Natural Language Processing Markers for Psychosis and Other Psychiatric Disorders Emerging Themes and Research Agenda From a Cross-Linguistic Workshop

Corona Hernández, Hugo; Corcoran, Cheryl; Achim, Amélie M.; de Boer, Janna N.; Boerma, Tessel; Brederoo, Sanne G.; Cecchi, Guillermo A.; Ciampelli, Silvia; Elvevåg, Brita; Fusaroli, Riccardo

DOI
10.1093/schbul/sbac215

Publication date
2023

Document Version
Final published version

Published in
Schizophrenia bulletin

Schizophrenia Bulletin vol. 49 suppl. 2 pp. S86–S92, 2023
https://doi.org/10.1093/schbul/sbac215

SUPPLEMENT ARTICLE

Natural Language Processing Markers for Psychosis and Other Psychiatric Disorders: Emerging Themes and Research Agenda From a Cross-Linguistic Workshop

Hugo Corona Hernández[1], Cheryl Corcoran[2], Amélie M. Achim[3], Janna N. de Boer[1,4], Tessel Boerma[5], Sanne G. Brederoo[1,6], Guillermo A. Cecchi[7], Silvia Ciampelli[1], Brita Elvevåg[8], Riccardo Fusaroli[9,10,11], Silvia Giordano[12], Mathias Hauglid[13], Arjan van Hessen[5,14], Wolfram Hinzen[15], Philipp Homan[16], Sybren F. de Kloet[17], Sanne Koops[1], Gina R. Kuperberg[18,19,20], Kritika Maheshwari[21,22], Natalia B. Mota[23,24], Alberto Parola[9,10], Roberta Rocca[10], Iris E. C. Sommer[1,6], Khiet Truong[14], Alban E. Voppel[1], Marieke van Vugt[25], Frank Wijnen[5], and Lena Palaniyappan[26,27,*]

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

[*]To whom correspondence should be addressed; 6875, Boul. LaSalle. Verdun (Quebec) H4H 1R3, Canada; tel: (514) 761-6131 Ext. 6116, fax: (514) 888-4458, e-mail: lena.palaniyappan@mcgill.ca

This workshop summary on natural language processing (NLP) markers for psychosis and other psychiatric disorders presents some of the clinical and research issues that NLP markers might address and some of the activities needed to move in that direction. We propose that the optimal development of NLP markers would occur in the context of research efforts to map out the underlying mechanisms of psychosis and other disorders. In this workshop, we identified some of the challenges to be addressed in developing and implementing NLP markers-based Clinical Decision Support Systems (CDSSs) in psychiatric practice, especially with respect to psychosis. Of note, a CDSS is meant to enhance decision-making by clinicians by providing additional relevant information primarily through software (although CDSSs are not without risks). In psychiatry, a field that relies on subjective clinical ratings that condense rich temporal behavioral information, the inclusion of computational quantitative NLP markers can plausibly lead to operationalized decision models in place of idiosyncratic ones, although ethical issues must always be paramount.

Key words: speech technology/implementation/digital markers/psychiatric practice/pathophysiology

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 31–November 4, 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 medicine,[1] 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 (ie, 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 (ie, semantic content), form (ie, 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 (eg, at home). While many developing markers are obtained using NLP techniques (eg, cosine semantic similarity metrics), markers derived using other computational approaches focused on human communication processes (eg, acoustics of speech signal and nonverbal 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 (eg, 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 (eg, 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 (eg, associations between cosine similarity metrics and tangentiality[4] in individuals at clinical risk for psychosis) and that are predictive of some outcomes of interest, such as transition to psychosis from risk states.[5] Robust external replications,[5] 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 actions[6] can be guided by a principled approach with three steps.[7] 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 (eg, 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

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 presymptomatic 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 nonhelp 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 related to swearing, anger, death, and a decrease in words related to work, friends, and health along with more first and second person pronouns.[40]	What is the best and actionable time frame for gathering relapse-prediction NLP markers data (e.g., every 2–4 weeks)?
	• Suicidality	• Improve assessment of suicidal ideation. • Prevent suicidal acts.	Among USA veterans, a combined set of acoustic and linguistic features improved detection of suicidal ideation.[41]	Can NLP markers predict suicidal ideation and behavior with greater accuracy than existing risk calculators? Can NLP markers accurately distinguish between suicidal ideation and nonsuicidal, negative thoughts?

Table 1. Continued

AI Task	Variable	Clinical Goal	Example of Candidate NLP Marker	Research Questions
Selection	• Optimal treatment	• Select an optimal treatment to increase the probability of recovery.	In individuals with depression, scores of words with emotional content were predictive of treatment success with psilocybin.[42]	Can NLP markers assist in the identification of the optimal treatment for a given patient?

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

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 tackled[8] so that interested parties can collaborate inter-institutionally[9] to advance the field.

NLP Markers and Mechanistic Research

Significant progress has been made in understanding the neural basis of language processing[10] and its interaction with neurocognitive processes such as attention[11] or memory.[12] Spoken language conveys information about impairments in thought and cognition in psychiatric disorders.[13] Thus, the mechanisms that underlie NLP markers might be in close proximity to the etiology of psychosis and other psychiatric disorders.[14] 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] (ie, 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 genetic[17] and developmental[18] processes. Furthermore, while an individual's anatomical[19] and cognitive[20] characteristics constrain its features, language production is influenced by pharmacological,[21] contextual,[22] and socio-demographic[23] 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 care.[24] 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 clinical decision support system (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 information,[24] and that training related to their acquisition and interpretation should have minimal burden on clinicians. Furthermore, along with accessibility to and acceptance of candidate markers by clinicians and patients,[25] any CDSS incorporating NLP markers should achieve expected standards of transparency, trust, and efficient and safe functioning[26] for regulatory approvals before widespread clinical use.[27] In the absence of a formal CDSS, clinical settings can implement NLP markers in pilot testing using human-in-the-loop iterative methodologies[28] 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 legislation.[29] 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-laws[30] 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 persist.[31] Specifically, unease about misuse (eg, discrimination) or potential harms (eg, 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 (eg, the European MONARCA project[32]), 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 (ie, clinicians and patients) will enable the implementation of informed modeling 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 article was enabled by a Distinguished Lorentz Fellowship granted to Iris Sommer in 2022 for Computational Linguistics to aid Diagnosis and Treatment Monitoring in Psychiatry. H.C.H. was supported by the Consejo Nacional de Ciencia y Tecnología (CONACyT, Mexican Government, scholarship number 739604). L.P. received research support from the Tanna Schulich Chair of Neuroscience and Mental Health (Schulich School of Medicine, Western University: 2019–2022); Canadian Institutes of Health Research (CIHR) Foundation Grant (154296); 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. A.P. is supported by a Marie Skłodowska-Curie Actions (MSCA)—Individuals Fellowship H2020-MSCA-IF-2018 (Grant agreement ID: 832518, Project: MOVES). A.M.A. received support from the Fonds de Recherche de Québec-Santé (FRQS). C.C. was supported by two grants of the National Institute of Mental Health: Using the RDoC Approach to Understand Thought Disorder: A Linguistic Corpus-Based Approach (5R01MH115332) and Thought Disorder and Social Cognition in Clinical Risk States for Schizophrenia (5R01MH107558). K.M. received funding from the NWO ELSA AI Lab Northern Netherlands.

Authorship contribution

H.C.H., I.E.C.S., and L.P. wrote the first draft of the manuscript. H.C.H., C.C., B.E., I.E.C.S., and L.P. 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. A.M.A., G.A.C., and G.R.K. were not able to attend the workshop.

Disclosures

L.P. reports personal fees from the Canadian Medical Association Journals for serving as chief editor, speaker/consultant fee from Janssen Canada and Otsuka Canada, SPMM 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. R.F. reports past consultant fees from F. Hoffmann-La Roche. N.B.M. works at Motrix, an EduTech startup, and has been a consultant to Boehringer Ingelheim. All other authors have no conflicts to disclose.

References

1. Vasudevan S, Saha A, Tarver ME, Patel B. Digital biomarkers: convergence of digital health technologies and biomarkers. *Npj Digit Med*. 2022;5(1):36. 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 [published online ahead of print, 2022 Nov 10]. *Schizophr Res*. 2022;S0920-9964(22)00362-0. 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 [published online ahead of print, 2022 Sep 21]. *Schizophr Res*. 2022;S0920-9964(22)00283-3. 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:10.1561/2200000083.
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*. Cambridge: 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, Zwicker 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;249:118795. 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. *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(1):27. 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(1):42. 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/fpsyt.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;3:17. 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*. 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*. Online edition, Oxford Academic; 2021. doi: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(1):3. 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*. 2022;S0920-9964(22)00245-6. 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(1):11. 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.