, VITAM Centre de Recherche en Santé Durable, Cervo Brain Research Centre, Québec, Canada; ⁴Utrecht University, University Medical Center Utrecht, Department of Psychiatry, Utrecht, the Netherlands; ⁵Utrecht University, Department of Languages, Literature and Communication, Institute for Language Sciences, Utrecht, the Netherlands; ⁶University Medical Center Groningen, University Center of Psychiatry, Groningen, the Netherlands; ⁷IBM T.J. Watson Research Center, Yorktown Heights, New York, United States of America; ⁸University of Tromsø - the Arctic University of Norway, Department of Clinical Medicine, Tromsø, Norway; ⁹Aarhus University, Department of Linguistics, Cognitive Science and Semiotics, School of Communication and Culture, Aarhus, Denmark; ¹⁰Aarhus University, Department of Culture, Interacting Minds Center, Cognition and Computation Communication, School of Culture and Society, Aarhus, Denmark; ¹¹University of Pennsylvania, Linguistic Data Consortium, Pennsylvania, United States of America; ¹²University of Applied Sciences and Arts of Southern Switzerland, Department of Innovative Technologies, Lugano, Switzerland; ¹³University of Tromsø - the Arctic University of Norway, Faculty of Law, Tromsø, Norway; ¹⁴University of Twente, Faculty of Electrical Engineering, Mathematics and Computer Science, Twente, the Netherlands; ¹⁵Institució Catalana de Recerca i Estudis Avançats and Universitat Pompeu Fabra, Barcelona, Spain; ¹⁶Psychiatric Hospital of the University of Zurich, Department of Psychiatry, Psychotherapy, and Psychosomatics, Zurich, Switzerland; ¹⁷Neurocast BV, Amsterdam, the Netherlands; ¹⁸Tufts University, Department of Psychology, Medford, Massachusetts, United States of America; ¹⁹Harvard Medical School, Department of Psychiatry, Boston, Massachusetts, United States of America; ²⁰Harvard Medical School, Massachusetts General Hospital, Boston, Massachusetts, United States of America; ²¹University Medical Centre Groningen, Department of Genetics, Groningen, the Netherlands; ²²Delft University of Technology, Ethics & Philosophy of Technology Section, Faculty of Technology, Policy, and Management, Delft, the Netherlands; ²³Federal University of Rio de Janeiro, Institute of Psychiatry, Rio de Janeiro, Brazil; ²⁴Research department at Motrix Lab – Motrix, Rio de Janeiro, Brazil; ²⁵University of Groningen, Bernoulli Institute of Mathematics, Computer Science & Artificial Intelligence, Groningen, the Netherlands; ²⁶McGill University, Douglas Mental Health University Institute, Department of Psychiatry, Montreal, Quebec, Canada; ²⁷Western University, Robarts Research Institute, Schulich School of Medicine and Dentistry, London, Ontario, Canada.

*To whom correspondence should be addressed: Douglas Mental Health University Institute, 6875, Boul. LaSalle. Verdun (Quebec) H4H 1R3. Email: lena.palaniyappan@mcgill.ca

Introduction

A multidisciplinary workshop entitled “Crosslinguistic speech patterns: biosocial markers of psychiatric disorders” was held with the support of a Distinguished Lorentz Fellowship granted to Iris Sommer, in conjunction with the DISCOURSE in Psychosis Consortium (October 31st-November 4th 2022, Leiden University, the Netherlands). We (the attendees) included clinical

practitioners and experts in diverse scientific disciplines, such as artificial intelligence (AI), clinical psychology, cognitive neurosciences, computational sciences, ethics, law, linguistics, psychiatry, and technology industry. A main aim of the workshop was to deliberate on potential challenges with respect to the discovery, characterization, validation, and potential utilization of natural language processing (NLP) markers for psychosis and other psychiatric disorders using computational technologies, with the ultimate goal of implementing them ethically in clinical settings. Related to this, we discussed who the main stakeholders key to this enterprise are, including individuals with lived experience, their families, the clinicians who serve them, research scientists with diverse areas of expertise, and ethicists. Ethical issues were discussed in detail, emphasizing their relationship to regulatory concerns that may differ by country and by stakeholder status.

NLP markers for psychiatric disorders

Definition and potential roles

Aligning with a broad characterization of markers in digital medicine¹, we agreed that an NLP marker is a digitally acquired, computationally derived, quantifiable measure or set of measures of human language production reflecting the state of biological, neurocognitive, and social processes that contribute to it. While acknowledging the breadth of oral and sign language-related processes (i.e., production and comprehension of spoken/sign/written language), we mostly focused on speech production for a few key reasons. In psychiatric practice, spoken language is considered to be indicative of mental states, which are reflected in its meaning (i.e., semantic content), form (i.e., grammar), and acoustic features. Metrics of spoken language can easily be derived from audio recordings obtained during routine clinical practice in psychiatry, as well as in naturalistic, ecologically valid contexts (e.g., at home). While many developing markers are obtained using NLP techniques (e.g., cosine semantic similarity metrics), markers derived using other computational approaches focused on human communication processes (e.g., acoustics of speech signal and non-verbal behaviors such as facial expression) are also included in the broad definition of NLP markers.

We recognized that NLP markers might have a *descriptive* role useful for screening, stratification in trials, and as a marker of outcome (e.g., prediction of relapse). In parallel, NLP markers might also have a *mechanistic* role, making them indicative of underlying pathological

mechanisms at cellular, physiological and/or circuit-based levels, which could lead to target engagement for the development of new therapeutics, and plausibly improve prediction accuracy, stratification and monitoring of treatment response.

NLP markers for clinical actions

A set of potential clinical actions and goals were nominated for the use of NLP markers in psychiatry (see Table 1), based on discussions of examples and existing avenues of research. These comprise mostly *descriptive* NLP markers that as yet are limited in accuracy, carrying the risk of both false positives and false negatives. It was agreed that much work needs to be done before any of these use cases could be implemented in the clinic, and that ethical issues, commensurate with other fields of *neurotechnology* that prioritize people’s neurorights^{2,3}, are paramount in developing NLP markers for psychiatric disorders.

The group agreed that the field as yet lacks comprehensive large-scale “candidate-selection” studies for several clinical decisions (e.g., treatment response monitoring and prediction of aggression/violence). We reviewed the promising proof-of-concept studies that support the construct validity of candidate NLP markers that correlate with standard clinical ratings (e.g., associations between cosine similarity metrics and tangentiality⁴ in individuals at clinical risk for psychosis) and that are predictive of some outcomes of interest, such as transition to psychosis from risk states⁵. Robust external replications⁵, prospective validations, cross-linguistic comparisons, and reliability estimates on assay performance are also needed, and clinical trials on integrating NLP markers with routine practice are yet to begin.

The measurement and evaluation of NLP markers for specific clinical actions⁶ can be guided by a principled approach with three steps⁷. First, current clinical knowledge, prior research results, and data-driven approaches should guide the selection of promising features to validate NLP markers for specific clinical actions. Second, optimal procedures for measuring those features should be defined. Third, arguments both in favor and against making changes in current clinical practice related to the employment of NLP markers should be thoroughly examined, addressing issues of validity, reliability, utility, acceptability, and costs.

Understanding the constraints of NLP markers on generalization (e.g., heterogeneity and inherent volunteer bias in training data) is crucial, requiring debiasing strategies during acquisition, training and validation stages and safeguards during implementation. There was general agreement

on the need to collect large diverse samples to determine how NLP markers generalize over populations varying in age, sex, ethnicity, and education, for instance. Constraints on implementation of NLP markers must be considered right from the start in developing predictive models for clinical use. Data-sharing obstacles should be tackled⁸ so that interested parties can collaborate inter-institutionally⁹ to advance the field.

NLP markers and mechanistic research

Significant progress has been made in understanding the neural basis of language processing¹⁰ and its interaction with neurocognitive processes such as attention¹¹ or memory¹². Spoken language conveys information about impairments in thought and cognition in psychiatric disorders¹³. Thus, the mechanisms that underlie NLP markers might be in close proximity to the etiology of psychosis and other psychiatric disorders¹⁴. To test this, there is a need for carefully designed hypothesis-driven experiments in clinical samples. By developing causal-mechanistic explanations for promising NLP markers^{15,16} (i.e., delineating the neural mechanisms that account for their characteristics), in the near future NLP markers could be used as proxy outcomes reflecting whether clinical interventions exert an effect on the underlying mechanisms of a given disorder.

Attendees highlighted that language production is the result of genetic¹⁷ and developmental¹⁸ processes. Further, while an individual's anatomical¹⁹ and cognitive²⁰ characteristics constrain its features, language production is influenced by pharmacological²¹, contextual²², and socio-demographic²³ factors. Therefore, we considered that, with respect to mechanistic investigations of candidate NLP markers, we must improve the consistency of how we acquire, preprocess, and analyze speech data, how we parse effect(s) of potential confounders on the characteristics of candidate NLP markers, and how we interpret candidate NLP markers to ensure robust replications. We acknowledged that candidate NLP markers could map onto multilevel biosocial causal frameworks, and group-aggregated results of NLP markers might be used as priors to inform any personalized care²⁴. Rigorous and large-scale clinical studies evaluating predictive models alongside experimental mechanistic studies should allow us to identify explainable candidate NLP markers.

Imagining a Clinical Decision Support System incorporating NLP markers

Discussions of a putative CDSS incorporating NLP markers highlighted that candidate markers must be validated with “ground truth” clinical rating scales, and evidence that they have real-life functional correlates should be provided. We also agreed that NLP markers must be integrated with other sources of clinical information²⁴, and that training related to their acquisition and interpretation should have minimal burden on clinicians. Further, along with accessibility to and acceptance of candidate markers by clinicians and patients²⁵, any CDSS incorporating NLP markers should achieve expected standards of transparency, trust, and efficient and safe functioning²⁶ for regulatory approvals before widespread clinical use²⁷. In the absence of a formal CDSS, clinical settings can implement NLP markers in pilot testing using human-in-the-loop iterative methodologies²⁸ to begin to flesh out these issues.

Ethical challenges

We anticipate the implementation of any CDSS incorporating NLP markers to face a series of ethical challenges (many of which have been debated for decades). Spoken language reflects psychological states and is considered to be personal data, raising nuanced concerns about data protection and privacy legislation²⁹. The use of audio and video recordings require us to adhere to a set of ethical principles to “preserve people’s privacy, identity, agency and equality”². Likewise, (inter)national AI-laws³⁰ should regulate the process of scaling up any putative CDSS incorporating NLP markers for routine use. Moreover, broader concerns over AI explainability, clinical reasoning, and patients’ autonomy also persist³¹. Specifically, unease about misuse (e.g., discrimination) or potential harms (e.g., missing a relapse event) arising from mistakes in utilizing NLP markers is widespread. In this context, NLP markers must also be first validated and assessed for accuracy, reliability, acceptability, scalability, utility and cost before any consideration can be made for making them an integral part of clinical care. All these ethical issues must be addressed in an explicit and transparent manner. Importantly, previous efforts have suggested that these challenges are surmountable (e.g., the European MONARCA project³²), but call for an interdisciplinary action plan.

Conclusions and future directions

Psychiatric practice is deeply rooted in human language and the communicative interchanges it allows. With unprecedented developments in digital health technology and NLP, we are now at the cusp of systematically building on language-related data to derive clinical benefits. Our consortium will work to build an alliance of lived-experience experts, clinicians, and caregivers in further collaborative work. Constructing benchmark transdiagnostic datasets requires sustained global multicenter collaborations. Researchers in the language sciences could inform the development of cross-linguistic NLP markers that incorporate phenomena of linguistic variation, thus increasing generalizability and avoiding the bias of underrepresenting certain languages or communities of speakers. Empirical cognitive neuroscience and psycholinguistic studies investigating the mechanistic basis of NLP markers can enhance their use in experimental medicine and treatment discoveries. The results could inspire novel linguistic remediations and speech and language therapy approach in psychiatry. A partnership of computational and data scientists with end-users (i.e., clinicians and patients) will enable the implementation of informed modelling pipelines fitting the needs of clinical use. Along with stakeholders in the health technology industry, we will work to improve the accessibility to and acceptability of acquisition and analytics procedures. The success of a safe and responsible use of any CDSS incorporating NLP markers requires support and guidance from ethicists, policy and legal experts, and regulatory bodies. With a commitment to act on these points, a diverse, inclusive, interdisciplinary and global collective for mental-health NLP markers can create the conditions to optimize health care with readily accessible and widely acceptable technology.

Funding

This paper was enabled by a Distinguished Lorentz Fellowship granted to Iris Sommer in 2022 for Computational Linguistics to aid Diagnosis and Treatment Monitoring in Psychiatry. HCH was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexican Government, scholarship number 739604). LP received research support from the Tanna Schulich Chair of Neuroscience and Mental Health (Schulich School of Medicine, Western University: 2019-2022); Monique H. Bourgeois Chair in Developmental Disorders and Graham Boeckh Foundation (Douglas Research Centre, McGill University) and salary award from the Fonds de recherche du Québec–Santé (FRQS). Part of the networking efforts of the DISCOURSE in Psychosis consortium

is funded by a grant from the Tannenbaum Open Science Initiative at the Neuro, McGill University. AP is supported by a Marie Skłodowska-Curie Actions (MSCA) – Individuals Fellowship H2020-MSCA-IF-2018 (Grant agreement ID: 832518, Project: MOVES).

Authorship contribution

HCH, IECS, and LP wrote the first draft of the manuscript. HCH, CC, BE, IECS, and LP developed the structure and arguments of the manuscript based on the contributions made by all of the online- and onsite-attendees of the workshop. All authors read, critically revised, and approved the final manuscript. AMA, GAC and GRK were not able to attend the workshop.

Disclosures

LP reports personal fees from the Canadian Medical Association Journals for serving as chief editor, speaker/consultant fee from Janssen Canada and Otsuka Canada, SPM Course Limited, UK, Canadian Psychiatric Association; book royalties from Oxford University Press; investigator-initiated educational grants from Janssen Canada, Sunovion and Otsuka Canada outside the submitted work. RF reports past consultant fees from F. Hoffmann-La Roche. NBM works at Motrix, an EduTech startup., and has been a consultant to Boehringer Ingelheim.

References

1. Vasudevan S, Saha A, Tarver ME, Patel B. Digital biomarkers: Convergence of digital health technologies and biomarkers. *Npj Digit Med.* 2022;5. doi:10.1038/s41746-022-00583-z
2. Yuste R, Goering S, Arcas BAY, et al. Four ethical priorities for neurotechnologies and AI. *Nature.* 2017;551(7679):159-163. doi:10.1038/551159a
3. Goering S, Klein E, Specker Sullivan L, et al. Recommendations for Responsible Development and Application of Neurotechnologies. *Neuroethics.* 2021;14:365-386. doi:10.1007/s12152-021-09468-6
4. Bilgrami ZR, Sarac C, Srivastava A, et al. Construct validity for computational linguistic metrics in individuals at clinical risk for psychosis: Associations with clinical ratings. *Schizophr Res.* 2022;245:90-96. doi:10.1016/j.schres.2022.01.019
5. Corcoran CM, Carrillo F, Fernández-Slezak D, et al. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry.* 2018;17(1):67-75.
6. Holmlund TB, Chandler C, Foltz PW, et al. Towards a temporospatial framework for measurements of disorganization in speech using semantic vectors. *Schizophr Res.* Published online 2022. doi:10.1016/j.schres.2022.09.020

7. Foltz PW, Chandler C, Diaz-Asper C, et al. Reflections on the nature of measurement in language-based automated assessments of patients' mental state and cognitive function. *Schizophr Res*. Published online 2022. doi:10.1016/j.schres.2022.07.011
8. Palaniyappan L, Alonso-Sanchez MF, MacWhinney B. Is Collaborative Open Science Possible With Speech Data in Psychiatric Disorders? *Schizophr Bull*. 2022;48(5):963-966. doi:10.1093/schbul/sbac058
9. Kairouz P, McMahan HB, Avent B, et al. Advances and open problems in federated learning. *Found Trends® Mach Learn*. 2021;14(1-2):1-210. doi:http://dx.doi.org/10.1561/22000000083
10. Hagoort P, Beckmann CF. Key issues and future directions: The neural architecture for language. In: Hagoort P, ed. *Human Language: From Genes and Brains to Behavior*. MIT Press; 2019:527-532.
11. Goller F, Choi S, Hong U, Ansorge U. Whereof one cannot speak: How language and capture of visual attention interact. *Cognition*. 2020;194:104023. doi:10.1016/j.cognition.2019.104023
12. Shain C, Blank IA, Fedorenko E, Gibson E, Schuler W. Robust Effects of Working Memory Demand during Naturalistic Language Comprehension in Language-Selective Cortex. *J Neurosci*. 2022;42(39):7412-7430. doi:10.1523/JNEUROSCI.1894-21.2022
13. de Boer JN, Brederoo SG, Voppel AE, Sommer IEC. Anomalies in language as a biomarker for schizophrenia: *Curr Opin Psychiatry*. 2020;33(3):212-218.
14. Uher R, Zwickler A. Etiology in psychiatry: embracing the reality of poly-gene-environmental causation of mental illness. *World Psychiatry*. 2017;16(2):121-129. doi:10.1002/wps.20436
15. Parola A, Simonsen A, Bliksted V, Fusaroli R. Voice patterns in schizophrenia: A systematic review and Bayesian meta-analysis. *Schizophr Res*. 2020;216:24-40. doi:10.1016/j.schres.2019.11.031
16. Fusaroli R, Grossman R, Bilenberg N, Cantio C, Jepsen JRM, Weed E. Toward a cumulative science of vocal markers of autism: A cross-linguistic meta-analysis-based investigation of acoustic markers in American and Danish autistic children. *Autism Res Off J Int Soc Autism Res*. 2022;15(4):653-664. doi:10.1002/aur.2661
17. Mekki Y, Guillemot V, Lemaître H, et al. The genetic architecture of language functional connectivity. *NeuroImage*. 2022;249. doi:10.1016/j.neuroimage.2021.118795
18. Rudolph JM, Leonard LB. Early Language Milestones and Specific Language Impairment. *J Early Interv*. 2016;38(1):41-58. doi:10.1177/1053815116633861
19. Dediu D, Janssen R, Moisik SR. Weak biases emerging from vocal tract anatomy shape the repeated transmission of vowels. *Nat Hum Behav*. 2019;3(10):1107-1115. doi:10.1038/s41562-019-0663-x
20. Shafto MA, Tyler LK. Language in the aging brain: The network dynamics of cognitive decline and preservation. *Science*. 2014;346(6209):583-587.
21. Fusaroli M, Simonsen A, Borrie SA, et al. Identifying medications underlying communication atypicalities in psychotic and affective disorders: A pharmacosurveillance study within the FDA Adverse Event Reporting System. Published online medRxiv 2022:2022.09.05.22279609. doi:10.1101/2022.09.05.22279609
22. Nölle J, Fusaroli R, Mills GJ, Tylén K. Language as shaped by the environment: linguistic construal in a collaborative spatial task. *Palgrave Commun*. 2020;6. doi:10.1057/s41599-020-0404-9

23. Palaniyappan L. More than a biomarker: could language be a biosocial marker of psychosis? *Npj Schizophr.* 2021;7. doi:10.1038/s41537-021-00172-1
24. Barron DS, Baker JT, Budde KS, et al. Decision Models and Technology Can Help Psychiatry Develop Biomarkers. *Front Psychiatry.* 2021;12:706655. doi:10.3389/fpsy.2021.706655
25. Brederoo SG, Nadema FG, Goedhart FG, et al. Implementation of automatic speech analysis for early detection of psychiatric symptoms: What do patients want? *J Psychiatr Res.* 2021;142:299-301. doi:10.1016/j.jpsychires.2021.08.019
26. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *Npj Digit Med.* 2020;3. doi:10.1038/s41746-020-0221-y
27. Koutsouleris N, Hauser TU, Skvortsova V, Choudhury MD. From promise to practice: towards the realisation of AI-informed mental health care. *Lancet Digit Health.* 2022;4(11):e829-e840. doi:10.1016/S2589-7500(22)00153-4
28. Chandler C, Foltz PW, Elvevåg B. Improving the Applicability of AI for Psychiatric Applications through Human-in-the-loop Methodologies. *Schizophr Bull.* 2022;48(5):949-957. doi:10.1093/schbul/sbac038
29. United Nations Conference on Trade and Development. Data Protection and Privacy Legislation Worldwide. Published December 14, 2021. <https://unctad.org/page/data-protection-and-privacy-legislation-worldwide>
30. Hauglid MK. What's That Noise? Interpreting Algorithmic Interpretation of Human Speech as a Legal and Ethical Challenge. *Schizophr Bull.* 2022;48(5):960-962. doi:10.1093/schbul/sbac008
31. Keeling G, Nyrup R. Explainable Machine Learning, Patient Autonomy, and Clinical Reasoning. In: Véliz C, ed. *The Oxford Handbook of Digital Ethics.* <https://doi.org/10.1093/oxfordhb/9780198857815.013.27>
32. Puiatti A, Mudda S, Giordano S, Mayora O. Smartphone-centred wearable sensors network for monitoring patients with bipolar disorder. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* ; 2011:3644-3647. doi:10.1109/IEMBS.2011.6090613
33. Huang KY, Wu CH, Su MH, Kuo YT. Detecting Unipolar and Bipolar Depressive Disorders from Elicited Speech Responses Using Latent Affective Structure Model. *IEEE Trans Affect Comput.* 2020;11(3):393-404. doi:10.1109/TAFFC.2018.2803178
34. Lewis CC, Boyd M, Puspitasari A, et al. Implementing Measurement-Based Care in Behavioral Health: A Review. *JAMA Psychiatry.* 2019;76(3):324-335. doi:10.1001/jamapsychiatry.2018.3329
35. Stanislawski ER, Bilgrami ZR, Sarac C, et al. Negative symptoms and speech pauses in youths at clinical high risk for psychosis. *Npj Schizophr.* 2021;7. doi:10.1038/s41537-020-00132-1
36. Birnbaum ML, Norel R, Van Meter A, et al. Identifying signals associated with psychiatric illness utilizing language and images posted to Facebook. *Npj Schizophr.* 2020;6. doi:10.1038/s41537-020-00125-0
37. Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal Acoustic Biomarkers of Depression Severity and Treatment Response. *Biol Psychiatry.* 2012;72(7):580-587. doi:10.1016/j.biopsych.2012.03.015

38. Ni Y, Barzman D, Bachtel A, Griffey M, Osborn A, Sorter M. Finding warning markers: Leveraging natural language processing and machine learning technologies to detect risk of school violence. *Int J Med Inf.* 2020;139:104137.
39. Silva AM, Limongi R, MacKinley M, Ford SD, Alonso-Sánchez MF, Palaniyappan L. Syntactic complexity of spoken language in the diagnosis of schizophrenia: A probabilistic Bayes network model. *Schizophr Res.* Published online 2022. doi:10.1016/j.schres.2022.06.011
40. Birnbaum ML, Ernala SK, Rizvi AF, et al. Detecting relapse in youth with psychotic disorders utilizing patient-generated and patient-contributed digital data from Facebook. *Npj Schizophr.* 2019;5. doi:10.1038/s41537-019-0085-9
41. Belouali A, Gupta S, Sourirajan V, et al. Acoustic and language analysis of speech for suicidal ideation among US veterans. *BioData Min.* 2021;14. doi:10.1186/s13040-021-00245-y
42. Carrillo F, Sigman M, Fernández Slezak D, et al. Natural speech algorithm applied to baseline interview data can predict which patients will respond to psilocybin for treatment-resistant depression. *J Affect Disord.* 2018;230:84-86. doi:10.1016/j.jad.2018.01.006

Table 1. Cases in point and scientific questions relevant to the validation and potential use of NLP markers in psychiatry.

AI task	Variable	Clinical goal	Example of candidate NLP marker	Research questions
Detection	•Diagnosis	•Establish a categorical diagnosis (despite questionable validity).	Emotion-related acoustic features in speech differentiate unipolar depression and bipolar disorder ³³ .	What are the likely pathognomonic NLP markers for the different psychiatric disorders?
	▪Symptoms	▪Improve detection and quantification of symptoms to more efficiently provide patients with Measurement-Based Care ³⁴ .	In CHR youths, pause length and percentage of pauses positively correlated with total severity of negative symptoms ³⁵ .	With what periodicity should the assessment of symptoms occur (e.g., once or twice per day) and for how long (e.g., one vs three months) to obtain reliable estimates?
	♦Warning signs	♦Identify CHR individuals timely. ♦Study pre-symptomatic phases of mental disorders.	Prior to the first psychiatric hospitalization of patients with SSD, a relative increase in the use of swearwords and words related to perceptual processes and negative emotions ³⁶ .	Are there transdiagnostic and pathognomonic early-warning NLP markers? Do early-warning NLP markers manifest similarly across the lifespan?
Monitoring	•Treatment effects	•Monitor response to treatment actively, including side effects. •Minimize adverse effects from medication and increase adherence to optimal treatment.	In adults with major depression, pause behavior and mean fundamental frequency (pitch) differentiated treatment response ³⁷ .	Can NLP markers that vary with a treatment effect provide sufficient information to make decisions regarding changing, augmenting, or discontinuing treatments?
Prediction	▪Aggression/violence	▪Reduce the number of injuries, amount of harm and damage resulting from aggression or violence. ▪Reduce the use of coercive measures against aggressive or violent individuals.	In youth referred for psychiatric risk assessment, features such as words related to violence and temporal phrases related to the frequency of violent thoughts or acts were significantly associated with the risk of school violence ³⁸ .	Might NLP markers be predictive of types of aggression/violence (e.g., verbal vs physical) and who the target is?
	♦Psychosis onset	♦Stratification of CHR individuals for targeted preventive interventions.	Prior to initial psychosis onset in CHR, decrease in semantic cosine similarity, greater variance in similarity, and less usage of possessive pronouns ⁵ .	How early should NLP markers be measured in order to predict onset reliably? What predictive value will NLP markers have in non-help seeking samples?
	•Prognosis	•Estimate the course of a patient's psychiatric disorder and/or the probability of recovery.	In first episode psychosis, a drop in syntactic complexity over 6 months indicated a later diagnosis of schizophrenia ³⁹ .	Which NLP markers best predict outcomes such as social functioning, symptoms' remission, or vocational recovery?
	▪Relapse	▪Estimate relapse to improve preventive care.	For patients with psychosis, in the month preceding relapse there was a relative increase in the use of words	What is the best and actionable time frame for gathering relapse-prediction NLP markers data (e.g., every 2-4 weeks)?

	•Suicidality	<ul style="list-style-type: none"> •Improve assessment of suicidal ideation. •Prevent suicidal acts. 	<p>related to swearing, anger, death, and a decrease in words related to work, friends, and health along with more first and second person pronouns⁴⁰.</p> <p>Among USA veterans, a combined set of acoustic and linguistic features improved detection of suicidal ideation⁴¹.</p>	<p>Can NLP markers predict suicidal ideation and behavior with greater accuracy than existing risk calculators?</p> <p>Can NLP markers accurately distinguish between suicidal ideation and non-suicidal, negative thoughts?</p>
Selection	•Optimal treatment	<ul style="list-style-type: none"> •Select an optimal treatment to increase the probability of recovery. 	<p>In individuals with depression, scores of words with emotional content were predictive of treatment success with psilocybin⁴².</p>	<p>Can NLP markers assist in the identification of the optimal treatment for a given patient?</p>

AI: artificial intelligence; **CHR:** clinical high risk; **NLP:** natural language processing; **SSD:** schizophrenia-spectrum disorders; **USA:** United States of America.