BMJ Open

Linguistic profile automated characterisation in pluripotential clinical high-risk mental state (CHARMS) conditions: methodology of a multicentre observational study

Luca Magnani ,[1,2] Luca Carmisciano,[3] Felice dell'Orletta,[4] Ornella Bettinardi,[5] Silvia Chiesa,[5] Massimiliano Imbesi,[5] Giuliano Limonta,[5] Elisa Montagna,[1,2] Ilaria Turone,[1,2] Dario Martinasso,[1,2] Andrea Aguglia,[1,2] Gianluca Serafini,[1,2] Mario Amore,[1,2] Andrea Amerio,[1,2] Alessandra Costanza ,[6,7] Francesca Sibilla,[1,2] Pietro Calcagno,[1,2] Sara Patti,[8] Gabriella Molino,[8] Andrea Escelsior,[1,2] Alice Trabucco,[1,2] Lisa Marzano ,[9] Dominique Brunato,[4] Andrea Amelio Ravelli,[4] Marco Cappucciati,[5,10] Roberta Fiocchi,[5] Gisella Guerzoni,[5] Davide Maravita,[5] Fabio Macchetti,[5] Elisa Mori,[5] Chiara Anna Paglia,[5] Federica Roscigno,[5] Antonio Saginario,[5] LNG-PSY Study Investigators

To cite: Magnani L, Carmisciano L, dell'Orletta F, *et al*. Linguistic profile automated characterisation in pluripotential clinical high-risk mental state (CHARMS) conditions: methodology of a multicentre observational study. *BMJ Open* 2023;**13**:e066642. doi:10.1136/bmjopen-2022-066642

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-066642).

Received 13 July 2022
Accepted 08 March 2023

For numbered affiliations see end of article.

Correspondence to
Dr Luca Magnani;
magnani1991@gmail.com

ABSTRACT

Introduction Language is usually considered the social vehicle of thought in intersubjective communications. However, the relationship between language and high-order cognition seems to evade this canonical and unidirectional description (ie, the notion of language as a simple means of thought communication). In recent years, clinical high at-risk mental state (CHARMS) criteria (evolved from the Ultra-High-Risk paradigm) and the introduction of the Clinical Staging system have been proposed to address the dynamicity of early psychopathology. At the same time, natural language processing (NLP) techniques have greatly evolved and have been successfully applied to investigate different neuropsychiatric conditions. The combination of at-risk mental state paradigm, clinical staging system and automated NLP methods, the latter applied on spoken language transcripts, could represent a useful and convenient approach to the problem of early psychopathological distress within a transdiagnostic risk paradigm.

Methods and analysis Help-seeking young people presenting psychological distress (CHARMS+/– and Clinical Stage 1a or 1b; target sample size for both groups n=90) will be assessed through several psychometric tools and multiple speech analyses during an observational period of 1-year, in the context of an Italian multicentric study. Subjects will be enrolled in different contexts: Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), Section of Psychiatry, University of Genoa—IRCCS Ospedale Policlinico San Martino, Genoa, Italy; Mental Health Department—territorial mental services (ASL 3—Genoa), Genoa, Italy; and Mental Health Department—territorial mental services (AUSL—Piacenza), Piacenza, Italy. The conversion rate to full-blown psychopathology (CS 2) will be evaluated over 2 years of clinical observation, to further confirm the predictive and discriminative value of CHARMS criteria and to verify the possibility of enriching them with several linguistic features, derived from a fine-grained automated linguistic analysis of speech.

Ethics and dissemination The methodology described in this study adheres to ethical principles as formulated in the Declaration of Helsinki and is compatible with International Conference on Harmonization (ICH)-good clinical practice. The research protocol was reviewed and approved by two different ethics committees (CER Liguria

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ Using validated diagnostic criteria, the study aims to improve the characterisation of early psychopathology with the results of a fine-grained analysis of language, hoping to define proper linguistic biomarkers.
⇒ The selection of relevant linguistic features is performed through a data-driven approach, without predefined cut-offs correlated to pathological significance.
⇒ Spoken-language data can be highly variable, and a key challenge concerns the optimisation of such a phase of data acquisition to allow the extraction of relevant information during the subsequent phase of textual processing.
⇒ The retention of participants for the entire duration of the observational period is a further challenge.
⇒ The assessment of conversion to full-blown psychopathology during the second year of prolonged observation may represent a methodological issue.

approval code: 591/2020—id.10993; Comitato Etico dell'Area Vasta Emilia Nord approval code: 2022/0071963). Participants will provide their written informed consent prior to study enrolment and parental consent will be needed in the case of participants aged less than 18 years old. Experimental results will be carefully shared through publication in peer-reviewed journals, to ensure proper data reproducibility.

Trial registration number DOI:10.17605/OSF.IO/BQZTN.

INTRODUCTION

Language, thought and human beings

Language is usually considered the social vehicle of thought in the context of intersubjective communications. This canonical interpretation of the thought-language relationship implicitly entails the priority of the first term. Therefore, in common medical practice, the verbalisation of delirious contents during an acute psychotic episode is notoriously considered as the manifestation of an underlying thought disorder.[1] On the other hand, interspecies cognitive studies revealed the existence of a fundamental gap between the expression of linguistic and non-linguistic contents.[2–4] The relationship between language and high-order cognition seems to evade the canonical unidirectional description and the common idea of a thought primacy. In fact, it is at least plausible that language acquisition exhibits a critical role for human cognitive development[5]: a subsequent deficit in cognitive functions has been experimentally linked to a primary insufficient development of linguistic skills.[6–8] Moreover, according to Tattersall's theory,[9 10] language should be regarded as the fact that permitted the transition across different cognitive phenotypes during the evolution of human species. In philosophy, many authors highlighted the deep constitutive character of language for humans.[11 12]

Language and phenomenal experience

As well known, consciousness definition is still an ongoing issue. Among other things, it could be simplistically represented as an active and fundamental background that synthetise experienced phenomena within a spatiotemporal schema. This said, the linguistic apparatus, analysable through advanced natural language processing (NLP) techniques, can be considered as something that shapes the product of this primary synthesis to further reduce the phenomenal complexity and to allow the emergence of a unique and well-defined subjective experience. Therefore, the conscious phenomenon offers itself as a synthetic elaboration of an unrefined experience, following a double logic (first transcendental and then formal-linguistic). In this context, language stops being exclusively a means of communication for predefined thoughts, thus becoming a refined tool of interaction with the experienced world, regardless of what it can be said about abstract thought function. In some recent interpretations this complex problem has been transposed within the theoretical framework of predictive brain[13 14] to show that 'verbal cues (even if self-generated) can act as highly flexible (and metabolically cheap) contexts (set of priors), to generate a predictive signal helping the system to process an input that is otherwise too weak or noisy".[15]

Table 1 Clinical staging system

Stage	Symptoms	Functioning
0	No current symptoms; increased risk of disorder	No historical change
1a	Mild or non-specific symptoms	Mild functional change/ decline; (GAF 70–100)
1b	Moderate but subthreshold symptoms	Functional decline (GAF<70)
2	Full-threshold disorder with moderate-to-severe symptoms	Functional decline (GAF<50)
3	Incomplete remission or relapse	Persistent functional decline
4	Severe, unremitting or refractory illness	Poor treatment effectiveness despite persistently intensive interventions

From Shah *et al*.[37]
GAF, Global Assessment of Functioning.

Psychopathology and language: floating on fluid psychopathological substrates

Most relevant psychopathological conditions appear to originate during early life stages.[16 17] At the same time, the boundaries between previously distinct psychopathological disorders seem to weaken.[18–20] The inadequacy of classical nosography has become progressively more evident, especially when it is applied to the dynamicity of early psychopathological phenomena.[21 22] A classification based on strictly defined diagnostic categories appears inadequate when considering the complex phenomenon of comorbidities[23 24] and symptomatological overlaps, frequently expressed during early stages of mental illness.[25–29] To address the complex world of early psychopathology the so-called Ultra-High-Risk paradigm has been proposed over the past decades,[30] originally developed to individuate schizophrenia prodromes.[31] More recently, the concept of at-risk mental state has been expanded to detect conditions of 'trans-diagnostic risk'.[32–35] Coherently, Hartmann and colleagues[36] proposed a new methodology based on the application of clinical high at-risk mental state (CHARMS) criteria and enriched with the introduction of a clinical staging system[37] (table 1). Specifically, the detection of the CHARMS criteria allows to identify adolescents and young adults (aged up to 25 years old) expressing a psychopathological condition of transdiagnostic risk, which corresponds to stage 1b in the context of the clinical staging system (CS 1b). Coherently with the transdiagnostic approach, the risk is referred to a generic 'exit syndrome', that is, a first episode of full-blown psychopathology, defined through the overcoming of some psychometric thresholds, as well as through the

verification of DSM-5 (Diagnostic and Statistical Manual of Mental Disorders - Fifth Edition) criteria for a specific disorder. Within the wide group of CHARMS+ / CS 1b subjects, some subcategories of 'attenuated syndromes' have been proposed by original authors (see table 2): the psychosis trait vulnerability group; the bipolar trait vulnerability group; the attenuated psychotic symptoms group; the attenuated (hypo)manic symptoms group; the moderate (attenuated) depression group; the attenuated borderline personality group; the brief limited intermittent psychotic symptom group. An ongoing observational study has been structured starting from this latter proposal, aimed to verify the predictive power of some novel risk category defined through the application of CHARMS criteria.[38] Preliminary results have recently been published.[39] To conceive our proposal and to first verify data reproducibility, we initially acquired the methodology described by Hartmann and colleagues.[36] However, we chose to further enrich the experimental apparatus to gather different information from a fine-grained analysis of the linguistic profile, derived from subjects' speech. This analysis will be conducted on textual data, revised by researchers, after a first automated transcription of the audio records, directly acquired during experimental assessment. In fact, language, as previously reinterpreted, can represent a further pathoplastic/pathogenetic factor, favouring the pathological crystallisation of phenomenal data.

Language production disturbances in psychosis and schizophrenia have been investigated since some early works promoted by Harrow and Quinlan[40] and Andreasen and Grove,[41] the sudden development of innovative methods of automated linguistic analysis promoted further investigations (for a comprehensive review see Corcoran *et al* 2020,[42] increasingly oriented towards the early stages of the disease. In fact, it seems that language alterations may soon represent valid and practical biomarkers to perform a multilayered assessment of psychotic risk and to offer more tailored interventions. The hope concerns the possibility of extending this approach within the abovementioned transdiagnostic risk paradigm.

NLP techniques and their application in neuropsychiatric conditions

As reported by Voleti and colleagues,[43] it is possible to identify different levels of complexity in linguistic analysis. For each level, several features have been proposed as potentially relevant in association with different neuropsychiatric conditions.

Lexical level

This level allows to extract features that account for the diversity and richness of lexicon used in a text. At this level, the following metrics are usually computed: (a) the **type/token ratio (TTR)**, a standard index of lexical variety; (b) the **Moving Average type-token ratio**,[44] considered as an 'advanced' TTR as it calculates lexical variety of a sample using a moving window that estimates TTRs for each successive window of fixed length. *Moving Average TTR, Brunét Index, Honoré's Statistic, part-of-speech (POS) tagging,* aim to quantify lexical diversity and density. These parameters have been mainly studied either for risk stratification or for the diagnosis of morbid conditions of purely neurological relevance.[45–49] These variables are quick and easy to assess. However, this simplicity is reflected in a reduced capacity of providing relevant information.

Morpho-syntactic level

This level allows to extract information from the POS-tagging step of linguistic annotation. In particular, the following variables are usually computed (a) the percentage distribution of morpho-syntactic categories (both functional and lexical); (b) the 'lexical density' index (ie, the proportion between functional words over the total number of words in a text).

Syntactic level

At this level whole propositions are examined to analyse the way in which words are organised in sentences and sentences in speeches. Considering the work of Mota and colleagues,[50] the researchers extrapolated objective parameters of language measurement, useful for quantifying the alterations characteristically found in specific morbid states. The set of verifiable syntactic features covers a wide range of properties which can be further grouped. For instance, features related to the parse tree structure (eg, maximum parse tree depth, average length of dependency links), to the use of specific syntactic relations (eg, use of coordination and subordination) and to canonicity effects (eg, relative order of subject and object with respect to the verb).

Semantic analysis

The analysis of linguistic expressions in relation to the meaning they acquire in speech. Among related NLP methods, one of the first developed is the Latent Semantic Analysis (LSA),[51] today carried out through the application of specific Machine Learning algorithms, exploiting artificial neural networks word2vec or GloVe.[52 53] These tools can probabilistically define, starting from the analysis of large textual corpora, the semantic content of individual words and develop a specific vocabulary. More recently, further algorithms have been developed that can operate similarly at the level of entire propositions (eg, sent2vec, InferSent, Universal Sentence Encoder - USE). One of the first studies carried out in this area aimed to measure, by applying LSA, the semantic coherence of the language of patients suffering from Formal Thought Disorder (FTD) of different severity.[54] Furthermore, through these methods the importance of semantic and pragmatic alterations in schizophrenia was confirmed.[55] In recent years, the characteristics of language potentially predictive for the onset of psychosis in clinically defined high-risk subjects (clinical high-risk, CHR) were also investigated.[56 57] Similarly, Rezaii and colleagues[58] defined a

Table 2 CHARMS subgroups of risk or early clinical phenotypes

Subgroup	Instrument	Description
Psychosis trait vulnerability group	SCID-5-PD, SOFAS, medical and family history	**Family history of psychosis in first-degree relative OR schizotypal personality disorder.** **AND** (SOFAS score of 50 or less for over 12 months **OR** SOFAS score at least 30% below previous level.)
Bipolar trait vulnerability group	SCID-5 medical and family history	**Depression+cyclothymic features/genetic risk group:** **Depression:** for at least 1 week: depressed mood or loss of interest or pleasure and at least two criteria from the list: (1) significant weight loss, (2) insomnia or hypersomnia nearly every day, (3) psychomotor retardation or agitation, (4) fatigue or loss of energy, (5) feelings of worthlessness or excessive or inappropriate guilt, (6) diminished ability to think or concentrate, (7) recurrent thoughts of death, recurrent suicidal ideation. The episode cannot be due to the direct physiological effects of a substance or medical condition and does not need to cause a clinically significant drop in functioning. **AND** **Cyclothymic features:** for a minimum of 6 months (lifetime) high and low mood (no more than two consecutive months without symptoms) and at least three criteria from the list: (1) decreased need for sleep (eg, feels rested after only 3 hours sleep), (2) increased energy, (3) inflated self-esteem or grandiosity, (4) increased goal directed activity, (5) restlessness, (6) more talkative than usual or pressure to keep talking, (7) unusual ideas, clear thinking, (8) troublesome behaviour, (9) inappropriate sense of humour. The episode cannot be due to the direct physiological effects of a substance or medical condition and does not need to cause a clinically significant drop in functioning. OR **Genetic risk:** first degree relative with bipolar disorder.
Attenuated psychotic symptoms group	CAARMS	**Intensity:** Global rating scale score of 3–5 on unusual thought content subscale, 3–5 on non-bizarre ideas subscale, 3–4 on perceptual abnormalities subscale and/or 4–5 on disorganised speech subscales (symptoms present for at least 1 week in the last year). **AND** **Frequency:** Scale score of 3–6 on unusual thought content, non-bizarre ideas, perceptual abnormalities and/or disorganised speech subscales. **OR** **Intensity:** Global rating scale score of 6 on unusual thought content subscale, 6 on non-bizarre ideas subscale, 5–6 on perceptual abnormalities subscale and/or 6 on disorganised speech subscales (symptoms present for at least 1 week in the last year). **AND** **Frequency:** Scale score of 3 on unusual thought content, non-bizarre ideas, perceptual abnormalities and/or disorganised speech subscales.
Attenuated (hypo)manic symptoms group	SCID-5	A distinct period of abnormally and persistently elevated, expansive or irritable mood and ≥2 (3 if irritable) of the following 'B' criteria for at least 2 days: Inflated self-esteem or grandiosity; decreased need for sleep (eg, feels rested after only 3 hours sleep); more talkative than usual or pressure to keep talking; flight of ideas or subjective experience that thought are racing; distractibility; increased goal-directed activity (either socially, at work or sexually) or psychomotor agitation; excessive involvement in pleasurable activities which have a high potential for painful consequences (eg, engaging in unrestrained buying sprees, sexual indiscretions or foolish business investments). The duration of this period can be maximal 3 days if ≥3 'B' criteria are met (≥4 if irritable) and it is associated with an unequivocal change in functioning that is uncharacteristic of the person when not symptomatic (criterion 'C'); and the disturbance in mood and the change in functioning is observable by others (criterion 'D'). The duration of this period can be maximal 6 days if: ► ≥3 'B' criteria are met and 'C' or 'D'. ► ≥3 'B' criteria are met and neither 'C' or 'D' are. ► 2 B criteria in any combination with 'C' and 'D'. The episode is not severe enough to cause marked impairment in social or occupational functioning, or to necessitate hospitalisation, and there are no psychotic features. The symptoms are not due to the direct physiological effects of a substance (eg, a drug of abuse, a medication or other treatment).

Continued

Table 2 Continued

Subgroup	Instrument	Description
Moderate (attenuated) depression group	SCID-5, QIDS-C	Major depressive episode (current or past) For at least 2 weeks: depressed mood or loss of interest or pleasure+at least five criteria from the list: (1) significant weight loss, (2) insomnia or hypersomnia nearly every day, (3) psychomotor retardation or agitation, (4) fatigue or loss of energy, (5) feelings of worthlessness or excessive or inappropriate guilt, (6) diminished ability to think or concentrate, (7) recurrent thoughts of death, recurrent suicidal ideation. The MDE must also: (1) not be due to the direct physiological effects of a substance, (2) cause a clinically significant drop in functioning and (3) not be better accounted for by bereavement. **AND** Current QIDS-C score: 11–15.
Attenuated borderline personality group	SCID-5-PD	For at least 6 months: at least two but less than five criteria from the list: 1. Frantic efforts to avoid real or imagined abandonment. 2. A pattern of unstable and intense interpersonal relationships characterised by alternating between extremes of idealisation and devaluation. 3. Identity disturbance: markedly and persistently unstable self-image or sense of self. 4. Impulsivity in at least two areas that are potentially self-damaging (eg, spending, sex, substance abuse, reckless driving and binge eating). This does not include suicidal or self-harming behaviour. 5. Recurrent suicidal behaviour, gestures or threats, or self-mutilating behaviour. 6. Affective instability due to a marked reactivity of mood—intense feelings that can last from a few hours to a few days. 7. Chronic feelings of emptiness. 8. Inappropriate intense anger or difficulty controlling anger. 9. Transient, stress-related paranoid ideas or severe dissociative symptoms.
Brief limited intermittent psychotic symptom group	CAARMS	**Intensity:** Global rating scale score of 6 on unusual thought content subscale, 6 on non-bizarre ideas subscale, 5 or 6 on perceptual abnormalities subscale and/or 6 on disorganised speech subscales (symptoms present for less than 1 week in the last year). **AND** **Frequency:** Frequency scale score of 4–6 on unusual thought content, non-bizarre ideas, perceptual abnormalities and/or disorganised speech subscales.

MDE, Major Depressive Epysode.

so-called digital phenotype useful for quantifying the risk of psychosis onset in CHR subjects. According to Morgan and colleagues,[59] different NLP measures may provide complementary information, being potentially associable to distinct aspects of mental disorders.

Aims and objectives

According to our theoretical speculations and inspired by previous works proposed by Hartmann and colleagues,[36 39] we designed a multicentre observational study with the following objectives:

Primary objective

- To investigate the alterations of multiple spoken language variables in subjects at pluripotential risk (CHARMS+ and Clinical Stage 1b) of developing a range of full-blown psychopathological disorders by estimating the associations between those variables and the conversion to full-blown psychopathological disorder in reference to a second group (CHARMS– and CS 1a), internally defined after the exclusion of the presence of a full-blown disorder (CS 2), as well as of a transdiagnostic risk condition (CHARMS+ and CS 1b).
- To prospectively confirm the predictive and discriminant validity of the CHARMS criteria in a sample of CHARMS+ and CS 1b subjects and in a second group (CHARMS– and CS 1a).

Secondary objectives

1. To develop a prediction model for the probability of conversion to full-blown disease (Clinical Stage 2), using CHARMS criteria, CHARMS subgroups, linguistic features and a data-driven subset of language markers.
2. To evaluate the predictive and discriminative capabilities of such model.

Crucially, the conversion to full-blown conditions (ie, Clinical stage 2) is referred to a set of 'exit-syndromes', in line with the original methodology (ie, psychotic disorder, bipolar disorder, depressive disorder and borderline personality disorder).

METHODS AND ANALYSIS

Participants and setting

We designed a longitudinal follow-up study with an observational period of 2 years. The expected sample size is n=180: 90 subjects who meet CHARMS criteria (Group A: CHARMS+; CS 1b) and 90 controls (Group B: CHARMS−; CS 1a). Potential participants are help-seeking people aged 14–25 who are referred to local mental health service. Subjects will be enrolled in different contexts:

- Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), Section of Psychiatry, University of Genoa—IRCCS Ospedale Policlinico San Martino, Genoa, Italy;
- Mental Health Department—territorial mental services (ASL 3 – Genoa), Genoa, Italy;
- Mental Health Department—territorial mental services (AUSL – Piacenza), Piacenza, Italy.

Group A (CHARMS+ and CS 1b): attenuated syndrome with moderate/subthreshold symptoms.

Inclusion criteria are: (1) the verified presence of a pluripotential 'at-risk' mental state—CHARMS+ and Clinical Stage 1b attenuated syndrome; (2) age between 14 and 25 Years; (3) native speakers (Italian) with a good understanding of spoken and written Italian language; (4) the ability to give informed consent by participants themselves and/or by parental authority holders in the case of a minor.

Exclusion criteria are: (1) a documented history of intellectual disability or a diagnosis of autism spectrum disorder; (2) current or previous diagnosed full-blown mental disorder—Clinical Stage ≥2; (3) the presence of relevant neurological disorders (including brain trauma, epilepsy, stroke, cerebral palsy); (4) a mental condition directly and exclusively induced by substances or by organic causes (in accordance with DSM-V criteria); (5) the identification through psychic examination of delusional persecutory thoughts, such that the experimental audio recording may be a source of discomfort for the enrolled subject.

Group B (CHARMS− and CS 1a): mild or non-specific symptoms.

Inclusion criteria are: (1) the exclusion of a pluripotential 'at-risk' mental state—CHARMS+ and Clinical Stage 1b attenuated syndrome; (2) age between 14 and 25 Years; (3) native speakers (Italian) with a good understanding of spoken and written Italian language; (4) the ability to give informed consent by participants themselves and/or by parental authority holders in case of a minor.

Exclusion criteria are: (1) a documented history of intellectual disability or a diagnosis of autism spectrum disorder; (2) current or previous diagnosed full-blown mental disorder—Clinical Stage≥2; (3) the presence of relevant neurological disorders (including brain trauma, epilepsy, stroke, cerebral palsy); (4) a mental condition directly and exclusively induced by substances or by organic causes (in accordance with DSM-V criteria).

Procedure and data acquisition

Baseline interview A—psychometric measures

For each enrolled subject, the necessary informed consent is acquired (for participants aged <18 years, parental consent will also be obtained). The baseline examination will be conducted by preselected and pretrained research team members and will involve a general neuropsychiatric evaluation and a structured recording of medical/family history. The following psychometric scales will be administered to finally verify inclusion/exclusion criteria and subjects eligibility:

- **Comprehensive Assessment of At-Risk Mental States[38]—semi-structured interview, seven different psychopathological domains rated according to a global rating score (0–6), a frequency score (0–6) and a substance use score (0–2). Italian validation.[60]
- **Structured Clinical Interview for DSM-5 (SCID-5)[61] and SCID-5 Personality Disorders (SCID-5_PD)[62]—semi-structured interviews for establishing clinical diagnoses (gold standard) based on DSM-5. Regarding the SCID-5_PD section, only the modules concerning borderline personality disorder and schizotypal personality disorder are considered for the purposes of this study.
- **Social and Occupational Functioning Scale (SOFAS)[63]—observer-rated (0–100) scale assessing the social and occupational functioning.
- Global Functioning Scale: Social and Role (GFS/GFR)[64 65] (clinician rated)—both deriving from Global Assessment of Functioning, GFS assesses (from 1—extreme dysfunction to 10—superior functioning) quantity and quality of social relationship, while GFR assesses (same scoring system) subject's performance in different contexts (school, work or home).
- Quick Inventory of Depressive Symptomatology-Clinicians rated (QIDS-C)[66]—clinician-rated 16-items questionnaire that assess the severity of depressive symptoms during the previous week.
- **Young Mania Rating Scale[67]—11 clinician-rated items that assess (gold-standard)[68] severity of manic symptomology over the previous 48 hours. Italian version.[69]
- Depression Anxiety Stress Scale 21 (DASS-21)[70]—self-report scale (short version of the 42-item DASS) to assess three domains (seven items for each domain) of negative affectivity referred to the past weeks. When the scale is administered in children and adolescents, only one (general) score is defined.[71]
- Bipolar Spectrum Diagnostic Scale (BSDS)[72]—self-rating narrative-based scale which assesses the entire bipolar spectrum, including subthreshold states of bipolar illness.[73]
- **Personality inventory for DSM-5, brief version[74]—self-rating screening[75 76] tools (25 items) for assessing in adult and adolescents five maladaptive personality trait-dimensions, described according to the alternative model of personality disorder.

- Davos Assessment of Cognitive Biases Scale (DACOBS)[77 78]—self-report scale with 42 items to assess the presence of possible cognitive biases, cognitive limitations and avoidance behaviours.
- Munich Cronotype Questionnaire (MCTQ)[79]—a self-report tool to assess information on sleep referred to work and work-free days and to quantitatively obtain a chronotype related to sleep intervals.
- **Insomnia Severity Index[80 81]—self-rating (seven items) instrument to assess night-time and daytime symptoms of insomnia in adults and adolescents.[82] Italian version.[83]

** Italian version available.

As indicated, some of these psychometric tools (ie, QIDS-C, SOFAS, GFS/GFR, DASS-21, BSDS, DACOBS and MCTQ) have not been officially validated in Italian, thus a preliminary internal translation was realised to reproduce as much as possible a certain methodology[36] and to eventually perform an internal validation of the abovementioned psychometric tools.

If a participant will exceed the threshold for a full-threshold disorder (Clinical Stage ≥2), then he/she will be excluded from the study and committed to the mental health service. On the contrary, if the subject will meet the CHARMS criteria (CHARMS+ and CS 1b) or falls below the CHARMS threshold (CHARMS– and CS 1a), then he/she will be included in the research programme (unless additional exclusion criteria are met).

Through the acquisition of psychometric variables and the application of CHARMS criteria it is also possible to verify for each subject in Group A the subgroup of risk, as proposed by Hartmann *et al*[86] (table 2).

Baseline interview B—speech recording

The baseline assessment includes a second part to be carried out after a few days (T0-b), according to a shared agenda and by the same research team members.

At T0-b, subjects of both groups will be first evaluated through the Montreal Cognitive Assessment scale.[84 85] Besides, the spoken language of enrolled subjects will be audio-recorded. Participants will describe four sequences of vignettes, picturing four logically linked events. In each sequence human individuals engaged in simple actions within contexts of daily life are represented. Two sequences were specifically created to be affectively neutral; a third one should be more emotionally salient; finally, a fourth sequence should express a less intuitive logic of transition between the single depicted moments. Then, they will be asked to answer four predefined questions, each related to a particular detail of each picture in a sequence. The free speech of each participant will also be recorded, eventually elicited with some questions, formulated according to narrative interview's recommendations (ie, phenomenological inquiry paradigm[86]

The recording sessions will last 30–45 min. Data will be acquired using the same recording device and using a free software (Auphonic), with the following settings:

- Format: CAF/WAV (PCM).
- Sample rate: 44 khz or 48 khz.
- Channel: MONO.
- Depth: 16 bits.

Time series

The first year of observation includes three phases of data acquisition for each participant.

During each phase, all relevant information about participants will be registered/updated. Linguistic data will be recorded at Tn-b, following the same methodology of acquisition described for T0-b. Data acquired at each time point are summed in table 3.

Crucially, at each different phase of the first year, subjects' conversion to full-blown disorder (Clinical Stage 2) will be verified. Each participant will be also periodically evaluated by mental health specialists (not directly involved in the study), who will eventually provide him/her with any appropriate therapeutic intervention. During the second year of observation subjects will be specifically assessed for conversion to Clinical Stage 2 any time a significative worsening of psychological status will be reported from the abovementioned standard periodical evaluation performed by external mental health teams.

Linguistic data processing and elaboration

At each phase of data acquisition, audio reports will be automated transcribed verbatim under the supervision of dedicated researchers. A first database of anonymous raw transcripts will be produced. In such a form, transcripts will be shared with our partner institution (Computational Linguistic Institute 'Antonio Zampolli', National Research Council (CNR), Pisa, Italy) to perform textual data processing. Starting from the transcripts, linguistic analyses will be performed along different levels. Raw text and (morpho-)syntactic analysis will be automatically carried out by means of Profiling-UD,[87] a multilingual web-based tool that provides a comprehensive assessment of language use. The tool performs a two-stage process: linguistic annotation (carried out by UDPipe)[88] according to Universal Dependencies (UD) formalism[89] and linguistic profiling. The annotated texts will be used as input to the further step, performed by the linguistic profiling component that defines the rules to extract and quantify the formal properties. The final output of the process is a vector-like representation that can comprises more than 120 linguistic features: (1) shallow features, for example, average length and counts of words and sentences, (2) morpho-syntactic features, for example, POS tagging distributions and inflectional properties of verbs or (3) more complex features obtained from syntactic parsing of the sentences, such as the use of subordination. The set of features from Profiling-UD has been derived from the literature on linguistic complexity, language acquisition and neurolinguistics, and have been successfully applied in a wide range of tasks and scenarios: from the automatic tracking of developmental patterns in child language acquisition[90 91] and the evolution of written language competence in school learners,[92 93] to

Table 3 Gantt chart

	Baseline		6 months		12 months	
Type of survey	T0/a	T0/b	T1/a	T1/b	T2/a	T2/b
Basic data						
Informed consent	X					
Recruitment form	X					
Therapeutic interventions	X		X		X	
Other relevant contingencies	X		X		X	
Interviews						
CAARMS	X		X		X	
SCID-5	X		X		X	
SCID-5-PD*	X		X		X	
QIDS-C	X		X		X	
SOFAS	X		X		X	
GSF/GFR	X		X		X	
YMRS	X		X		X	
Self- administered						
DASS-21	X		X		X	
DACOBS	X					
PID-5-BF	X		X		X	
BSDS	X					
ISI	X		X		X	
Audio recording		X		X		X
MoCA		X		X		X

*Borderline personality disorder module and schizotypal personality disorder module.
BSDS, Bipolar Spectrum Diagnostic Scale; CAARMS, The Comprehensive Assessment of At-Risk Mental States; DACOBS, Davos Assessment of Cognitive Biases Scale; DASS-21, Depression Anxiety Stress Scale (21 items version); GFR, Global Functioning Scale: Role; GFS, Global Functioning Scale: Social; ISI, Insomnia Severity Index; MoCA, Montreal Cognitive Assessment; PID-5-BF, The Personality Inventory for DSM-5, Brief Version; QIDS—C, Quick Inventory of Depressive Symptomatology—Clinician rated; SCID-5, Structured Clinical Interview for DSM-5; SCID-5-PD, Structured Clinical Interview for DSM-5 Personality Disorders.; Sleep Questionnaire, Insomnia Severity Index and Munich ChronoType Questionnaire; SOFAS, Social and Occupational Functioning Scale; YMRS, Young Mania Rating Scale.

the prediction of behavioural and cognitive impairments based on the detection of relevant linguistic markers from clinical tests.[46 94]

Furthermore, semantic representations of each transcript will be computed for both single sentence and the whole session level. To this end, we will rely on state-of-the-art neural network architecture, for example, Transformers models,[95] which have shown massive improvements in NLP. Particularly, Natural Language Understanding models based on this technology, such as those of the BERT family[96] have defined new states of the art in many tasks (eg, GLUE collection of benchmark tasks). The main advantage of these recent models over previous methodologies is that the embedding of a word is not fixed but computed for every occurrence based on its lexical contour; they can also applied in pathological contexts[97]. We plan to exploit a pretrained BERT model for the Italian language that has been trained on a huge corpus of more than 13 billions of words, that is, 'bert-base-italian-cased' to encode semantic information of words and sentences. Following Corcoran *et al*,[42] we plan to analyse the coherence in the flow of subject speech by computing the semantic closeness of contiguous sentences (ie, the cosine distance between the embedding representations of the sentences).

Duration of the study

The recruitment kick-off is scheduled for July 2022; the database lock will be carried out according to the achievement of a sufficient sample size. The minimum expected duration of the observational period for a non-dropped out participant is 2 years. Preliminary data referred to each participant will be analysed at the end of recruitment phase. After this first passage data could possibly undergo a correction process due to the potential delayed conversion to full-blown disease (Clinical Stage 2) during the second year of observation.

Patient and public involvement

Research questions and outcomes were defined to address the complexity of early psychopathology and

to better correspond to help-seeking young people needs, frequently expressed in real-world mental health settings. The richness of patients' symptoms descriptions and their expressive urgency guided the development of the study design, prompting us to focus our attention on the characteristics of spoken language. At the end of the period of data acquisition, the results of the experimental investigations could help inform patients' primary advisers, potentially optimising the care offer. Furthermore, during the individual assessment, the exceeding of predefined psychopathological thresholds (conversion to full-blown disease—Clinical Stage 2) will be communicated to the patient's advisers to immediately adopt appropriate therapeutic measures.

Statistical analyses

Estimated sample size and statistical power

As reported by Hartmann and colleagues,[36] literature'-based expectations of the 1-year transition rate in the CHARMS+ and CHARMS− groups are, respectively, 20% and 3%. To detect such a 6.7-fold increase as significant with 90% power, 5% significance level and 20% drop-out rate, a total of 180 subjects are required. Hence, we defined an expected sample size of n=180 (n=90 group A and n=90 group B).

Power calculations were performed by simulation with R software V.4.0.2. All the supporting material is available at a publicly repository.

Statistical analyses post data acquisition

Primary analysis will determine whether the rate of patient's conversion to full-blown disease (Clinical Stage 2) in CHARMS+ patient's group differs from the rate in CHARMS− group. Pearson's χ^2 test will be performed on the 2×2 contingency table of patient's group (CHARMS+ or CHARMS−) and the occurrence of the Stage 2 conversion over a fixed follow-up time (which is one or 2 years of observation in preliminary and final analyses, respectively).

In the presence of heterogeneity in drop-out rates between the two patient's groups the main analysis will be performed both in the complete case data set and under the assumption of the conversion rate to 1 (most conservative approach) for all drop-out patients.

In the presence of significative imbalances of patient's characteristics between the two patient's groups a multivariable logistical regression analysis will be used to adjust for potential confounders. Both raw and adjusted analysis will be reported with OR, 95% CIs and p values. A sensitivity analysis considering the conversion event as a time dependent outcome will be preplanned to make full usage of all available follow-up periods and make each day of observation contribute to the final conversion rate estimation.

In the case of difference between the rate of Stage 2 conversion in CHARMS+ and CHARMS− patients three classification analyses will be performed to detect: (1) the clinical alterations, (2) the CHARMS subgroups and (3) the spoken language features most associated to the Stage 2 conversion regardless the CHARMS group. Within these three analyses variable selection procedures might be applied, such as: bidirectional stepwise, ridge, lasso or elastic net to identify the best combination of predictors to identify Stage 2 conversion. Receiver Operating Characteristic curve, the area below such curve and metrics derived from the 2×2 prediction-observation confusion matrix (such as sensitivity and specificity, positive and negative predictive values) will be used to estimate model's predictive and discriminative capabilities. Internal validation procedures such as k-fold cross validation may also be required to improve model generalisability. If needed, external information source such as real-world prevalence rates may be used to contextualise model performance.

Due to the high number of tested, there is a high risk of labelling some false, spurious, associations as significative. Therefore, we define a three-step strategy aimed at mitigating this risk. First, we prospectively list a set of characteristics (N=41, online supplemental material) that will be tested for interaction with CHARMS group in Stage 2 conversion rate. Second, the p values to test these prespecified characteristic associations to the risk of Stage 2 conversion will be presented using Bonferroni correction for multiple comparisons based on the number of variables in the list (ie, 41, regardless of any data collection issues that may emerge during the study execution). Third, due to the data driven nature of the audio processing approach, other unplanned analyses on characteristics yet to be defined are expected to be performed; such analyses will be explicitly labelled as exploratory, and the reader will be acknowledged in the result presentation (and trough this protocol) to carefully look at the findings merely as 'hypothesis generating'.

All analyses will be performed with R software (or equivalent statistical software) and uploaded in a public repository to guarantee the transparency and replicability of any finding.

Ethics and dissemination

The methodology described in this study adheres to ethical principles as formulated in the Declaration of Helsinki and is compatible with International Conference on Harmonization (ICH)-good clinical practice. The research protocol was reviewed and approved by two different Ethics Committees (CER Liguria approval code: 591/2020 – id.10993; Comitato Etico dell'Area Vasta Emilia Nord approval code: 2022/0071963). Participants will provide their written informed consent prior to study enrolment and parental consent will be needed in the case of participants aged less than 18 years old. Experimental results will be carefully shared through publication in peer-reviewed journals, to ensure proper data reproducibility.

Author affiliations
[1]IRCCS Ospedale Policlinico San Martino, Genoa, Italy
[2]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), Section of Psychiatry, University of Genoa, Genoa, Italy
[3]Department of Health Sciences (DISSAL), Section of Biostatistics, University of Genoa, Genoa, Italy
[4]Italian Natural Language Processing Lab, Institute of Computational Linguistics "Antonio Zampolli", CNR di Pisa, Pisa, Italy
[5]Department of Mental Health and Pathological Addictions, Piacenza Local Authority, Piacenza, Italy
[6]Department of Psychiatry, Faculty of Medicine, University of Geneva (UNIGE), Geneva, Switzerland
[7]Department of Psychiatry, Service of Adult Psychiatry (SPA), University Hospital of Geneva (HUG), Geneva, Switzerland
[8]Department of Mental Health and Pathological Addictions, Genoa Local Authority, Genoa, Liguria, Italy
[9]Departement of Psychology, School of Science and Technology, Middlesex University, London, UK
[10]Early Psychosis: Interventions and Clinical-Detection (EPIC) Lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, Kings College London, London, UK

Twitter Lisa Marzano @lisa_marzano

Acknowledgements This work was developed within the framework of the DINOGMI Department of Excellence of MIUR 2018-2022 (Law 232/2016).

Collaborators The LNG-PSY Study Investigators: Alessandra Costanza (Department of Psychiatry, Faculty of Medicine, University of Geneva (UNIGE), Geneva, Switzerland). Francesca Sibilla, Pietro Calcagno (Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), Section of Psychiatry, University of Genoa, Genoa, Italy. IRCCS Ospedale Policlinico San Martino, Genoa, Italy). Sara Patti, Gabriella Molino (Department of Mental Health and Pathological Addictions, Genoa Local Health Authority, Genoa, Italy). Andrea Escelsior (Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), Section of Psychiatry, University of Genoa, Genoa, Italy. IRCCS Ospedale Policlinico San Martino, Genoa, Italy). Alice Trabucco (Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), Section of Psychiatry, University of Genoa, Genoa, Italy). Lisa Marzano (Department of Psychology, School of Science and Technology, Middlesex University, London, UK). Dominique Brunato, Andrea Amelio Ravelli (Italian Natural Language Processing Lab, Institute of Computational Linguistics 'Antonio Zampolli' (ILC-CNR), Pisa, Italy). Marco Cappucciati (Department of Mental Health and Pathological Addictions, Piacenza Local Health Authority, Piacenza, Italy. Early Psychosis: Interventions and Clinical-detection (EPIC) lab, Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK). Roberta Fiocchi, Gisella Guerzoni, Davide Maravita, Fabio Macchetti, Elisa Mori, Chiara Anna Paglia, Federica Roscigno, Antonio Saginario (Department of Mental Health and Pathological Addictions, Piacenza Local Health Authority, Piacenza, Italy).

Contributors The authors have all contributed to the manuscript equally. LMag and AAm conceptualised and designed the study. LMag, EMon, IT, AT and DMart wrote the first draft of the protocol. OB, SC, MI and GL carried out a first methodological integration to optimise the experimental design. GM, SP, PC, FS, MC, RF, GG, DMara, FM, EMor, CAP, FR and AS further corrected the design to adapt it to different experimental settings, included in the multicentric context. LC conceptualised the planning of statistical analyses. FD, DB and AAR from Italian Natural Language Processing Lab provided a detailed project for linguistic data extraction and analysis. MA, GS, LMag, LMar, AC, AAg and AE carefully revised and approved the final version of the manuscript. Furthermore, a special thanks goes to the patient's primary advisers for their contribution in reporting enrollable subjects.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Luca Magnani http://orcid.org/0000-0002-3620-2021
Alessandra Costanza http://orcid.org/0000-0001-6387-6462
Lisa Marzano http://orcid.org/0000-0001-9735-3512

REFERENCES

1 Bleuler E. *Dementia praecox, oder Gruppe der Schizophrenien*. 1911.
2 Hauser MD. *The evolution of communication*. Cambridge, MA: MIT Press, 1996: 760.
3 Penn DC, Holyoak KJ, Povinelli DJ. Darwin's mistake: explaining the discontinuity between human and nonhuman minds. *Behav Brain Sci* 2008;31:109–30.
4 Tomasello M. *Origins of human communication*. Cambridge, MA: MIT Press, 2010: 408.
5 Vouloumanos A, Waxman SR. Listen up! speech is for thinking during infancy. *Trends Cogn Sci* 2014;18:642–6.
6 Eigsti I-M, de Marchena AB, Schuh JM, *et al*. Language acquisition in autism spectrum disorders: a developmental review. *Res Autism Spectr Disord* 2011;5:681–91.
7 Schaller S, Sacks O. *A man without words*. Berkeley: University of California Press, 1991: 204.
8 Humphries T, Kushalnagar P, Mathur G, *et al*. Ensuring language acquisition for deaf children: what linguists can do. *Language* 2014;90:e31–52.
9 Tattersall I. *The world from beginnings to 4000 BCE*. Oxford: Oxford University Press, 2008.
10 Tattersall I. An evolutionary context for the emergence of language. *Language Sciences* 2014;46:199–206.
11 Heidegger M, Krell DF. Basic writings: from being and time (1927) to the task of thinking; 1964.
12 Cassirer E. *An essay on man an introduction to a philosophy of human culture*. Yale University Press, 1944.
13 Rao RPN, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat Neurosci* 1999;2:79–87.
14 Helmholtz H. *Helmholtz's treatise on physiological optics*. Dover Publications, 1962.
15 Lupyan G, Clark A. Words and the world: predictive coding and the language-perception-cognition interface. *Curr Dir Psychol Sci* 2015;24:279–84.
16 Kessler RC, Berglund P, Demler O, *et al*. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National comorbidity survey replication. *Arch Gen Psychiatry* 2005;62:593.
17 Jones PB. Adult mental health disorders and their age at onset. *Br J Psychiatry* 2013;202:s5–10.
18 Potash JB. Carving chaos: genetics and the classification of mood and psychotic syndromes. *Harv Rev Psychiatry* 2006;14:47–63.
19 Potash JB, Bienvenu OJ. Neuropsychiatric disorders: shared genetics of bipolar disorder and schizophrenia. *Nat Rev Neurol* 2009;5:299–300.
20 Ivleva EI, Morris DW, Moates AF, *et al*. Genetics and intermediate phenotypes of the schizophrenia--bipolar disorder boundary. *Neurosci Biobehav Rev* 2010;34:897–921.
21 McGorry PD, Hickie IB, Yung AR, *et al*. Clinical staging of psychiatric disorders: a heuristic framework for choosing earlier, safer and more effective interventions. *Aust N Z J Psychiatry* 2006;40:616–22.
22 McGorry P. Transition to adulthood: the critical period for pre-emptive, disease-modifying care for schizophrenia and related disorders. *Schizophr Bull* 2011;37:524–30.

23 Kessler RC, Birnbaum H, Demler O, *et al*. The prevalence and correlates of nonaffective psychosis in the national comorbidity survey replication (NCS-R). *Biol Psychiatry* 2005;58:668–76.
24 Kessler RC, Ormel J, Petukhova M, *et al*. Development of lifetime comorbidity in the world health organization world mental health surveys. *Arch Gen Psychiatry* 2011;68:90–100.
25 Merikangas KR, Herrell R, Swendsen J, *et al*. Specificity of bipolar spectrum conditions in the comorbidity of mood and substance use disorders. *Arch Gen Psychiatry* 2008;65:47.
26 Merikangas KR, He J-P, Burstein M, *et al*. Lifetime prevalence of mental disorders in U.S. adolescents: results from the National comorbidity survey replication -- adolescent supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry* 2010;49:980–9.
27 Merikangas KR, Cui L, Kattan G, *et al*. Mania with and without depression in a community sample of US adolescents. *Arch Gen Psychiatry* 2012;69:943–51.
28 Murray GK, Jones PB. Psychotic symptoms in young people without psychotic illness: mechanisms and meaning. *Br J Psychiatry* 2012;201:4–6.
29 Ormel J, Raven D, van Oort F, *et al*. Mental health in Dutch adolescents: a trails report on prevalence, severity, age of onset, continuity and co-morbidity of DSM disorders. *Psychol Med* 2015;45:345–60.
30 Yung AR, McGorry PD. The prodromal phase of first-episode psychosis: past and current conceptualizations. *Schizophr Bull* 1996;22:353–70.
31 Malla AK, Norman RMG. Prodromal symptoms in schizophrenia. *Br J Psychiatry* 1994;164:487–93.
32 McGorry PD, Mei C. Ultra-high-risk paradigm: lessons learnt and new directions. *Evid Based Ment Health* 2018;21:131–3.
33 Lin A, Wood SJ, Nelson B, *et al*. Outcomes of nontransitioned cases in a sample at ultra-high risk for psychosis. *Am J Psychiatry* 2015;172:249–58.
34 Rutigliano G, Valmaggia L, Landi P, *et al*. Persistence or recurrence of non-psychotic comorbid mental disorders associated with 6-year poor functional outcomes in patients at ultra high risk for psychosis. *J Affect Disord* 2016;203:101–10.
35 Beck K, Andreou C, Studerus E, *et al*. Clinical and functional long-term outcome of patients at clinical high risk (CHR) for psychosis without transition to psychosis: a systematic review. *Schizophr Res* 2019;210:39–47.
36 Hartmann JA, Nelson B, Spooner R, *et al*. Broad clinical high-risk mental state (charms): methodology of a cohort study validating criteria for pluripotent risk. *Early Interv Psychiatry* 2019;13:379–86.
37 Shah JL, Scott J, McGorry PD, *et al*. Transdiagnostic clinical staging in youth mental health: a first international consensus statement. *World Psychiatry* 2020;19:233–42.
38 Yung AR, Yuen HP, McGorry PD, *et al*. Mapping the onset of psychosis: the comprehensive assessment of at-risk mental states. *Aust N Z J Psychiatry* 2005;39:964–71.
39 Hartmann JA, McGorry PD, Destree L, *et al*. Pluripotential risk and clinical staging: theoretical considerations and preliminary data from a transdiagnostic risk identification approach. *Front Psychiatry* 2020;11:553578.
40 Harrow M, Quinlan D. Is disordered thinking unique to schizophrenia? *Arch Gen Psychiatry* 1977;34:15–21.
41 Andreasen NC, Grove WM. Thought, language, and communication in schizophrenia: diagnosis and prognosis. *Schizophr Bull* 1986;12:348–59.
42 Corcoran CM, Mittal VA, Bearden CE, *et al*. Language as a biomarker for psychosis: a natural language processing approach. *Schizophr Res* 2020;226:158–66.
43 Voleti R, Liss JM, Berisha V. A review of automated speech and language features for assessment of cognitive and thought disorders. *IEEE J Sel Top Signal Process* 2020;14:282–98.
44 Covington MA, McFall JD. Cutting the gordian knot: the moving-average type–token ratio (MATTR). *J Quant Linguist* 2010;17:94–100.
45 Asgari M, Kaye J, Dodge H. Predicting mild cognitive impairment from spontaneous spoken utterances. *Alzheimers Dement (N Y)* 2017;3:219–28.
46 Roark B, Mitchell M, Hollingshead K. Syntactic complexity measures for detecting mild cognitive impairment - ACL anthology. *BioNLP* 2007:1–8.
47 Bucks RS, Singh S, Cuerden JM, *et al*. Analysis of spontaneous, conversational speech in dementia of Alzheimer type: evaluation of an objective technique for analysing lexical performance. *Aphasiology* 2000;14:71–91.
48 Fraser KC, Meltzer JA, Rudzicz F. Linguistic features identify Alzheimer's disease in narrative speech. *J Alzheimers Dis* 2016;49:407–22.
49 Sadeghian R, Schaffer JD, Zahorian SA. Speech processing approach for diagnosing dementia in an early stage. INTERSPEECH 2017; ISCA, 2017:2705–9
50 Mota NB, Vasconcelos NAP, Lemos N, *et al*. Speech graphs provide a quantitative measure of thought disorder in psychosis. *PLoS One* 2012;7:e34928.
51 Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. *Discourse Processes* 1998;25:259–84.
52 Mikolov T, Chen K, Corrado G, *et al*. *Efficient estimation of word representations in vector space*. ICLR, 2013.
53 Pennington J, Socher R, Manning CD. Glove: global vectors for word representation. EMNLP 2014 - 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference; 2014:1532–43
54 Elvevåg B, Foltz PW, Weinberger DR, *et al*. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophr Res* 2007;93:304–16.
55 Covington MA, He C, Brown C, *et al*. Schizophrenia and the structure of language: the linguist's view. *Schizophr Res* 2005;77:85–98.
56 Bedi G, Carrillo F, Cecchi GA, *et al*. Automated analysis of free speech predicts psychosis onset in high-risk youths. *NPJ Schizophr* 2015;1:15030.
57 Corcoran CM, Carrillo F, Fernández-Slezak D, *et al*. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry* 2018;17:67–75.
58 Rezaii N, Walker E, Wolff P. A machine learning approach to predicting psychosis using semantic density and latent content analysis. *NPJ Schizophr* 2019;5:9.
59 Morgan SE, Diederen K, Vértes PE, *et al*. Natural language processing markers in first episode psychosis and people at clinical high-risk. *Transl Psychiatry* 2021;11:630.
60 Pelizza L, Paterlini F, Azzali S, *et al*. The approved Italian version of the comprehensive assessment of at-risk mental states (CAARMS-ITA): field test and psychometric features. *Early Interv Psychiatry* 2019;13:810–7.
61 First MB, Williams JB, Karg RS, *et al*. Structured clinical interview for DSM-5—research version (SCID-5 for DSM-5, research version; SCID-5-RV) [Preprint at]. Arlington, VA: American Psychiatric Association, 2015: 1–94.
62 First MB, Williams JBW, Benjamin LS, *et al*. Structured clinical interview for DSM-5 personality disorders [Preprint at]. 2016.
63 Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. *Am J Psychiatry* 1992;149:1148–56.
64 Auther A, Smith C, Cornblatt B. Global functioning: social scale (GF: social) [Preprint at]. 2006.
65 Niendam TA, Bearden CE, Johnson JK, *et al*. Global functioning: role scale (GF: role) [Preprint at]. 2006.
66 Rush AJ, Trivedi MH, Ibrahim HM, *et al*. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 2003;54:573–83.
67 Young RC, Biggs JT, Ziegler VE, *et al*. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* 1978;133:429–35.
68 Lam RW, Michalaak EE, Swinson RP. *Assessment scales in depression, mania, and anxiety*. Taylor & Francis, 2005.
69 Palma A, Pancheri P. Scale di valutazione e di misura dei sintomi psichiatrici. In: Masson Italia M, ed. *Trattato Italiano di Psichiatria*. 1999.
70 Lovibond SH, Lovibond PF. Manual for the depression anxiety stress scales; 1995.
71 Patrick J, Dyck M, Bramston P. Depression anxiety stress scale: is it valid for children and adolescents? *J Clin Psychol* 2010;66:996–1007.
72 Nassir Ghaemi S, Miller CJ, Berv DA, *et al*. Sensitivity and specificity of a new bipolar spectrum diagnostic scale. *J Affect Disord* 2005;84:273–7.
73 Baldassano CF. Assessment tools for screening and monitoring bipolar disorder. *Bipolar Disord* 2005;7 Suppl 1:8–15.
74 Krueger RF, Derringer J, Markon KE, *et al*. Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychol Med* 2012;42:1879–90.
75 Anderson JL, Sellbom M, Salekin RT. Utility of the personality inventory for DSM-5-brief form (PID-5-BF) in the measurement of maladaptive personality and psychopathology. *Assessment* 2018;25:596–607.
76 Fossati A, Somma A, Borroni S, *et al*. A head-to-head comparison of the personality inventory for DSM-5 (PID-5) with the personality diagnostic questionnaire-4 (PDQ-4) in predicting the general level of personality pathology among community dwelling subjects. *J Pers Disord* 2016;30:82–94.

77 van der Gaag M, Schütz C, Ten Napel A, *et al*. Development of the davos assessment of cognitive biases scale (DACOBS). *Schizophr Res* 2013;144:63–71.
78 Bastiaens T, Claes L, Smits D, *et al*. The cognitive biases questionnaire for psychosis (CBQ-P) and the davos assessment of cognitive biases (DACOBS): validation in a flemish sample of psychotic patients and healthy controls. *Schizophr Res* 2013;147:310–4.
79 Roenneberg T, Wirz-Justice A, Merrow M. Life between clocks: daily temporal patterns of human chronotypes. *J Biol Rhythms* 2003;18:80–90.
80 Bastien CH, Vallières A, Morin CM. Validation of the insomnia severity index as an outcome measure for insomnia research. *Sleep Med* 2001;2:297–307.
81 Morin CM, Belleville G, Bélanger L, *et al*. The insomnia severity index: psychometric indicators to detect insomnia cases and evaluate treatment response. *Sleep* 2011;34:601–8.
82 Chung KF, Kan KKK, Yeung WF. Assessing insomnia in adolescents: comparison of insomnia severity index, Athens insomnia scale and sleep quality index. *Sleep Med* 2011;12:463–70.
83 Castronovo V, Galbiati A, Marelli S, *et al*. Validation study of the Italian version of the insomnia severity index (ISI). *Neurol Sci* 2016;37:1517–24.
84 Nasreddine ZS, Phillips NA, Bédirian V, *et al*. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;53:695–9.
85 Pirrotta F, Timpano F, Bonanno L, *et al*. Italian validation of montreal cognitive assessment. *Eur J Psychol Assess* 2015;31:131–7.
86 Ben-David S, Birnbaum ML, Eilenberg ME, *et al*. The subjective experience of youths at clinically high risk of psychosis: a qualitative study. *Psychiatr Serv* 2014;65:1499–501.
87 Brunato D, Cimino A, Dell'Orletta F, *et al*. Profiling-UD: a tool for linguistic profiling of texts. In: ACL Anthology. 2020: 7145–51.
88 Straka M, Hajič J, Straková J. UDPipe: trainable pipeline for processing CoNLL-U files performing tokenization, morphological analysis, POS tagging and parsing - ACL anthology. In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. 2016: 4290–7.
89 Nivre J, de Marneffe M-C, Ginter F, *et al*. Universal dependencies v1: a multilingual treebank collection; 2016.
90 Lu X. Automatic measurement of syntactic complexity in child language acquisition. *IJCL* 2009;14:3–28.
91 Lubetich S, Sagae K. Data-driven measurement of child language development with simple syntactic templates - ACL anthology. In: *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*. 2014: 2151–60.
92 Weiss Z, Meurers D. Analyzing linguistic complexity and accuracy in academic language development of German across elementary and secondary school. In: *ACL 2019 - Innovative Use of NLP for Building Educational Applications, BEA 2019 - Proceedings of the 14th Workshop*. 2019: 380–93.
93 Miaschi A, Brunato D, Dell'Orletta F, *et al*. What makes my model perplexed? A linguistic investigation on neural language models perplexity [Online]. *Proceedings of Deep Learning Inside Out (DeeLIO): The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures* 2021:40–7.
94 Prud'hommeaux E. T, Roark B, Black L. M, *et al*. Classification of atypical language in autism. *Proceedings of the 2nd Workshop on Cognitive Modeling and Computational Linguistics* 2011:88–96.
95 Vaswani A, Shazeer N, Parmer N, *et al*. Attention is all you need. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Red Hook, NY, USA: Curran Associates Inc, 2017: 6000–10. Available: https://papers.nips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
96 Devlin J, Chang M-W, Lee K, *et al*. Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North* 2019:4171–86.
97 Costanza A, Amerio A, Aguglia A, *et al*. "Hard to say, hard to understand, hard to live": possible associations between neurologic language impairments and suicide risk. *Brain Sci* 2021;11:1594.