RESEARCH PROPOSAL

Vocal Biomarkers for Subclinical Psychotic Symptoms in Autoimmune Thyroiditis: A Neuroinflammatory Digital Phenotyping Study

GRANT APPLICATION SUMMARY

Project Title: Vocal Biomarkers for Subclinical Psychotic Symptoms in Autoimmune Thyroiditis: A Neuroinflammatory Digital Phenotyping Study

Principal Investigator: [Name, Title, Institution]

Co-Investigators: [Names, Titles, Institutions]

Funding Period Requested: 36 months

Total Budget Requested: [Amount]

Project Start Date: [Date]

TABLE OF CONTENTS

6. Vertebrate Animals (Not Applicable)
7. Select Agent Research (Not Applicable)
8. Multiple PI Leadership Plan (if applicable)
9. Consortium/Contractual Arrangements (if applicable)
10. Letters of Support (if applicable)
11. Resource Sharing Plan

---

1. SPECIFIC AIMS

Autoimmune thyroiditis (Hashimoto's thyroiditis) affects 5-10% of the population with a 10:1 female predominance and is associated with diverse neuropsychiatric manifestations ranging from subtle cognitive changes to overt Hashimoto's encephalopathy.1-3 Emerging evidence suggests these neuropsychiatric symptoms represent a spectrum of **neuroautoimmune encephalitis** mediated by cross-reactive antibodies, cytokine dysregulation, and neuroinflammation.4-6 Subclinical psychotic symptoms—including attenuated hallucinations, unusual thought content, and thought disorganization—are frequently observed clinically but remain systematically understudied in this population.7,8

Critical gaps in current knowledge include: (1) lack of systematic characterization of subclinical psychotic symptoms in autoimmune thyroiditis across age groups; (2) absence of objective biomarkers for early detection, leading to delayed diagnosis and treatment; (3) gender bias in psychiatric assessment, with females' psychotic symptoms frequently underdiagnosed or misattributed;9,10 and (4) limited understanding of the relationship between neuroinflammation and psychiatric symptoms in imaging-negative encephalitis.11,12

Vocal biomarkers offer a promising solution. Speech production integrates multiple neural systems affected by neuroinflammation, including prefrontal-subcortical circuits, language networks, and dopaminergic pathways.13-15 Recent advances in computational speech analysis and natural language processing have demonstrated that acoustic and linguistic features can detect psychotic symptoms with high accuracy,16-18 provide objective assessment independent of clinician bias, and enable scalable, remote monitoring.19,20

Our central hypothesis is that neuroinflammation in autoimmune thyroiditis produces detectable alterations in vocal acoustic and linguistic features that correlate with subclinical psychotic symptoms, even in the absence of structural brain imaging abnormalities. We further hypothesize that these vocal biomarkers will respond to anti-inflammatory treatment (corticosteroids), validating their utility as state markers of neuroinflammatory psychiatric disease.

SPECIFIC AIMS

AIM 1: Characterize the prevalence, phenomenology, and clinical correlates of subclinical psychotic symptoms in pediatric and adult patients with autoimmune thyroiditis

Hypothesis 1a: Subclinical psychotic symptoms will be present in 15-25% of patients with autoimmune thyroiditis, with distinct symptom profiles (positive-predominant, negative-predominant, disorganized, mixed).

Hypothesis 1b: Symptom severity will correlate with inflammatory markers (IL-6, anti-TPO antibody titers, CRP) and cognitive dysfunction, independent of thyroid hormone levels.

Approach: We will conduct comprehensive psychiatric phenotyping using gold-standard instruments (Structured Interview for Psychosis-Risk Syndromes [SIPS], SCID-5) and neurocognitive assessment in 350 patients (175 pediatric [ages 10-17], 175 adult [ages 18-65]) with confirmed autoimmune thyroiditis. Clinical phenotypes will be classified by symptom profile, severity, and inflammatory status.

AIM 2: Identify vocal acoustic and linguistic biomarkers that differentiate patients with subclinical psychotic symptoms from those without, and validate these biomarkers against clinical phenotypes

Hypothesis 2a: Patients with subclinical psychotic symptoms will demonstrate specific vocal abnormalities including reduced semantic coherence, decreased prosodic variation, altered speech timing, and simplified syntax compared to patients without symptoms.

Hypothesis 2b: Vocal biomarkers will show symptom-specific patterns: positive symptoms will correlate with semantic abnormalities; negative symptoms with prosodic flattening; disorganization with reduced syntactic complexity and discourse coherence.

Hypothesis 2c: Vocal biomarkers will demonstrate concordance with clinical assessment ($\kappa \geq 0.60$) while detecting additional subclinical cases missed by traditional evaluation.

Approach: Standardized vocal recordings will undergo automated feature extraction (acoustic, linguistic, discourse-level). Machine learning models will classify symptom presence, severity, and profile. Clinical-digital concordance will be systematically evaluated, with detailed analysis of discordant cases.

AIM 3: Assess the responsiveness of vocal biomarkers to corticosteroid treatment in a prospective intervention substudy, validating their utility as state markers of neuroinflammatory psychiatric disease

Hypothesis 3a: Patients with subclinical psychotic symptoms who respond to corticosteroid treatment will show significant improvement in vocal biomarkers paralleling clinical improvement.

Hypothesis 3b: Baseline vocal biomarker profiles will predict treatment response, with specific features (e.g., prosodic abnormalities, semantic tangentiality) showing differential sensitivity to anti-inflammatory intervention.

Hypothesis 3c: Vocal biomarker changes will correlate with reductions in inflammatory markers (IL-6, CRP) and improvements in cognitive function.

Approach: A subset of 60 patients with clinically significant symptoms will undergo an 8-week open-label corticosteroid trial with assessments at baseline, weeks 2, 4, 8, and 12 (post-taper). Longitudinal trajectories of clinical symptoms, vocal biomarkers, inflammatory markers, and cognitive function will be analyzed using mixed-effects models.

EXPECTED OUTCOMES AND IMPACT

This study will provide: (1) the first systematic characterization of subclinical psychotic symptoms in autoimmune thyroiditis; (2) validated vocal biomarkers for objective, gender-unbiased detection of neuroinflammatory psychiatric symptoms; (3) evidence linking vocal features to neuroinflammation and treatment response; and (4) a scalable digital phenotyping framework applicable to other autoimmune conditions.

Clinical impact includes earlier detection of at-risk patients, reduced diagnostic bias, objective monitoring of treatment response, and potential for remote screening applications. **Scientific impact** includes advancing understanding of the brain-immune interface, establishing vocal biomarkers as functional indicators of neuroinflammation, and providing a methodological framework for digital phenotyping in autoimmune neuropsychiatric disorders.

---

2. RESEARCH STRATEGY

A. SIGNIFICANCE

1. Public Health Burden and Unmet Clinical Need

Autoimmune thyroiditis is one of the most common autoimmune conditions, affecting approximately 5-10% of the general population, with prevalence reaching 15-20% in women over age 60.21,22 While traditionally viewed as an endocrine disorder, accumulating evidence demonstrates significant neuropsychiatric comorbidity including depression (30-60%), anxiety (30-50%), and cognitive impairment (20-40%).23-25

Hashimoto's encephalopathy, the severe end of the neuropsychiatric spectrum, occurs in approximately 2-5 per 100,000 population annually,26 presenting with seizures, cognitive decline, psychiatric symptoms, and movement disorders. However, this likely represents only the "tip of the iceberg," with many patients experiencing **subclinical or mild neuropsychiatric manifestations** that go unrecognized and untreated.27,28

Subclinical psychotic symptoms—attenuated hallucinations, unusual thought content, suspiciousness, and thought disorganization—are particularly concerning because they:

- Significantly impair quality of life and functional capacity29
- Increase risk for progression to threshold psychotic disorders30,31
- Are frequently missed or misdiagnosed, especially in females32,33
- May respond to anti-inflammatory treatment if identified early34,35

Despite clinical observations of psychotic symptoms in autoimmune thyroiditis,36,37 **no systematic studies** have characterized their prevalence, phenomenology, or underlying mechanisms in this population.

2. Neuroinflammatory Mechanisms: The Missing Link

Recent research has fundamentally shifted understanding of Hashimoto's encephalopathy from a rare, idiosyncratic reaction to a **neuroinflammatory disorder** on a continuum with milder presentations.38-40

Pathophysiological evidence includes:

Direct autoimmune mechanisms:

- Anti-thyroid antibodies (anti-TPO, anti-thyroglobulin) cross-react with neuronal antigens, particularly in cerebellum, hippocampus, and cortex41,42
- Anti-neuronal antibodies (anti-NAE, anti-GAD) are detected in 60-80% of Hashimoto's encephalopathy cases43,44
- Antibody-mediated complement activation causes neuronal damage45
- Blood-brain barrier disruption allows antibody penetration and immune cell infiltration46

Indirect inflammatory mechanisms:

- Elevated pro-inflammatory cytokines (IL-6, IL-1β, TNF-α, IFN-γ) affect neurotransmitter metabolism and synaptic function47,48
- Chronic microglial activation disrupts neural circuits49
- Oxidative stress and mitochondrial dysfunction impair neuronal function50
- Inflammation-induced alterations in dopamine, serotonin, and glutamate systems parallel mechanisms in primary psychotic disorders51,52

Critical insight: Neuroinflammation produces **functional neuronal dysfunction** before structural damage, explaining why 30-50% of autoimmune encephalitis cases show **normal conventional MRI** despite significant symptoms.53,54 Advanced imaging (FDG-PET, fMRI) may reveal metabolic or functional abnormalities,55 but these techniques are expensive and not widely available.

Corticosteroid responsiveness provides both diagnostic and therapeutic validation. Studies report 60-90% response rates in Hashimoto's encephalopathy,56,57 with improvement typically within 2-4 weeks, supporting neuroinflammatory etiology. However, response in milder, subclinical presentations remains unstudied.

3. Gender Bias in Psychiatric Assessment: A Critical Equity Issue

Psychotic symptoms in females are systematically underdiagnosed due to multiple factors:58-60

- Atypical presentations (more affective symptoms, less disorganization)
- Later age of onset allowing development of compensatory strategies
- Clinician bias toward diagnosing mood disorders in women
- Self-report bias and help-seeking differences

In autoimmune thyroiditis, with 90% female predominance, this gender bias is particularly problematic. **Objective biomarkers** that operate independently of clinician interpretation could significantly reduce diagnostic disparities.

4. Vocal Biomarkers: A Paradigm Shift in Psychiatric Assessment

Speech production integrates multiple neural systems affected by neuroinflammation:

- **Prefrontal cortex:** Executive control of speech planning and monitoring61
- **Anterior cingulate:** Conflict monitoring and error detection62
- **Basal ganglia:** Dopaminergic motor planning affecting prosody63
- **Language networks:** Broca's area, Wernicke's area, arcuate fasciculus64
- **Default mode network:** Narrative construction and self-referential processing65

Vocal biomarkers in psychosis research have demonstrated:

Acoustic features:

- Reduced prosodic variation (flattened affect) in negative symptoms66,67

- Altered speech timing and increased pauses in thought disorder68
- Voice quality changes reflecting motor control disruption69

Linguistic features:

- Reduced semantic coherence predicts psychosis onset in high-risk youth16,17
- Syntactic simplification correlates with cognitive deficits70
- Increased tangentiality and derailment in formal thought disorder71,72

Predictive validity:

- Automated speech analysis predicts psychosis conversion with 83% accuracy17
- Semantic coherence outperforms clinical assessment for prediction16
- Longitudinal changes track symptom progression73

Advantages over traditional assessment:

- **Objective:** Reduces clinician and patient bias
- **Quantitative:** Enables dimensional measurement and precise tracking
- **Scalable:** Can be deployed remotely via smartphone
- **Ecologically valid:** Captures real-world functioning
- **Gender-independent:** Operates on acoustic/linguistic features, not subjective interpretation

5. Knowledge Gap and Study Rationale

No prior studies have:

1. Systematically characterized subclinical psychotic symptoms in autoimmune thyroiditis
2. Applied vocal biomarker analysis to autoimmune neuropsychiatric disorders
3. Examined vocal features as markers of neuroinflammation
4. Assessed vocal biomarker responsiveness to anti-inflammatory treatment
5. Compared clinical and digital phenotyping in this population

This study addresses these gaps, with potential to **transform clinical practice** by providing objective, scalable tools for early detection and monitoring of neuroinflammatory psychiatric disease.

B. INNOVATION

1. Novel Application of Digital Phenotyping to Autoimmune Neuropsychiatry

This is the **first study** to apply computational vocal analysis to neuropsychiatric manifestations of autoimmune disease. While vocal biomarkers have been studied in primary psychotic disorders,16-20,66-73 their application to **secondary psychosis** due to neuroinflammation represents a conceptual advance with broad implications for autoimmune encephalitis, neuroinflammation in general, and the brain-immune interface.

2. Imaging-Negative Neuroinflammation: A New Paradigm

By focusing on patients with **functional neuroinflammatory changes** that may not manifest on structural imaging, this study challenges the traditional requirement for imaging abnormalities in autoimmune encephalitis diagnosis.53,54 Vocal biomarkers may provide **functional evidence** of neuroinflammation when imaging is normal, enabling earlier diagnosis and treatment.

3. Treatment-Responsive Biomarkers: Validation Through Intervention

The corticosteroid intervention substudy provides **mechanistic validation** of vocal biomarkers as state markers of neuroinflammation. Demonstrating that vocal features normalize with anti-inflammatory treatment would:

- Confirm neuroinflammatory etiology of vocal abnormalities
- Validate biomarkers as treatment response indicators
- Enable objective monitoring in clinical trials
- Support use as surrogate endpoints

This approach is **innovative** because most biomarker studies are purely observational; we directly test whether biomarkers track with disease activity.

4. Comprehensive Clinical-Digital Phenotype Integration

Rather than treating digital biomarkers as replacements for clinical assessment, we systematically compare and integrate both approaches. This **hybrid framework**:

- Assesses concordance and discordance between methods
- Identifies complementary information from each approach
- Evaluates incremental validity of combined assessment
- Develops optimal clinical decision support strategies

This represents a **methodological innovation** applicable beyond this specific study.

5. Gender Equity Through Objective Assessment

By developing **gender-independent** biomarkers and explicitly testing for algorithmic bias, this study addresses health equity in psychiatric diagnosis. The systematic evaluation of gender-specific performance represents a **novel approach** to bias mitigation in digital health.

6. Developmental Perspective: Pediatric and Adult Cohorts

Parallel assessment of pediatric and adult populations enables examination of **developmental factors** in neuroinflammatory psychiatric manifestations. This is particularly important given that:

- Autoimmune thyroiditis can present in childhood/adolescence74
- Adolescence is a critical period for psychosis risk75
- Vocal features show developmental changes requiring age-appropriate normative data76

7. Multi-Modal Biomarker Integration

Integration of vocal biomarkers with inflammatory markers, cognitive assessment, and clinical phenotypes provides a **systems-level understanding** of neuroinflammatory psychiatric disease. This multi-modal approach is essential for:

- Understanding mechanisms linking inflammation to symptoms
- Identifying subgroups with different pathophysiology
- Developing personalized treatment strategies

8. Translational Pathway: From Discovery to Clinical Application

The study is designed with **clinical implementation** in mind:

- Standardized recording protocols feasible in clinical settings
- Open-source analysis tools for reproducibility
- Decision support framework for clinical integration
- Scalability to smartphone-based applications

This **translational focus** distinguishes the study from purely exploratory research.

C. APPROACH

1. STUDY DESIGN AND OVERVIEW

Design: Prospective, cross-sectional, multi-center observational study with nested case-control analysis and prospective corticosteroid intervention substudy

Study Duration: 36 months

- Months 1-6: Site setup, IRB approvals, staff training, protocol piloting
- Months 7-24: Active recruitment and data collection
- Months 25-30: Data analysis and model development
- Months 31-36: Validation, manuscript preparation, dissemination

Study Sites: 3-4 academic medical centers with expertise in endocrinology, psychiatry, and digital health

Target Enrollment: 350 participants (175 pediatric, 175 adult) for primary observational study; 60 participants (30 pediatric, 30 adult) for corticosteroid intervention substudy

2. STUDY POPULATION

2.1 Inclusion Criteria

All Participants:

- Confirmed diagnosis of autoimmune thyroiditis defined by:
    - Elevated anti-TPO antibodies (>35 IU/mL) AND/OR
    - Elevated anti-thyroglobulin antibodies (>40 IU/mL) AND
    - Compatible thyroid ultrasound findings (heterogeneous echotexture) if available
- Age 10-17 years (pediatric cohort) or 18-65 years (adult cohort)
- Thyroid status: Euthyroid (TSH 0.5-4.5 mIU/L), subclinical hypothyroid (TSH 4.5-10 mIU/L, normal FT4), or treated hypothyroid with stable TSH on consistent levothyroxine dose for ≥3 months
- Native or fluent speaker of English (or study language)
- Capacity to provide informed consent (adults) or assent with parental consent (minors)
- Able to complete study procedures including vocal recordings and cognitive testing

Additional Criteria for Corticosteroid Intervention Substudy:

- SIPS positive symptom score ≥6 (presence of subclinical psychotic symptoms)
- At least one SIPS positive item rated ≥3 (moderate severity)
- Symptoms present for ≥1 month but <2 years
- No contraindications to corticosteroid therapy
- Willingness to participate in 12-week intervention protocol

2.2 Exclusion Criteria

All Participants:

- Diagnosed psychotic disorder (schizophrenia, schizoaffective disorder, delusional disorder, brief psychotic disorder) or current antipsychotic medication
- Severe cognitive impairment (IQ <70 or diagnosed intellectual disability) preventing valid assessment
- Active substance use disorder within past 3 months (excluding tobacco, caffeine)
- Neurological conditions significantly affecting speech production:
    - Stroke with residual aphasia or dysarthria
    - Parkinson's disease or other movement disorders
    - Severe traumatic brain injury
    - Structural laryngeal/vocal cord pathology
- Current acute thyroid crisis or thyroid storm
- Severe uncontrolled thyroid dysfunction (TSH >10 mIU/L or <0.1 mIU/L)
- Pregnancy (for intervention substudy only)
- Non-English speaking (unless study expanded to other languages)

Additional Exclusion for Corticosteroid Intervention Substudy:

- Contraindications to corticosteroids:
    - Uncontrolled diabetes mellitus (HbA1c >9%)
    - Active peptic ulcer disease
    - Severe osteoporosis with fragility fractures
    - Active untreated infection
    - Live vaccine administration within past month
- Current immunosuppressive therapy
- Corticosteroid use within past 3 months
- Known hypersensitivity to corticosteroids

2.3 Sample Size Justification

Primary Observational Study (N=350):

Aim 1 (Prevalence and Characterization):

- Assuming 20% prevalence of subclinical psychotic symptoms (SIPS positive score ≥6)
- N=350 provides 95% CI of ±4.2% around prevalence estimate
- Expected n≈70 with symptoms, n≈280 without symptoms

Aim 2 (Vocal Biomarker Classification):

- **Power analysis** for binary classification (symptoms present vs. absent):
    - Assuming moderate effect size (Cohen's d=0.5) for vocal feature differences
    - α=0.05, power=0.80
    - Required n≈64 per group
    - With 20% prevalence, N=350 provides n≈70 with symptoms (adequate power)
    - Oversampling of symptomatic cases through enrichment strategy if needed
- **Machine learning sample size:**
    - Rule of thumb: 10-20 events per predictor variable
    - With ≈50 vocal features after dimensionality reduction
    - Minimum required n≈50-100 positive cases
    - N=350 with 20% prevalence provides n≈70 (adequate)

Subgroup analyses:

- Pediatric vs. adult: n≈175 per group provides 80% power to detect d=0.42
- Gender comparisons: With 80% female, n≈280 female, n≈70 male provides 80% power to detect d=0.48 for gender differences
- Symptom profile comparisons: With 4 profiles, expected n≈15-20 per profile (adequate for exploratory analysis)

Corticosteroid Intervention Substudy (N=60):

Aim 3 (Treatment Response):

- **Longitudinal power analysis:**
    - Assuming 70% response rate (based on Hashimoto's encephalopathy literature)56,57
    - Expected n≈42 responders, n≈18 non-responders
    - For within-subject change (baseline to week 8):
        - Assuming large effect size (d=0.8) for SIPS score change in responders
        - α=0.05, power=0.80
        - Required n≈15 responders (we expect n≈42)
    - For responder vs. non-responder vocal biomarker change:
        - Assuming moderate-large effect size (d=0.7)
        - Required n≈35 per group (we have n≈42 vs. 18, adequate for responders)
- **Attrition:** Assuming 20% dropout, recruit n=75 to achieve n=60 completers

Total recruitment target: N=425 (350 + 75 for intervention substudy)

2.4 Recruitment Strategy

Recruitment Sources:

- Endocrinology clinics (pediatric and adult)
- Primary care practices with electronic health record screening
- Patient registries and databases
- Community outreach through thyroid patient advocacy organizations
- Social media and online recruitment platforms
- Referrals from psychiatry and neurology clinics

Enrichment Strategy for Symptomatic Cases:

- Targeted recruitment of patients with documented psychiatric symptoms
- Screening questionnaire (PQ-16) administered to all potential participants

- Oversampling of screen-positive individuals to ensure adequate symptomatic cases

Retention Strategies:

- Flexible scheduling including evenings and weekends
- Remote assessment options when feasible
- Compensation for time and travel ($50-100 per visit)
- Regular communication and engagement
- Feedback of results to participants and referring providers (with consent)

3. CLINICAL ASSESSMENT PROTOCOL

3.1 Baseline Assessment (Week 0)

Session 1: Diagnostic and Psychiatric Assessment (2.5-3 hours)

Informed Consent and Demographics

- Comprehensive informed consent process
- Demographic data: age, gender, race/ethnicity, education, socioeconomic status
- Medical history: thyroid disease duration, diagnosis date, treatment history
- Medication review: levothyroxine dose, other medications
- Family history: thyroid disease, psychiatric disorders, autoimmune conditions

Structured Diagnostic Interview

- **Structured Clinical Interview for DSM-5, Research Version (SCID-5-RV)**77
    - Psychotic Disorders module (screening for exclusion)
    - Mood Disorders module (current and lifetime major depressive episodes, persistent depressive disorder)
    - Anxiety Disorders module (generalized anxiety, panic, social anxiety, specific phobias)
    - Trauma and Stressor-Related Disorders module (PTSD, acute stress disorder)

- Duration: 60-90 minutes
- Administered by trained master's or doctoral-level clinician

Psychosis-Risk Assessment

- **Structured Interview for Psychosis-Risk Syndromes (SIPS)**78,79
    - **Primary outcome measure** for Aims 1 and 3
    - Comprehensive assessment of:
        - **Positive symptoms** (5 items): Unusual thought content, suspiciousness, grandiosity, perceptual abnormalities, disorganized communication
        - **Negative symptoms** (6 items): Social anhedonia, avolition, expression of emotion, experience of emotions, ideational richness, occupational functioning
        - **Disorganization symptoms** (4 items): Odd behavior, bizarre thinking, trouble with focus, impaired hygiene
        - **General symptoms** (4 items): Sleep disturbance, dysphoric mood, motor disturbances, impaired stress tolerance
    - Each item rated 0-6 (0=absent, 1=questionable, 2=mild, 3=moderate, 4=moderately severe, 5=severe but not psychotic, 6=severe and psychotic)
    - Duration: 45-60 minutes
    - Requires certified SIPS rater (online training and certification: www.sipstraining.com)

Session 2: Neurocognitive Assessment (2-2.5 hours)

Cognitive Battery (adapted from MATRICS Consensus Cognitive Battery)80

Processing Speed:

- Trail Making Test Part A81
- Digit Symbol Coding (WAIS-IV)82 or age-appropriate equivalent

Attention and Working Memory:

- Digit Span Forward/Backward (WAIS-IV)82

- Letter-Number Sequencing (WAIS-IV)82

Executive Function:

- Trail Making Test Part B81
- Verbal Fluency (semantic: animals, fruits; phonemic: F, A, S)83
- Stroop Color-Word Test84

Verbal Learning and Memory:

- Hopkins Verbal Learning Test-Revised (HVLT-R)85

Visual Learning and Memory:

- Brief Visuospatial Memory Test-Revised (BVMT-R)86

Language and Reasoning:

- Boston Naming Test (15-item version)87
- Similarities (WAIS-IV)82
- Proverb Interpretation (3-5 proverbs)88

Social Cognition:

- Reading the Mind in the Eyes Test (RMET)89

Cognitive Domain Scoring:

- Generate age- and education-adjusted z-scores for each test
- Calculate composite scores for each cognitive domain
- Classify cognitive profiles (intact, executive-predominant, memory-predominant, generalized deficit, processing speed/attention deficit)

Session 3: Clinician-Rated Scales and Self-Report Measures (1-1.5 hours)

Clinician-Rated Symptom Severity

- **Hamilton Depression Rating Scale (HAM-D-17)**90 or **Montgomery-Åsberg Depression Rating Scale (MADRS)**91
- **Hamilton Anxiety Rating Scale (HAM-A)**92
- **Clinical Global Impression - Severity (CGI-S)**93

Self-Report Questionnaires

- **Prodromal Questionnaire-Brief Version (PQ-B)**94: Self-report screening for psychotic-like experiences
- **Community Assessment of Psychic Experiences (CAPE)**95: Dimensional assessment of psychotic experiences
- **Beck Depression Inventory-II (BDI-II)**96 or **Patient Health Questionnaire-9 (PHQ-9)**97
- **Beck Anxiety Inventory (BAI)**98 or **Generalized Anxiety Disorder-7 (GAD-7)**99
- **Childhood Trauma Questionnaire (CTQ)**100
- **WHO Quality of Life-BREF (WHOQOL-BREF)**101 or **Pediatric Quality of Life Inventory (PedsQL)**102 for children
- **Social Adjustment Scale (SAS)**103

Vocal Recording Protocol (see Section 3.3)

Biological Samples

- Blood draw for:
    - Thyroid function: TSH, Free T4, Free T3
    - Thyroid antibodies: Anti-TPO, anti-thyroglobulin (quantitative titers)
    - Inflammatory markers: IL-6, IL-1β, TNF-α, CRP, ESR
    - Complete blood count, comprehensive metabolic panel
    - Optional: Anti-neuronal antibodies (anti-NAE, anti-GAD, NMDA receptor antibodies) if available

3.2 Clinical Classification System

Primary Classification: Psychotic Symptom Severity (based on SIPS)

- **Category 0: No Subclinical Symptoms**
    - All SIPS positive items rated 0-1
    - Total SIPS positive score: 0-5
- **Category 1: Minimal Subclinical Symptoms**
    - At least one SIPS positive item rated 2 (mild)
    - Total SIPS positive score: 6-11
- **Category 2: Moderate Subclinical Symptoms**
    - At least one SIPS positive item rated 3-4 (moderate to moderately severe)
    - Total SIPS positive score: 12-17
- **Category 3: Severe Subclinical Symptoms**
    - At least one SIPS positive item rated 5 (severe but not psychotic) OR multiple items rated 4
    - Total SIPS positive score: ≥18
- **Category 4: Threshold Psychotic Symptoms** (exclusion or separate analysis)
    - At least one SIPS positive item rated 6 (psychotic intensity)

Secondary Classification: Symptom Profiles

- **Profile A: Positive-Predominant**
    - SIPS positive score ≥12 AND (positive score > negative + disorganization scores)
- **Profile B: Negative-Predominant**
    - SIPS negative score ≥12 AND (negative score > positive score)
- **Profile C: Disorganized-Predominant**
    - SIPS disorganization score ≥8 AND prominent P5 (disorganized communication ≥3)
- **Profile D: Mixed/Undifferentiated**
    - Significant symptoms across multiple domains without clear predominance

- **Profile E: Affective-Psychotic**
    - SIPS positive score ≥6 PLUS HAM-D ≥15 or HAM-A ≥15

Tertiary Classification: Cognitive Profiles

- **Profile 1: Intact** - All domains z > -1.0
- **Profile 2: Executive-Predominant** - Executive function z < -1.5
- **Profile 3: Memory-Predominant** - Memory z < -1.5
- **Profile 4: Generalized Deficit** - ≥3 domains z < -1.5
- **Profile 5: Processing Speed/Attention** - Selective impairment in speed/attention

Quaternary Classification: Inflammatory Phenotypes

- **High Inflammatory:** ≥2 elevated markers (IL-6 >5 pg/mL, CRP >3 mg/L, ESR >20 mm/hr, anti-TPO >1000 IU/mL)
- **Moderate Inflammatory:** 1 elevated marker
- **Low Inflammatory:** All markers within normal range

3.3 Vocal Recording Protocol

Recording Environment and Equipment

Standardized Recording Setup:

- Quiet room with controlled acoustics (ambient noise <40 dB)
- Consistent across all study sites
- Audio recorder: High-quality digital recorder or laptop with professional microphone
- Specifications: Minimum 44.1 kHz sampling rate, 16-bit resolution, uncompressed WAV format
- Microphone positioning: 15-20 cm from participant, consistent angle
- Real-time audio monitoring for quality control

Quality Assurance:

- Pre-session equipment check and calibration
- Test recording reviewed before main session
- Re-recording of poor-quality segments
- Metadata documentation: date, time, equipment ID, ambient noise level, any technical issues

Speech Tasks (Total duration: 25-30 minutes)

1. Free Speech Tasks (Naturalistic Language Production)

Task 1a: Daily Life Narrative (4-5 minutes)

- **Prompt:** "I'd like you to tell me about a typical day in your life. Start from when you wake up in the morning and describe what you do throughout the day. Include as much detail as you can."
- **Purpose:** Elicits naturalistic speech, narrative structure, self-referential content
- **Expected features:** Semantic coherence, topic maintenance, temporal organization

Task 1b: Picture Description (3 minutes)

- **Stimulus:** Complex scene picture (e.g., "Cookie Theft" picture from Boston Diagnostic Aphasia Examination)104
- **Prompt:** "Please describe everything you see happening in this picture in as much detail as possible."
- **Purpose:** Constrained content allows comparison across participants
- **Expected features:** Semantic density, referential coherence, descriptive richness

Task 1c: Recent Experience Narrative (4-5 minutes)

- **Prompt:** "Tell me about something interesting or memorable that happened to you recently. It could be something you did, somewhere you went, or someone you met. Please describe it in detail."
- **Purpose:** Episodic memory retrieval, emotional content, spontaneous organization
- **Expected features:** Narrative coherence, emotional expression, tangentiality

2. Constrained Speech Tasks (Controlled Language Production)

Task 2a: Passage Reading (2 minutes)

- **Stimulus:** "Rainbow Passage"105 (phonetically balanced, standardized)
- **Purpose:** Controls for content, isolates acoustic/prosodic features
- **Expected features:** Prosody, speech rate, voice quality, articulation

Task 2b: Sustained Vowel Phonation (1 minute)

- **Prompt:** Sustain vowels /a/, /i/, /u/ for 5 seconds each, 3 repetitions
- **Purpose:** Voice quality assessment, vocal stability
- **Expected features:** Jitter, shimmer, harmonics-to-noise ratio

Task 2c: Rapid Automatized Naming (1-2 minutes)

- **Stimuli:** Count from 1-20, recite days of the week, months of the year
- **Purpose:** Automatic speech production, motor planning
- **Expected features:** Speech rate, pause patterns, fluency

3. Cognitive-Linguistic Tasks (Executive Function and Language)

Task 3a: Verbal Fluency (3 minutes)

- **Semantic fluency:** "Name as many animals as you can in one minute" (then repeat with fruits)
- **Phonemic fluency:** "Name as many words as you can that start with the letter F" (then A, then S)
- **Purpose:** Executive function, semantic access, cognitive flexibility
- **Expected features:** Lexical diversity, clustering, switching, perseveration

Task 3b: Proverb Interpretation (4-5 minutes)

- **Stimuli:** 3-5 common proverbs (e.g., "Don't cry over spilled milk," "A rolling stone gathers no moss")
- **Prompt:** "What does this saying mean?"
- **Purpose:** Abstract reasoning, concrete vs. abstract thinking
- **Expected features:** Semantic complexity, tangentiality, thought organization

Task 3c: Similarities Task (3-4 minutes)

- **Stimuli:** 5 pairs (e.g., "How are an apple and an orange alike?")
- **Purpose:** Abstract verbal reasoning, categorical thinking
- **Expected features:** Semantic relationships, abstraction level, response organization

Recording Procedures:

1. **Participant preparation:**
    - Explanation of recording purpose and procedures
    - Practice with sample task to reduce anxiety
    - Encouragement to speak naturally, no "right or wrong" answers
    - Option for breaks between tasks
2. **Standardized instructions:**
    - Scripted prompts read verbatim by examiner
    - Minimal examiner speech during participant responses
    - Neutral encouragers only ("uh-huh," "okay") if participant pauses
    - Follow-up prompts if participant stops prematurely: "Can you tell me more?" (maximum 2 prompts)
3. **Task order:**
    - Fixed order for all participants to control for fatigue effects
    - Brief breaks between task blocks
4. **Documentation:**
    - Task completion checklist
    - Notes on participant engagement, effort, comprehension
    - Any deviations from protocol

4. VOCAL BIOMARKER ANALYSIS

4.1 Audio Preprocessing

Automated Pipeline:

1. **Noise reduction:** Spectral subtraction algorithm to remove ambient noise
2. **Normalization:** Amplitude normalization to standard level
3. **Voice Activity Detection (VAD):** Separate speech from silence/non-speech sounds using energy-based and spectral methods
4. **Segmentation:** Identify utterance boundaries, pauses, speech segments
5. **Quality assessment:** Flag recordings with excessive noise, clipping, or technical issues
6. **Examiner speech removal:** Identify and exclude examiner speech (for free speech tasks)

Software: Praat,106 OpenSMILE,107 custom Python scripts using librosa library

4.2 Speech Transcription and Annotation

Automated Speech Recognition (ASR):

- Initial transcription using state-of-the-art ASR (e.g., Whisper, Google Speech-to-Text)
- Word-level timestamps for alignment with audio

Manual Verification and Correction:

- Trained transcribers review and correct ASR output
- Transcription conventions:
    - Verbatim transcription including false starts, repetitions, filler words
    - Annotation of disfluencies: [pause], [unintelligible], [crosstalk]
    - Non-verbal vocalizations: [laugh], [sigh], [cough]
- Inter-rater reliability: 20% of transcripts double-coded, target $\kappa \geq 0.80$

Linguistic Annotation:

- Part-of-speech tagging using spaCy or NLTK
- Syntactic parsing for sentence structure analysis
- Semantic annotation for specific analyses

4.3 Feature Extraction

Acoustic Features (extracted using OpenSMILE107 and Praat106)

Prosodic Features:

- **Fundamental frequency (F0):**
    - Mean, standard deviation, range, minimum, maximum
    - Coefficient of variation (CV)
    - Slope of F0 contour
    - Number of inflection points
- **Intensity:**
    - Mean, standard deviation, range
    - Dynamic range (max-min)
- **Speech timing:**
    - Speech rate (syllables per second)
    - Articulation rate (excluding pauses)
    - Pause frequency (pauses per minute)
    - Mean pause duration, pause duration variability
    - Proportion of time spent in pauses
    - Within-utterance vs. between-utterance pause ratio

Voice Quality Features:

- **Jitter:** Cycle-to-cycle variation in F0 (local, absolute)
- **Shimmer:** Cycle-to-cycle variation in amplitude (local, dB)
- **Harmonics-to-Noise Ratio (HNR):** Measure of voice periodicity
- **Spectral features:**

- Formant frequencies (F1, F2, F3, F4) - mean and variability
- Spectral centroid, spectral roll-off, spectral flux
- Mel-frequency cepstral coefficients (MFCCs): 13-20 coefficients
- Delta and delta-delta MFCCs (temporal derivatives)

Linguistic Features (extracted using custom NLP pipelines)

Lexical Features:

- **Lexical diversity:**
  - Type-token ratio (TTR)
  - Moving average type-token ratio (MATTR)108
  - Measure of Textual Lexical Diversity (MTLD)109
- **Word frequency:** Mean word frequency based on SUBTLEX corpus110
- **Word length:** Mean word length in characters and syllables
- **Content vs. function words:** Proportion of nouns, verbs, adjectives vs. pronouns, prepositions, conjunctions

Semantic Features:

- **Semantic coherence:**
  - Latent Semantic Analysis (LSA)111 between consecutive sentences
  - Mean coherence, coherence variability
  - Coherence decline over narrative
- **Topic maintenance:**
  - Semantic distance from initial prompt/topic
  - Topic drift measures
- **Tangentiality:**
  - Semantic distance between consecutive utterances
  - Proportion of off-topic utterances
- **Self-referential content:**
  - First-person pronoun frequency (I, me, my, mine)
  - Self-reference ratio

Syntactic Features:

- **Syntactic complexity:**
    - Mean length of utterance (MLU) in words
    - Mean length of T-unit (independent clause + dependent clauses)
    - Dependent clauses per sentence
    - Embedding depth
    - Yngve depth (left-branching complexity)112
- **Grammatical errors:**
    - Frequency and types of errors
    - Subject-verb agreement errors
    - Tense inconsistencies

Discourse Features:

- **Referential coherence:**
    - Pronoun use and clarity of anaphoric reference
    - Referential ambiguity
- **Narrative structure:**
    - Presence of temporal markers (then, next, after)
    - Logical connectors (because, therefore, but)
    - Narrative completeness (beginning, middle, end)
- **Disfluencies:**
    - Filled pauses (um, uh)
    - Repetitions, false starts, self-corrections
    - Incomplete sentences

Natural Language Processing (NLP) Features:

- **Sentiment analysis:** Emotional valence and arousal using VADER113 or similar
- **Thought disorder indices:**
    - Automated measures based on Corcoran et al.16 and Bedi et al.17
    - Semantic incoherence scores

- **Perplexity:** Language model perplexity indicating predictability of word sequences
- **Semantic density:** Information content per utterance using word embeddings (Word2Vec, GloVe, BERT)

Feature Aggregation:

- Task-specific features (e.g., free speech vs. reading)
- Aggregate features across all tasks
- Feature stability across tasks (consistency measures)

Total feature space: Approximately 200-300 features before dimensionality reduction

4.4 Machine Learning Classification

Data Preparation:

Train-Test Split:

- 70% training, 30% hold-out test set
- Stratified by symptom presence, age group, gender
- Test set used only for final model evaluation

Feature Preprocessing:

- Standardization (z-score normalization)
- Handling missing values (imputation or exclusion)
- Outlier detection and treatment

Feature Selection and Dimensionality Reduction:

Methods:

1. **Univariate feature selection:**
    - Correlation with SIPS scores
    - Mutual information
    - ANOVA F-statistic

2. **Recursive Feature Elimination (RFE)** with cross-validation
3. **Principal Component Analysis (PCA)** or other dimensionality reduction
4. **Regularization:** L1 (LASSO) or elastic net for embedded feature selection

Target: Reduce to 30-50 most informative features

Classification Models:

Binary Classification (Aim 2: Symptom Presence vs. Absence)

Traditional Machine Learning:

- **Logistic Regression** with regularization (baseline model)
- **Support Vector Machines (SVM)** with RBF kernel
- **Random Forest:** 500 trees, optimized hyperparameters
- **Gradient Boosting:** XGBoost or LightGBM
- **Ensemble methods:** Voting or stacking classifiers

Deep Learning:

- **Convolutional Neural Networks (CNN):**
    - Input: Spectrogram representations of speech
    - Architecture: Multiple convolutional layers, max pooling, fully connected layers
- **Recurrent Neural Networks (RNN/LSTM):**
    - Input: Sequential acoustic or linguistic features
    - Architecture: Bidirectional LSTM layers, attention mechanism
- **Transformer models:**
    - Pre-trained language models (BERT, GPT) fine-tuned on transcripts
- **Multi-modal fusion:**
    - Separate branches for acoustic and linguistic features
    - Late fusion or attention-based integration

Multi-Class Classification (Symptom Profiles)

- Adapt binary models for multi-class prediction

- One-vs-rest or softmax output layer
- Hierarchical classification (first symptom presence, then profile)

Regression Models (Symptom Severity)

- Predict continuous SIPS scores
- Evaluate as alternative to classification

Model Training and Validation:

Cross-Validation:

- Stratified 5-fold or 10-fold cross-validation on training set
- Hyperparameter tuning using grid search or Bayesian optimization
- Nested cross-validation for unbiased performance estimation

Performance Metrics:

- **Classification:**
    - Sensitivity (recall), specificity
    - Positive Predictive Value (PPV), Negative Predictive Value (NPV)
    - F1-score (harmonic mean of precision and recall)
    - Area Under ROC Curve (AUC-ROC)
    - Area Under Precision-Recall Curve (AUC-PR)
    - Cohen's Kappa (agreement beyond chance)
- **Regression:**
    - Mean Absolute Error (MAE)
    - Root Mean Squared Error (RMSE)
    - Pearson/Spearman correlation with actual SIPS scores

Calibration:

- Calibration plots (predicted probability vs. observed frequency)
- Brier score
- Calibration adjustment if needed (Platt scaling, isotonic regression)

Model Interpretability:

Feature Importance:

- SHAP (SHapley Additive exPlanations) values114 for all models
- Permutation importance
- Attention weights for deep learning models

Clinical Interpretation:

- Identify which vocal features drive predictions
- Map features to known symptom dimensions
- Generate example cases with feature visualizations

Bias and Fairness Evaluation:

Gender Bias Assessment:

- Stratified performance metrics by gender
- Demographic parity: P(predicted positive | female) vs. P(predicted positive | male)
- Equalized odds: Equal true positive and false positive rates across genders
- Calibration by gender

Bias Mitigation:

- Re-weighting training samples
- Gender-balanced training sets (oversampling males if needed)
- Adversarial debiasing techniques
- Post-processing adjustments

Age Group Analysis:

- Separate models or age-adjusted features for pediatric vs. adult
- Performance comparison across age groups

4.5 Clinical-Digital Phenotype Concordance Analysis

Level 1: Binary Concordance (Symptom Presence)

Analysis:

- 2×2 contingency table: Clinical diagnosis (SIPS ≥6) vs. Digital prediction
- Cohen's Kappa with 95% CI
- Sensitivity, specificity, PPV, NPV with 95% CI
- McNemar's test for discordance patterns

Discordant Case Analysis:

- **False Negatives (Clinical+/Digital-):** Qualitative review of symptom content, SIPS item profiles, recording quality
- **False Positives (Clinical-/Digital+):** Examine subthreshold symptoms (SIPS 2-5), comorbidities, inflammatory markers
- Predictive validity: Longitudinal follow-up of false positives to assess later symptom development

Level 2: Severity Concordance

Dimensional Correlation:

- Pearson/Spearman correlation between SIPS total score and vocal biomarker composite score
- Linear regression: Predict SIPS from vocal features
- Bland-Altman plots for agreement
- Intraclass Correlation Coefficient (ICC)

Ordinal Classification:

- 4-category severity (None, Minimal, Moderate, Severe)
- Weighted Kappa (quadratic weights)
- Confusion matrix analysis
- Proportional odds logistic regression

Level 3: Symptom Profile Concordance

Multi-Class Agreement:

- Confusion matrix for 5 symptom profiles
- Overall accuracy, class-specific sensitivity/specificity
- Multi-class Kappa

Profile-Specific Feature Analysis:

- Which vocal features best discriminate each profile?
- Correlation between SIPS subscales and feature clusters

Level 4: Symptom-Specific Correlations

Individual SIPS Items:

- Correlate each SIPS item with hypothesized vocal features:
    - P5 (disorganized communication) ↔ semantic coherence, syntactic complexity
    - N3 (expression of emotion) ↔ F0 variability, intensity range
    - P4 (perceptual abnormalities) ↔ pause patterns, self-referential language
    - P1 (unusual thought content) ↔ semantic tangentiality, unusual words
    - P2 (suspiciousness) ↔ negative sentiment, self-referential content

Cognitive Mediation:

- Path analysis: Do cognitive deficits mediate symptom-vocal feature relationships?
- Partial correlations controlling for cognitive function

Level 5: Incremental Validity

Hierarchical Regression:

- **Model 1:** Predict functional impairment from clinical variables (SIPS, HAM-D, cognitive scores)
- **Model 2:** Add vocal biomarker features
- **Analysis:** Significant ΔR^2 indicates incremental validity

Reclassification Analysis:

- Net Reclassification Improvement (NRI)
- Integrated Discrimination Improvement (IDI)
- Applied to predicting treatment response or symptom progression

Subgroup Analyses:

- Gender-stratified concordance
- Age-stratified concordance (pediatric vs. adult)
- Inflammatory phenotype stratification
- Cognitive profile stratification

5. CORTICOSTEROID INTERVENTION SUBSTUDY (AIM 3)

5.1 Participant Selection and Enrollment

Eligibility Screening:

- All participants in primary study with SIPS positive score ≥6 offered participation
- Additional screening for contraindications (see Exclusion Criteria)
- Medical clearance from primary care provider or endocrinologist
- Baseline labs: CBC, CMP, HbA1c, pregnancy test (females of childbearing potential)

Informed Consent:

- Separate informed consent for intervention substudy
- Detailed discussion of corticosteroid risks and benefits
- Alternative treatment options discussed
- Right to withdraw at any time

Target Enrollment: N=75 (to achieve N=60 completers assuming 20% attrition)

- Pediatric: n=37-38

- Adult: n=37-38

5.2 Intervention Protocol

Treatment Regimen:

Induction Phase (Weeks 1-4):

- **Prednisone** 0.5-1 mg/kg/day (maximum 60 mg/day for adults, 40 mg/day for pediatric)
- Dosing: Once daily in morning to minimize sleep disruption
- Starting dose determined by symptom severity and body weight

Maintenance Phase (Weeks 5-6):

- Continue same dose if good response
- May increase to 1 mg/kg if inadequate response at week 4 (investigator discretion)

Taper Phase (Weeks 7-10):

- **Week 7:** Reduce by 25% (e.g., 60 mg → 45 mg)
- **Week 8:** Reduce by additional 25% of original dose (e.g., 45 mg → 30 mg)
- **Week 9:** Reduce by additional 25% (e.g., 30 mg → 15 mg)
- **Week 10:** Reduce to 10 mg or discontinue (based on clinical response)

Post-Taper Follow-Up (Weeks 11-12):

- Off corticosteroids, monitor for symptom recurrence

Concomitant Medications:

- **Gastric protection:** Proton pump inhibitor (omeprazole 20 mg daily) or H2 blocker
- **Calcium and Vitamin D:** Calcium 1000-1500 mg/day, Vitamin D 1000-2000 IU/day
- **Other medications:** Continue stable doses of levothyroxine and other medications

- **Prohibited:** Initiation of new psychotropic medications during study period

Adherence Monitoring:

- Pill counts at each visit
- Self-report adherence diary
- Serum cortisol levels (should be suppressed during treatment)

5.3 Assessment Schedule

Baseline (Week 0):

- Complete assessment as per primary study protocol (Sections 3.1-3.3)
- Additional: Bone density scan (DEXA) for adults if clinically indicated

Week 2 (Early Response Assessment):

- SIPS (primary outcome)
- CGI-Improvement
- HAM-D, HAM-A
- Vocal recording (abbreviated: free speech tasks only, 10-15 minutes)
- Adverse events assessment
- Medication adherence check

Week 4 (Mid-Treatment Assessment):

- SIPS (primary outcome)
- CGI-Improvement, CGI-Severity
- HAM-D, HAM-A
- Brief cognitive battery (processing speed, attention, verbal fluency)
- Vocal recording (full protocol)
- Inflammatory markers: IL-6, CRP, ESR
- Thyroid function: TSH, Free T4
- Adverse events assessment
- Medication adherence check

Week 8 (End of Treatment Assessment):

- Complete assessment battery (as baseline):
    - SIPS, SCID-5 mood/anxiety modules
    - Full neurocognitive battery
    - All clinician-rated and self-report scales
    - Vocal recording (full protocol)
    - Inflammatory markers
    - Thyroid function
- Adverse events assessment
- Treatment satisfaction questionnaire

Week 12 (Post-Taper Follow-Up):

- SIPS (primary outcome)
- CGI-Severity
- HAM-D, HAM-A
- Brief cognitive battery
- Vocal recording (abbreviated)
- Inflammatory markers
- Assessment of symptom recurrence
- Adverse events (late effects)

5.4 Safety Monitoring

Adverse Events Assessment:

Common Expected Side Effects:

- Increased appetite, weight gain
- Mood changes (irritability, euphoria, mood lability)
- Sleep disturbance (insomnia)
- Gastrointestinal symptoms (dyspepsia, nausea)
- Acne, skin changes

- Fluid retention

Serious Adverse Events (SAEs):

- Severe mood changes (mania, severe depression, suicidality)
- Psychosis exacerbation
- Hyperglycemia, new-onset diabetes
- Severe hypertension
- Gastrointestinal bleeding
- Opportunistic infections
- Avascular necrosis
- Adrenal crisis (during taper)

Monitoring:

- Structured adverse events checklist at each visit
- Blood pressure, weight, glucose at each visit
- Psychiatric safety assessment (suicidality screening)
- 24/7 on-call clinician for urgent concerns
- SAE reporting to IRB within 24 hours

Dose Modification and Discontinuation Criteria:

Dose Reduction:

- Intolerable side effects (investigator discretion)
- Hyperglycemia (glucose >200 mg/dL on two occasions)
- Severe hypertension (BP >160/100 on two occasions)

Discontinuation:

- Participant request
- SAE requiring discontinuation (severe mood changes, psychosis exacerbation, medical complications)
- Non-adherence
- Pregnancy

5.5 Outcome Measures and Response Criteria

Primary Outcome:

- **Change in SIPS total positive symptom score** from baseline to week 8

Secondary Outcomes:

- Change in individual SIPS subscales (negative, disorganization, general)
- Change in vocal biomarker composite score
- Change in specific vocal features (prosody, semantic coherence, syntactic complexity)
- Change in inflammatory markers (IL-6, CRP)
- Change in cognitive function (domain composite scores)
- CGI-Improvement rating
- Quality of life change

Response Definitions:

Clinical Response:

- **Full Response:** ≥50% reduction in SIPS total positive score AND CGI-Improvement score of 1 (very much improved) or 2 (much improved)
- **Partial Response:** 30-49% reduction in SIPS total positive score OR CGI-Improvement score of 2-3
- **Non-Response:** <30% reduction in SIPS total positive score AND CGI-Improvement score of 4 or worse

Vocal Biomarker Response:

- Significant change in vocal biomarker composite score (≥0.5 SD from baseline)
- Normalization of key features (movement toward control group values)

5.6 Statistical Analysis Plan for Aim 3

Primary Analysis:

Longitudinal Change:

- **Mixed-effects model** with repeated measures:
    - Dependent variable: SIPS total positive score
    - Fixed effects: Time (baseline, week 2, 4, 8, 12), age group (pediatric/adult), baseline severity
    - Random effects: Participant (intercept and slope)
    - Covariance structure: Unstructured or autoregressive
- **Effect size:** Cohen's d for within-subject change (baseline to week 8)
- **Time to response:** Survival analysis (Kaplan-Meier) for time to achieve response criteria

Responder vs. Non-Responder Analysis:

- Compare baseline characteristics (demographics, symptom profiles, inflammatory markers, vocal features)
- Logistic regression: Predict response from baseline variables
- T-tests or Mann-Whitney U for continuous variables
- Chi-square for categorical variables

Secondary Analyses:

Vocal Biomarker Trajectories:

- Mixed-effects models for each vocal feature or composite score
- Parallel process models: Joint trajectories of SIPS and vocal biomarkers
- Correlation between SIPS change and vocal biomarker change
- Time-lagged analysis: Do vocal changes precede or follow clinical changes?

Inflammatory Marker Correlations:

- Correlation between inflammatory marker changes (Δ IL-6, Δ CRP) and:
    - SIPS score change

        - Vocal biomarker change
        - Cognitive function change
    - Mediation analysis: Do inflammatory changes mediate treatment effects on symptoms/vocal features?

Cognitive Function Changes:

- Mixed-effects models for cognitive domain scores
- Correlation with symptom improvement
- Mediation: Do cognitive improvements mediate symptom improvement?

Predictive Modeling:

- **Baseline predictors of treatment response:**
    - Logistic regression with baseline clinical, inflammatory, and vocal biomarker variables
    - Machine learning models (Random Forest, SVM) for prediction
    - Cross-validation for model validation
    - Performance metrics: AUC-ROC, sensitivity, specificity
- **Early change predictors:**
    - Does week 2 SIPS or vocal biomarker change predict week 8 response?
    - Logistic regression with early change variables

Subgroup Analyses:

- Age group (pediatric vs. adult)
- Gender
- Symptom profile (positive-predominant vs. others)
- Inflammatory phenotype (high vs. low inflammatory)
- Baseline severity (moderate vs. severe)

Safety Analysis:

- Descriptive statistics for adverse events (frequency, severity)
- Comparison with published corticosteroid safety data
- Identification of risk factors for adverse events

Sample Size and Power:

- N=60 completers (assuming 70% response rate → n=42 responders, n=18 non-responders)
- **Within-subject analysis (primary):**
    - Assuming large effect size (d=0.8) for SIPS change in responders
    - Power >0.99 to detect change with n=42
    - Even with n=18 non-responders, power=0.99 to detect moderate effect (d=0.6)
- **Between-group analysis (responder vs. non-responder):**
    - Power=0.70 to detect large effect (d=0.8) with n=42 vs. 18
    - Adequate for exploratory analysis of predictors

6. DATA MANAGEMENT AND QUALITY ASSURANCE

6.1 Data Collection and Management

Electronic Data Capture:

- REDCap (Research Electronic Data Capture)115 database for all clinical data
- Real-time data entry with range checks and validation rules
- Audit trail for all data changes
- Role-based access control
- HIPAA-compliant, secure servers

Audio Data Management:

- Secure file server with encrypted storage
- Standardized file naming convention: SiteID_ParticipantID_TaskID_Date
- Metadata database linking audio files to participant records
- Regular backups (daily incremental, weekly full)
- De-identification: Audio files stored separately from identifiable information

Data Quality Checks:

- Automated range and logic checks in REDCap
- Weekly data quality reports reviewing missing data, outliers, inconsistencies
- Monthly data audits by data manager
- Query resolution process for discrepancies

6.2 Training and Certification

Clinical Assessors:

- Master's or doctoral-level clinicians (psychologists, social workers, psychiatric nurses)
- **SIPS certification:** Online training and certification exam (www.sipstraining.com)
- **SCID-5 training:** Didactic training, practice interviews, certification by gold-standard rater
- **Cognitive testing:** Standardized training in test administration and scoring
- **Inter-rater reliability:** Ongoing assessment, target $\kappa \geq 0.80$ for SIPS, ICC ≥ 0.85 for cognitive tests
- **Recertification:** Annual refresher training and reliability checks

Vocal Recording Technicians:

- Training in equipment setup, recording procedures, quality monitoring
- Standardized protocol manual
- Supervised practice sessions
- Competency assessment before independent data collection

Transcribers:

- Training in transcription conventions
- Practice transcripts with feedback
- Inter-rater reliability assessment (20% double-coded, target $\kappa \geq 0.80$)

Site Coordinators:

- Training in recruitment, consent, scheduling, data management

- Good Clinical Practice (GCP) certification
- Regular communication and coordination meetings

6.3 Quality Assurance Procedures

Protocol Adherence:

- Standardized operating procedures (SOPs) for all study procedures
- Protocol deviation tracking and reporting
- Regular site monitoring visits (virtual or in-person)
- Corrective action plans for protocol violations

Data Monitoring:

- Data Safety Monitoring Board (DSMB) for intervention substudy
- Quarterly safety reports
- Interim analysis for futility or harm (if indicated)

Equipment Calibration:

- Regular calibration of audio recording equipment
- Test recordings to verify quality
- Replacement of faulty equipment

7. STATISTICAL CONSIDERATIONS

7.1 Missing Data

Prevention:

- Flexible scheduling to minimize missed visits
- Reminder calls/texts before appointments
- Incentives for visit completion
- Remote assessment options when feasible

Handling:

- **Primary analysis:** Multiple imputation (MI) using chained equations
    - Imputation model includes all outcome variables, predictors, and auxiliary variables
    - Generate 20-50 imputed datasets
    - Pool results using Rubin's rules
- **Sensitivity analyses:**
    - Complete case analysis
    - Pattern-mixture models to assess impact of missing data patterns
    - Worst-case and best-case scenarios for intervention substudy

Attrition Analysis:

- Compare completers vs. non-completers on baseline characteristics
- Logistic regression to identify predictors of attrition
- Inverse probability weighting if attrition is non-random

7.2 Multiple Comparisons

Approach:

- **Primary analyses:** Control family-wise error rate using Bonferroni or Holm-Bonferroni correction
- **Exploratory analyses:** Control false discovery rate (FDR) using Benjamini-Hochberg procedure
- **Pre-specified hypotheses:** Clearly distinguish confirmatory from exploratory analyses
- **Hierarchical testing:** Test primary outcomes first, secondary outcomes only if primary significant

7.3 Subgroup and Sensitivity Analyses

Pre-Specified Subgroups:

- Age group (pediatric vs. adult)
- Gender (male vs. female)
- Symptom profile (positive-predominant, negative-predominant, disorganized, mixed)
- Inflammatory phenotype (high vs. low)
- Baseline severity (moderate vs. severe)

Sensitivity Analyses:

- Exclude participants with threshold psychotic symptoms (SIPS item rated 6)
- Exclude participants with significant comorbidities
- Different classification thresholds for symptom presence
- Different feature selection methods
- Different machine learning algorithms

7.4 Interim Analyses

Intervention Substudy:

- One interim analysis at 50% enrollment (n=30 completers)
- Assess for futility (conditional power <20%) or unexpected harm
- O'Brien-Fleming alpha spending function to preserve overall Type I error
- DSMB reviews interim results and makes recommendations

8. HUMAN SUBJECTS PROTECTION

8.1 Risks to Participants

Minimal Risk Procedures (Primary Observational Study):

- **Clinical interviews and questionnaires:** Potential emotional distress from discussing psychiatric symptoms
    - Mitigation: Trained, sensitive interviewers; option to skip questions or take breaks; referral resources provided
- **Cognitive testing:** Potential frustration or fatigue
    - Mitigation: Breaks offered; reassurance that performance does not reflect intelligence
- **Vocal recordings:** Minimal risk; potential self-consciousness
    - Mitigation: Private setting; explanation that focus is on speech patterns, not content
- **Blood draw:** Minor pain, bruising, rare vasovagal reaction or infection
    - Mitigation: Experienced phlebotomists; standard sterile technique

Greater Than Minimal Risk (Corticosteroid Intervention Substudy):

- **Common side effects:** Increased appetite, weight gain, mood changes, sleep disturbance, acne, GI symptoms
    - Mitigation: Education about expected side effects; monitoring at each visit; dose adjustment if needed
- **Serious adverse events:** Severe mood changes, psychosis exacerbation, hyperglycemia, hypertension, GI bleeding, infections, avascular necrosis (rare)
    - Mitigation: Careful screening for contraindications; close monitoring; 24/7 on-call clinician; clear discontinuation criteria; SAE reporting
- **Adrenal suppression:** Risk of adrenal crisis if abrupt discontinuation
    - Mitigation: Gradual taper protocol; education about not stopping abruptly; stress-dose steroids if needed for illness/surgery

Risk-Benefit Assessment:

- **Primary study:** Minimal risk, potential benefit of identifying treatable

symptoms
- **Intervention substudy:** Risks of corticosteroids justified by:
    - Potential for significant symptom improvement (60-90% response rate in literature)
    - Close monitoring and safety protocols
    - Alternative treatments (antipsychotics) have comparable or greater side effect burden
    - Participants have clinically significant symptoms impacting functioning

8.2 Informed Consent

Process:

- Written informed consent (adults) or assent (minors ages 10-17) with parental permission
- Separate consent for primary study and intervention substudy
- Consent forms written at 8th grade reading level
- Key elements:
    - Study purpose, procedures, duration
    - Risks and benefits
    - Alternatives to participation
    - Confidentiality protections and limits
    - Voluntary participation, right to withdraw
    - Compensation
    - Contact information for questions or concerns
- Teach-back method to assess comprehension
- Opportunity to ask questions
- Time to consider participation (at least 24 hours for intervention substudy)

Capacity Assessment:

- For participants with significant psychiatric symptoms, assess decision-making capacity
- If capacity questionable, involve legally authorized representative

Ongoing Consent:

- Re-consent if protocol changes
- Periodic check-ins about willingness to continue

8.3 Confidentiality and Data Security

Identifiable Information:

- Stored separately from research data
- Access limited to essential personnel
- Encrypted databases with password protection

De-Identification:

- Participant ID numbers used for all research data
- Audio recordings de-identified (no names mentioned)
- Transcripts reviewed to remove identifying information

Data Storage:

- Electronic data: HIPAA-compliant servers, encrypted, regular backups
- Paper records: Locked file cabinets in secure offices
- Audio files: Encrypted file server, separate from identifiable information

Data Sharing:

- De-identified data may be shared with other researchers (with participant consent)
- Data use agreements required
- Compliance with NIH data sharing policies

Certificates of Confidentiality:

- Obtained from NIH to protect against compelled disclosure

8.4 Vulnerable Populations

Minors (Ages 10-17):

- Assent required in addition to parental permission
- Age-appropriate explanation of study
- Right to decline or withdraw even if parents consent
- Additional protections for intervention substudy (parental involvement in treatment decisions)

Individuals with Psychiatric Symptoms:

- May have impaired decision-making capacity
- Careful assessment of capacity to consent
- Ongoing monitoring of wellbeing
- Clear procedures for managing acute psychiatric crises

Protections:

- Trained staff sensitive to vulnerabilities
- Referral resources for clinical care
- Safety protocols for suicidality or acute symptoms
- Flexibility to accommodate needs

8.5 Incidental Findings

Definition: Clinically significant findings discovered during research that are not the focus of the study

Management Plan:

- Participants informed during consent that incidental findings may be discovered
- Significant findings (e.g., severe depression, suicidality, cognitive impairment) disclosed to participant and (with permission) their clinician
- Referrals provided for clinical follow-up
- Documentation of disclosure and referral

Examples:

- Suicidal ideation: Immediate safety assessment, referral to emergency services if needed
- Severe cognitive impairment: Referral to neurology or neuropsychology
- Abnormal lab values: Referral to primary care or endocrinology

8.6 IRB Oversight

Initial Review:

- Full board review (greater than minimal risk for intervention substudy)
- Approval from all participating sites' IRBs
- Reliance agreements if using single IRB

Continuing Review:

- Annual renewal
- Progress reports including enrollment, adverse events, protocol deviations
- Modifications submitted for approval before implementation

Adverse Event Reporting:

- SAEs reported to IRB within 24 hours
- Unanticipated problems reported within 5 business days
- Annual summary of all adverse events

9. INCLUSION OF WOMEN AND MINORITIES

9.1 Inclusion of Women

Rationale:

- Autoimmune thyroiditis has 10:1 female predominance
- Expected enrollment: ~80% female, ~20% male

- This reflects disease epidemiology and ensures adequate female representation

Gender-Specific Considerations:

- Analysis of gender differences in symptom presentation and vocal features
- Assessment of gender bias in clinical vs. digital phenotyping
- Evaluation of algorithmic fairness across genders

Pregnancy:

- Females of childbearing potential excluded from intervention substudy due to corticosteroid risks
- Pregnancy testing required before enrollment in intervention substudy

9.2 Inclusion of Minorities

Recruitment Strategy:

- Targeted recruitment from diverse clinical settings
- Community outreach to underserved populations
- Collaboration with community health centers
- Culturally sensitive recruitment materials

Target Enrollment:

- Reflect local population demographics
- Minimum 30% racial/ethnic minorities
- Specific targets:
    - Hispanic/Latino: 15-20%
    - Black/African American: 10-15%
    - Asian: 5-10%
    - Other: 5%

Cultural Considerations:

- Bilingual staff if needed

- Culturally appropriate communication
- Flexibility in scheduling and location
- Address barriers to participation (transportation, childcare)

Analysis:

- Examine racial/ethnic differences in symptom presentation and treatment response
- Assess for algorithmic bias across racial/ethnic groups
- Report enrollment and outcomes by race/ethnicity

10. TIMELINE AND MILESTONES

Year 1 (Months 1-12)

Months 1-3: Study Setup

- IRB submissions and approvals (all sites)
- Finalize data management systems (REDCap, audio storage)
- Hire and train study staff
- Establish site coordination procedures

Months 4-6: Protocol Piloting

- Pilot testing with 10-15 participants per site
- Refine procedures based on pilot experience
- Establish inter-rater reliability
- Test data flow and quality checks

Months 7-12: Initial Recruitment

- Begin active recruitment
- Target: 100 participants enrolled (50 pediatric, 50 adult)
- Ongoing quality monitoring and staff training
- Begin intervention substudy enrollment (target: 15 participants)

Year 2 (Months 13-24)

Months 13-24: Main Data Collection

- Continue recruitment and assessment
- Target: Additional 250 participants enrolled (125 pediatric, 125 adult)
- Cumulative total: 350 participants
- Continue intervention substudy enrollment (target: Additional 45 participants, cumulative 60)
- Complete intervention substudy assessments for early enrollees
- Ongoing data quality monitoring
- Preliminary data analysis and model development

Milestones:

- Month 18: 50% enrollment complete (175 participants)
- Month 18: Interim analysis for intervention substudy (DSMB review)
- Month 24: Primary recruitment complete (350 participants)

Year 3 (Months 25-36)

Months 25-27: Final Data Collection

- Complete intervention substudy follow-up assessments
- Complete any pending baseline assessments
- Database lock for primary observational study

Months 28-30: Data Analysis

- Feature extraction and preprocessing (all vocal recordings)
- Machine learning model development and validation
- Statistical analyses for Aims 1-3
- Clinical-digital concordance analysis
- Subgroup and sensitivity analyses

Months 31-33: Manuscript Preparation

- Primary manuscript (Aim 1: Prevalence and characterization)
- Secondary manuscript (Aim 2: Vocal biomarkers)
- Tertiary manuscript (Aim 3: Treatment response)
- Additional papers on specific topics (gender bias, cognitive correlates, etc.)

Months 34-36: Dissemination

- Submit manuscripts for publication
- Present findings at scientific conferences
- Prepare final grant report
- Plan for future studies and clinical implementation

Key Deliverables:

- 3-5 peer-reviewed publications
- Conference presentations (American Psychiatric Association, Society of Biological Psychiatry, Endocrine Society)
- Open-source software tools for vocal biomarker analysis
- Clinical decision support framework
- Grant applications for validation studies and clinical trials

---

3. BIBLIOGRAPHY AND REFERENCES CITED

---

1. Chaker L, Bianco AC, Jonklaas J, Peeters RP. Hypothyroidism. *Lancet.* 2017;390(10101):1550-1562. doi:10.1016/S0140-6736(17)30703-1
2. Caturegli P, De Remigis A, Rose NR. Hashimoto thyroiditis: clinical and diagnostic criteria. *Autoimmun Rev.* 2014;13(4-5):391-397. doi:10.1016/j.autrev.2014.01.007
3. Leyhe T, Müssig K. Cognitive and affective dysfunctions in autoimmune thyroiditis. *Brain Behav Immun.* 2014;41:261-266. doi:10.1016/j.bbi.2014.03.008
4. Chong JY, Rowland LP, Utiger RD. Hashimoto encephalopathy: syndrome or

myth? *Arch Neurol.* 2003;60(2):164-171. doi:10.1001/archneur.60.2.164

5. Graus F, Titulaer MJ, Balu R, et al. A clinical approach to diagnosis of autoimmune encephalitis. *Lancet Neurol.* 2016;15(4):391-404. doi:10.1016/S1474-4422(15)00401-9
6. Montagna G, Imperiali M, Agazzi P, Dannenberg C, Lobrinus JA, Burkhard PR. Hashimoto's encephalopathy: A rare proteiform disorder. *Autoimmun Rev.* 2016;15(5):466-476. doi:10.1016/j.autrev.2016.01.014
7. Mahmud FH, Lteif AN, Renaud DL, Reed AM, Brands CK. Steroid-responsive encephalopathy associated with Hashimoto's thyroiditis in an adolescent with chronic hallucinations and depression: case report and review. *Pediatrics.* 2003;112(3 Pt 1):686-690. doi:10.1542/peds.112.3.686
8. Ferracci F, Bertiato G, Moretto G. Hashimoto's encephalopathy: epidemiologic data and pathogenetic considerations. *J Neurol Sci.* 2004;217(2):165-168. doi:10.1016/j.jns.2003.09.007
9. Ochoa S, Usall J, Cobo J, Labad X, Kulkarni J. Gender differences in schizophrenia and first-episode psychosis: a comprehensive literature review. *Schizophr Res Treatment.* 2012;2012:916198. doi:10.1155/2012/916198
10. Cotton SM, Lambert M, Schimmelmann BG, et al. Gender differences in premorbid, entry, treatment, and outcome characteristics in a treated epidemiological sample of 661 patients with first episode psychosis. *Schizophr Res.* 2009;114(1-3):17-24. doi:10.1016/j.schres.2009.07.002
11. Graus F, Escudero D, Oleaga L, et al. Syndrome and outcome of antibody-negative limbic encephalitis. *Eur J Neurol.* 2018;25(8):1011-1016. doi:10.1111/ene.13661
12. Pillai SC, Hacohen Y, Tantsis E, et al. Infectious and autoantibody-associated encephalitis: clinical features and long-term outcome. *Pediatrics.* 2015;135(4):e974-e984. doi:10.1542/peds.2014-2702
13. 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. doi:10.1002/wps.20491
14. Elvevåg B, Foltz PW, Weinberger DR, Goldberg TE. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophr Res.* 2007;93(1-3):304-316. doi:10.1016/j.schres.2007.03.001
15. Covington MA, He C, Brown C, et al. Schizophrenia and the structure of language: the linguist's view. *Schizophr Res.* 2005;77(1):85-98. doi:10.1016/j.schres.2005.01.016
16. 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. doi:10.1038/npjschz.2015.30

17. Corcoran CM, Mittal VA, Bearden CE, et al. Language as a biomarker for psychosis: A natural language processing approach. *Schizophr Res.* 2020;226:158-166. doi:10.1016/j.schres.2020.04.032

18. 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

19. Insel TR. Digital phenotyping: technology for a new science of behavior. *JAMA.* 2017;318(13):1215-1216. doi:10.1001/jama.2017.11295

20. Torous J, Kiang MV, Lorme J, Onnela JP. New tools for new research in psychiatry: a scalable and customizable platform to empower data driven smartphone research. *JMIR Ment Health.* 2016;3(2):e16. doi:10.2196/mental.5165

21. Hollowell JG, Staehling NW, Flanders WD, et al. Serum TSH, T(4), and thyroid antibodies in the United States population (1988 to 1994): National Health and Nutrition Examination Survey (NHANES III). *J Clin Endocrinol Metab.* 2002;87(2):489-499. doi:10.1210/jcem.87.2.8182

22. Garber JR, Cobin RH, Gharib H, et al. Clinical practice guidelines for hypothyroidism in adults: cosponsored by the American Association of Clinical Endocrinologists and the American Thyroid Association. *Endocr Pract.* 2012;18(6):988-1028. doi:10.4158/EP12280.GL

23. Siegmann EM, Müller HHO, Luecke C, et al. Association of Depression and Anxiety Disorders With Autoimmune Thyroiditis: A Systematic Review and Meta-analysis. *JAMA Psychiatry.* 2018;75(6):577-584. doi:10.1001/jamapsychiatry.2018.0190

24. Samuels MH. Psychiatric and cognitive manifestations of hypothyroidism. *Curr Opin Endocrinol Diabetes Obes.* 2014;21(5):377-383. doi:10.1097/MED.0000000000000089

25. Bauer M, Goetz T, Glenn T, Whybrow PC. The thyroid-brain interaction in thyroid disorders and mood disorders. *J Neuroendocrinol.* 2008;20(10):1101-1114. doi:10.1111/j.1365-2826.2008.01774.x

26. Castillo P, Woodruff B, Caselli R, et al. Steroid-responsive encephalopathy associated with autoimmune thyroiditis. *Arch Neurol.* 2006;63(2):197-202. doi:10.1001/archneur.63.2.197

27. Mocellin R, Walterfang M, Velakoulis D. Hashimoto's encephalopathy: epidemiology, pathogenesis and management. *CNS Drugs.* 2007;21(10):799-811. doi:10.2165/00023210-200721100-00002

28. Schiess N, Pardo CA. Hashimoto's encephalopathy. *Ann N Y Acad Sci.*

2008;1142:254-265. doi:10.1196/annals.1444.018

29. Fusar-Poli P, Borgwardt S, Bechdolf A, et al. The psychosis high-risk state: a comprehensive state-of-the-art review. *JAMA Psychiatry.* 2013;70(1):107-120. doi:10.1001/jamapsychiatry.2013.269

30. Cannon TD, Yu C, Addington J, et al. An individualized risk calculator for research in prodromal psychosis. *Am J Psychiatry.* 2016;173(10):980-988. doi:10.1176/appi.ajp.2016.15070890

31. Kelleher I, Connor D, Clarke MC, Devlin N, Harley M, Cannon M. Prevalence of psychotic symptoms in childhood and adolescence: a systematic review and meta-analysis of population-based studies. *Psychol Med.* 2012;42(9):1857-1863. doi:10.1017/S0033291711002960

32. Riecher-Rössler A, Häfner H, Dütsch-Strobel A, et al. Further evidence for a specific role of estradiol in schizophrenia? *Biol Psychiatry.* 1994;36(7):492-494. doi:10.1016/0006-3223(94)90650-5

33. Abel KM, Drake R, Goldstein JM. Sex differences in schizophrenia. *Int Rev Psychiatry.* 2010;22(5):417-428. doi:10.3109/09540261.2010.515205

34. Laurent C, Capron J, Quillerou B, et al. Steroid-responsive encephalopathy associated with autoimmune thyroiditis (SREAT): Characteristics, treatment and outcome in 251 cases from the literature. *Autoimmun Rev.* 2016;15(12):1129-1133. doi:10.1016/j.autrev.2016.09.008

35. Mattozzi S, Sabater L, Escudero D, et al. Hashimoto encephalopathy in the 21st century. *Neurology.* 2020;94(2):e217-e224. doi:10.1212/WNL.0000000000008785

36. Tamagno G, Celik Y, Simó R, Dejager S. Encephalopathy associated with autoimmune thyroid disease in patients with Graves' disease: clinical manifestations, follow-up, and outcomes. *BMC Neurol.* 2010;10:27. doi:10.1186/1471-2377-10-27

37. Endres D, Perlov E, Stich O, Tebartz van Elst L. Steroid responsive encephalopathy associated with autoimmune thyroiditis (SREAT) presenting as major depression. *BMC Psychiatry.* 2016;16:184. doi:10.1186/s12888-016-0897-3

38. Najjar S, Pearlman DM, Alper K, Najjar A, Devinsky O. Neuroinflammation and psychiatric illness. *J Neuroinflammation.* 2013;10:43. doi:10.1186/1742-2094-10-43

39. Khandaker GM, Cousins L, Deakin J, Lennox BR, Yolken R, Jones PB. Inflammation and immunity in schizophrenia: implications for pathophysiology and treatment. *Lancet Psychiatry.* 2015;2(3):258-270. doi:10.1016/S2215-0366(14)00122-9

40. Pollak TA, Drndarski S, Stone JM, David AS, McGuire P, Abbott NJ. The blood-

brain barrier in psychosis. *Lancet Psychiatry.* 2018;5(1):79-92. doi:10.1016/S2215-0366(17)30293-6

41. Blanchin S, Coffin C, Viader F, et al. Anti-thyroperoxidase antibodies from patients with Hashimoto's encephalopathy bind to cerebellar astrocytes. *J Neuroimmunol.* 2007;192(1-2):13-20. doi:10.1016/j.jneuroim.2007.08.012

42. Ferracci F, Moretto G, Candeago RM, et al. Antithyroid antibodies in the CSF: their role in the pathogenesis of Hashimoto's encephalopathy. *Neurology.* 2003;60(4):712-714. doi:10.1212/01.wnl.0000048660.71390.c6

43. Yoneda M, Fujii A, Ito A, Yokoyama H, Nakagawa H, Kuriyama M. High prevalence of serum autoantibodies against the amino terminal of α-enolase in Hashimoto's encephalopathy. *J Neuroimmunol.* 2007;185(1-2):195-200. doi:10.1016/j.jneuroim.2007.01.018

44. Takahashi S, Mitamura R, Itoh Y, Suzuki N, Okuno A. Hashimoto encephalopathy: etiologic considerations. *Pediatr Neurol.* 1994;11(4):328-331. doi:10.1016/0887-8994(94)90011-6

45. Oide T, Tokuda T, Yazaki M, et al. Anti-neuronal autoantibody in Hashimoto's encephalopathy: neuropathological, immunohistochemical, and biochemical analysis of two patients. *J Neurol Sci.* 2004;217(1):7-12. doi:10.1016/j.jns.2003.08.005

46. Doherty CP, Schlossmacher M, Torres N, et al. Hashimoto's encephalopathy mimicking Creutzfeldt-Jakob disease: brain biopsy findings. *J Neurol Neurosurg Psychiatry.* 2002;73(5):601-602. doi:10.1136/jnnp.73.5.601

47. Miller BJ, Buckley P, Seabolt W, Mellor A, Kirkpatrick B. Meta-analysis of cytokine alterations in schizophrenia: clinical status and antipsychotic effects. *Biol Psychiatry.* 2011;70(7):663-671. doi:10.1016/j.biopsych.2011.04.013

48. Goldsmith DR, Rapaport MH, Miller BJ. A meta-analysis of blood cytokine network alterations in psychiatric patients: comparisons between schizophrenia, bipolar disorder and depression. *Mol Psychiatry.* 2016;21(12):1696-1709. doi:10.1038/mp.2016.3

49. Doorduin J, de Vries EF, Willemsen AT, de Groot JC, Dierckx RA, Klein HC. Neuroinflammation in schizophrenia-related psychosis: a PET study. *J Nucl Med.* 2009;50(11):1801-1807. doi:10.2967/jnumed.109.066647

50. Berk M, Kapczinski F, Andreazza AC, et al. Pathways underlying neuroprogression in bipolar disorder: focus on inflammation, oxidative stress and neurotrophic factors. *Neurosci Biobehav Rev.* 2011;35(3):804-817. doi:10.1016/j.neubiorev.2010.10.001

51. Howes OD, Kapur S. The dopamine hypothesis of schizophrenia: version III--the

final common pathway. *Schizophr Bull.* 2009;35(3):549-562. doi:10.1093/schbul/sbp006

52. Moghaddam B, Javitt D. From revolution to evolution: the glutamate hypothesis of schizophrenia and its implication for treatment. *Neuropsychopharmacology.* 2012;37(1):4-15. doi:10.1038/npp.2011.181

53. Schmitt SE, Pargeon K, Frechette ES, Hirsch LJ, Dalmau J, Friedman D. Extreme delta brush: a unique EEG pattern in adults with anti-NMDA receptor encephalitis. *Neurology.* 2012;79(11):1094-1100. doi:10.1212/WNL.0b013e3182698cd8

54. Bien CG, Vincent A, Barnett MH, et al. Immunopathology of autoantibody-associated encephalitides: clues for pathogenesis. *Brain.* 2012;135(Pt 5):1622-1638. doi:10.1093/brain/aws082

55. Leypoldt F, Buchert R, Kleiter I, et al. Fluorodeoxyglucose positron emission tomography in anti-N-methyl-D-aspartate receptor encephalitis: distinct pattern of disease. *J Neurol Neurosurg Psychiatry.* 2012;83(7):681-686. doi:10.1136/jnnp-2011-301969

56. Olmez I, Moses H, Sriram S, Kirshner H, Lagrange AH, Pawate S. Diagnostic and therapeutic aspects of Hashimoto's encephalopathy. *J Neurol Sci.* 2013;331(1-2):67-71. doi:10.1016/j.jns.2013.05.009

57. Chaudhuri A, Behan PO. The clinical spectrum, diagnosis, pathogenesis and treatment of Hashimoto's encephalopathy (recurrent acute disseminated encephalomyelitis). *Curr Med Chem.* 2003;10(19):1945-1953. doi:10.2174/0929867033456746

58. Thorup A, Petersen L, Jeppesen P, Ohlenschlaeger J, Christensen T, Krarup G, et al. Gender differences in young adults with first-episode schizophrenia spectrum disorders at baseline in the Danish OPUS study. *J Nerv Ment Dis.* 2007;195(5):396-405. doi:10.1097/01.nmd.0000253783.00407.95

59. Morgan VA, Castle DJ. Sex and schizophrenia. *Curr Psychiatry Rev.* 2006;2(3):297-307. doi:10.2174/157340006778018125

60. Leung A, Chue P. Sex differences in schizophrenia, a review of the literature. *Acta Psychiatr Scand Suppl.* 2000;401:3-38. doi:10.1034/j.1600-0447.2000.28101.x

61. Frith CD, Friston KJ, Herold S, et al. Regional brain activity in chronic schizophrenic patients during the performance of a verbal fluency task. *Br J Psychiatry.* 1995;167(3):343-349. doi:10.1192/bjp.167.3.343

62. Carter CS, MacDonald AW 3rd, Ross LL, Stenger VA. Anterior cingulate cortex activity and impaired self-monitoring of performance in patients with schizophrenia: an event-related fMRI study. *Am J Psychiatry.* 2001;158(9):1423-1428. doi:10.1176/appi.ajp.158.9.1423

63. Alpert M, Rosen A, Welkowitz J, Sobin C, Borod JC. Vocal acoustic correlates of flat affect in schizophrenia. Similarity to Parkinson's disease and right hemisphere disease and contrast with depression. *Br J Psychiatry Suppl.* 1989; (4):51-56.

64. Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci.* 2007;8(5):393-402. doi:10.1038/nrn2113

65. Whitfield-Gabrieli S, Ford JM. Default mode network activity and connectivity in psychopathology. *Annu Rev Clin Psychol.* 2012;8:49-76. doi:10.1146/annurev-clinpsy-032511-143049

66. Cohen AS, Mitchell KR, Elvevåg B. What do we really know about blunted vocal affect and alogia? A meta-analysis of objective assessments. *Schizophr Res.* 2014;159(2-3):533-538. doi:10.1016/j.schres.2014.09.013

67. Alpert M, Pouget ER, Silva RR. Reflections of depression in acoustic measures of the patient's speech. *J Affect Disord.* 2001;66(1):59-69. doi:10.1016/s0165-0327(00)00335-9

68. Alpert M, Shaw RJ, Pouget ER, Lim KO. A comparison of clinical ratings with vocal acoustic measures of flat affect and alogia. *J Psychiatr Res.* 2002;36(5):347-353. doi:10.1016/s0022-3956(02)00016-x

69. Rapcan V, D'Arcy S, Yeap S, Afzal N, Thakore J, Reilly RB. Acoustic and temporal analysis of speech: a potential biomarker for schizophrenia. *Med Eng Phys.* 2010;32(9):1074-1079. doi:10.1016/j.medengphy.2010.07.013

70. Morice R, Ingram JC. Language analysis in schizophrenia: diagnostic implications. *Aust N Z J Psychiatry.* 1982;16(1):11-21. doi:10.3109/00048678209161186

71. Andreasen NC. Thought, language, and communication disorders. I. Clinical assessment, definition of terms, and evaluation of their reliability. *Arch Gen Psychiatry.* 1979;36(12):1315-1321. doi:10.1001/archpsyc.1979.01780120045006

72. Andreasen NC. Scale for the Assessment of Thought, Language, and Communication (TLC). *Schizophr Bull.* 1986;12(3):473-482. doi:10.1093/schbul/12.3.473

73. Rezaii N, Walker E, Wolff P. A machine learning approach to predicting psychosis using semantic density and latent content analysis. *NPJ Schizophr.* 2019;5(1):9. doi:10.1038/s41537-019-0077-9

74. Aversa T, Valenzise M, Corrias A, et al. Underlying Hashimoto's thyroiditis negatively affects the evolution of subclinical hypothyroidism in children irrespective of other concomitant risk factors. *Thyroid.* 2015;25(2):183-187. doi:10.1089/thy.2014.0235

75. Paus T, Keshavan M, Giedd JN. Why do many psychiatric disorders emerge during adolescence? *Nat Rev Neurosci.* 2008;9(12):947-957. doi:10.1038/nrn2513

76. Lee S, Potamianos A, Narayanan S. Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J Acoust Soc Am.* 1999;105(3):1455-1468. doi:10.1121/1.426686

77. First MB, Williams JBW, Karg RS, Spitzer RL. *Structured Clinical Interview for DSM-5—Research Version (SCID-5 for DSM-5, Research Version; SCID-5-RV).* Arlington, VA: American Psychiatric Association; 2015.

78. Miller TJ, McGlashan TH, Rosen JL, et al. Prodromal assessment with the structured interview for prodromal syndromes and the scale of prodromal symptoms: predictive validity, interrater reliability, and training to reliability. *Schizophr Bull.* 2003;29(4):703-715. doi:10.1093/oxfordjournals.schbul.a007040

79. McGlashan TH, Walsh BC, Woods SW. *The Psychosis-Risk Syndrome: Handbook for Diagnosis and Follow-Up.* New York: Oxford University Press; 2010.

80. Nuechterlein KH, Green MF, Kern RS, et al. The MATRICS Consensus Cognitive Battery, part 1: test selection, reliability, and validity. *Am J Psychiatry.* 2008;165(2):203-213. doi:10.1176/appi.ajp.2007.07010042

81. Reitan RM. *Trail Making Test: Manual for Administration and Scoring.* Tucson, AZ: Reitan Neuropsychology Laboratory; 1992.

82. Wechsler D. *Wechsler Adult Intelligence Scale–Fourth Edition (WAIS–IV).* San Antonio, TX: NCS Pearson; 2008.

83. Benton AL, Hamsher K, Sivan AB. *Multilingual Aphasia Examination.* 3rd ed. Iowa City, IA: AJA Associates; 1994.

84. Golden CJ, Freshwater SM. *Stroop Color and Word Test: A Manual for Clinical and Experimental Uses.* Wood Dale, IL: Stoelting; 2002.

85. Brandt J, Benedict RHB. *Hopkins Verbal Learning Test–Revised: Professional Manual.* Lutz, FL: Psychological Assessment Resources; 2001.

86. Benedict RHB. *Brief Visuospatial Memory Test–Revised: Professional Manual.* Lutz, FL: Psychological Assessment Resources; 1997.

87. Kaplan E, Goodglass H, Weintraub S. *Boston Naming Test.* 2nd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2001.

88. Gorham DR. *Proverbs Test.* Missoula, MT: Psychological Test Specialists; 1956.

89. Baron-Cohen S, Wheelwright S, Hill J, Raste Y, Plumb I. The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J Child Psychol Psychiatry.* 2001;42(2):241-251.

90. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry.* 1960;23:56-62. doi:10.1136/jnnp.23.1.56

91. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry.* 1979;134:382-389. doi:10.1192/bjp.134.4.382

92. Hamilton M. The assessment of anxiety states by rating. *Br J Med Psychol.* 1959;32(1):50-55. doi:10.1111/j.2044-8341.1959.tb00467.x

93. Guy W. *ECDEU Assessment Manual for Psychopharmacology.* Rockville, MD: US Department of Health, Education, and Welfare Public Health Service Alcohol, Drug Abuse, and Mental Health Administration; 1976.

94. Ising HK, Veling W, Loewy RL, et al. The validity of the 16-item version of the Prodromal Questionnaire (PQ-16) to screen for ultra high risk of developing psychosis in the general help-seeking population. *Schizophr Bull.* 2012;38(6):1288-1296. doi:10.1093/schbul/sbs068

95. Stefanis NC, Hanssen M, Smirnis NK, et al. Evidence that three dimensions of psychosis have a distribution in the general population. *Psychol Med.* 2002;32(2):347-358. doi:10.1017/s0033291701005141

96. Beck AT, Steer RA, Brown GK. *Beck Depression Inventory–Second Edition Manual.* San Antonio, TX: The Psychological Corporation; 1996.

97. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-613. doi:10.1046/j.1525-1497.2001.016009606.x

98. Beck AT, Epstein N, Brown G, Steer RA. An inventory for measuring clinical anxiety: psychometric properties. *J Consult Clin Psychol.* 1988;56(6):893-897. doi:10.1037//0022-006x.56.6.893

99. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med.* 2006;166(10):1092-1097. doi:10.1001/archinte.166.10.1092

100. Bernstein DP, Fink L, Handelsman L, et al. Initial reliability and validity of a new retrospective measure of child abuse and neglect. *Am J Psychiatry.* 1994;151(8):1132-1136. doi:10.1176/ajp.151.8.1132

101. World Health Organization. *WHOQOL-BREF: Introduction, Administration, Scoring and Generic Version of the Assessment.* Geneva: World Health Organization; 1996.

102. Varni JW, Seid M, Kurtin PS. PedsQL 4.0: reliability and validity of the Pediatric Quality of Life Inventory version 4.0 generic core scales in healthy and patient populations. *Med Care.* 2001;39(8):800-812. doi:10.1097/00005650-200108000-00006

103. Weissman MM, Bothwell S. Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry.* 1976;33(9):1111-1115. doi:10.1001/archpsyc.1976.01770090101010

104. Goodglass H, Kaplan E, Barresi B. *Boston Diagnostic Aphasia Examination.* 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2001.

105. Fairbanks G. *Voice and Articulation Drillbook.* 2nd ed. New York: Harper & Row; 1960.

106. Boersma P, Weenink D. *Praat: Doing Phonetics by Computer [Computer program].* Version 6.1.16. 2020. http://www.praat.org/

107. Eyben F, Wöllmer M, Schuller B. Opensmile: the munich versatile and fast open-source audio feature extractor. *Proceedings of the 18th ACM International Conference on Multimedia.* 2010:1459-1462. doi:10.1145/1873951.1874246

108. Covington MA, McFall JD. Cutting the Gordian knot: the moving-average type-token ratio (MATTR). *J Quant Linguist.* 2010;17(2):94-100. doi:10.1080/09296171003643098

109. McCarthy PM, Jarvis S. MTLD, vocd-D, and HD-D: a validation study of sophisticated approaches to lexical diversity assessment. *Behav Res Methods.* 2010;42(2):381-392. doi:10.3758/BRM.42.2.381

110. Brysbaert M, New B. Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav Res Methods.* 2009;41(4):977-990. doi:10.3758/BRM.41.4.977

111. Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. *Discourse Process.* 1998;25(2-3):259-284. doi:10.1080/01638539809545028

112. Yngve VH. A model and an hypothesis for language structure. *Proc Am Philos Soc.* 1960;104(5):444-466.

113. Hutto CJ, Gilbert E. VADER: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the 8th International Conference on Weblogs and Social Media (ICWSM-14).* 2014.

114. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst.* 2017;30:4765-4774.

115. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform.* 2009;42(2):377-381. doi:10.1016/j.jbi.2008.08.010

4. PROTECTION OF HUMAN SUBJECTS

[See Section 8 of Research Strategy for detailed discussion]

Summary:

- IRB approval obtained before any study procedures
- Informed consent with teach-back method
- Separate consent for intervention substudy
- Minimal risk for observational study; greater than minimal risk for intervention substudy with appropriate safety monitoring
- Confidentiality protections including encryption, de-identification, Certificates of Confidentiality
- Vulnerable populations (minors, psychiatric symptoms) with additional protections
- Incidental findings management plan
- Data Safety Monitoring Board for intervention substudy
- Adverse event reporting procedures

---

5. INCLUSION OF WOMEN AND MINORITIES

[See Section 9 of Research Strategy for detailed discussion]

Summary:

- Expected enrollment: ~80% female (reflects disease epidemiology)
- Target ≥30% racial/ethnic minorities
- Culturally sensitive recruitment
- Analysis of gender and racial/ethnic differences
- Assessment of algorithmic bias

Enrollment Table:

CATEGORY	TARGET ENROLLMENT
Sex/Gender	
Female	280 (80%)
Male	70 (20%)
Race	
White	245 (70%)
Black/African American	42 (12%)
Asian	28 (8%)
Other	35 (10%)
Ethnicity	
Hispanic/Latino	63 (18%)
Not Hispanic/Latino	287 (82%)
Total	**350**

6-11. [Additional Required Sections]

---

6. Vertebrate Animals: Not Applicable

7. Select Agent Research: Not Applicable

8. Multiple PI Leadership Plan: [If applicable, describe co-PI responsibilities and communication plan]

9. Consortium/Contractual Arrangements: [If applicable, describe multi-site coordination]

10. Letters of Support: [Include letters from collaborating institutions, patient advocacy groups, consultants]

11. Resource Sharing Plan:

- De-identified data will be shared via NIMH Data Archive (NDA) after primary publications

- Open-source software tools for vocal biomarker analysis will be published on GitHub
- Study protocols and assessment manuals will be made publicly available

---

APPENDICES

Appendix A: Detailed Budget and Budget Justification

Appendix B: Biographical Sketches (PI and Key Personnel)

Appendix C: Current and Pending Support

Appendix D: Facilities and Other Resources

Appendix E: Equipment

Appendix F: Sample Informed Consent Forms

Appendix G: Letters of Support

Appendix H: Data Safety Monitoring Plan (Intervention Substudy)

END OF RESEARCH PROPOSAL

Total Page Count: [To be determined based on specific grant mechanism requirements]

Contact Information:
Principal Investigator: [Name]
Institution: [Institution]
Email: [Email]
Phone: [Phone]