# Longitudinal Inference of Multiscale Markers in Psychosis: From Hippocampal Centrality to Functional Outcome

Totzek, J.F.[1,2,3], Chakravarty, M. M.[1,2,4], Joober, R.[1,2], Malla, A.[1,2], Shah, J.L.[1,2], Raucher-Chéné, D.[1,2], Young, A.L.[5], Hernaus, D.[6], Lepage, M.[1,2±] & Lavigne, K.M.[1,2,7±*]

[1]Department of Psychiatry, McGill University, Montreal, Quebec, Canada
[2]Douglas Research Centre, Montreal, Quebec, Canada
[3]Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands
[4]Department of Biological and Biomedical Engineering, McGill University, Montreal, Quebec, Canada
[5]Department of Neuroimaging, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
[6]Department of Psychiatry & Neuropsychology, School for Mental Health and NeuroScience MHeNS, Maastricht University, Maastricht, The Netherlands
[7]Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada

± Martin Lepage and Katie Lavigne share senior authorship.

* Corresponding Author: Katie M. Lavigne. F1128.1, 6875 Boulevard LaSalle, Frank B. Common Pavilion, Douglas Research Centre, Montreal, QC, H4H 1R3, Canada. E-mail: katie.lavigne@mcgill.ca.

Word Count: 6538

## Abstract

Psychosis represents a heterogeneous collection of biological and behavioural alterations that evolve over time. We propose a multiscale disease progression model of psychosis, in which hippocampal-cortical dysconnectivity precedes impaired episodic memory and social cognition, worsening negative symptoms and lowering functional outcome. In two cross-sectional datasets of first- and multi-episode psychosis (163 patients; 117 controls), we applied a recently developed machine-learning algorithm, SuStaIn, which uniquely integrates clustering and disease progression modeling. SuStaIn identified three patient subtypes, with Subtype 0 showing normalrange performance on all variables. In comparison, Subtype 1 showed lower episodic memory, social cognition, functional outcome, and higher negative symptoms, while Subtype 2 showed lower hippocampal-cortical connectivity. Subtype 1 deteriorated from (social) cognition to symptoms, functioning and hippocampal-cortical dysconnectivity, while Subtype 2 deteriorated from hippocampal-cortical dysconnectivity to (social) cognition, functioning and symptoms. This first application of SuStaIn in a multiscale model of psychiatry provides distinguishable disease trajectories of hippocampal-cortical connectivity, which might drive heterogeneous behavioural alterations in psychosis.


*Keywords:* hippocampal-cortical connectivity; episodic memory; social cognition; negative symptoms; functional outcome; machine-learning

## Introduction

Schizophrenia and related psychotic disorders can be characterized along the three dimensions of positive, negative and cognitive symptoms (1, 2). Positive symptoms describe abnormal perceptions, thoughts, and behaviors (i.e., hallucinations, delusions, disorganization) while negative symptoms describe a significant reduction of typical behaviors and motivation (i.e., asociality) (1). Cognitive symptoms tend to arise before the onset of positive and negative symptomatology (3, 4) and are characterized by an overall deficit in neurocognitive functioning, and particularly in episodic memory (2). Due to its broad range of symptoms, psychosis is among the most disabling mental disorders and is a leading cause of disability worldwide (5). The most prominent measure of disability caused by psychotic disorders is functional outcome, which assesses the influence of disease on social and occupational functioning (6, 7). Within these two domains of functioning, affected individuals have been shown to have fewer stable friendships, lowered marriage rates and difficulties with social interactions (7). Further, psychosis is associated with high unemployment rates, decreased levels of productivity and an overall increase in occupational stress (7). Beyond the considerable impact on the affected individual and their social surroundings, psychosis also carries an immense economic burden caused by the often-persisting difficulties of maintaining employment following a first-episode of psychosis (FEP) (8, 9). Considering that employment is additionally linked to increased quality of life (10), identifying disease trajectories up to and including poor functional outcomes of psychosis is crucial.

Multiscale neuroscience is a relatively novel framework which conceptualizes psychiatric disorders as the result of aberrant interactions within and across multiple biopsychosocial scales (11, 12). The smallest scale is the micro-scale, which covers the neurotransmitter systems and genetic vulnerabilities that are thought to be involved in disease pathophysiology. The next scale

is the meso-scale, which describes structural and functional brain organization in disease, including aberrant connectivity (i.e., interactions between brain regions). The manifest-scale covers behavioral components of mental illness, and most importantly the cognitive and clinical symptoms (e.g., impaired memory and negative symptoms). At the broadest level is the socialscale, which describes the degree of functional impairment caused by a disorder (12). The key assumption of multiscale neuroscience is that there exist natural bidirectional interactions between these scales, meaning that dysfunction in one scale influences and reflects functionality of higher and lower-order scales (11). In the context of functional outcomes of psychosis, multiscale neuroscience predicts that poor functional outcomes may be driven by aberrant cross-scale interactions that span the brain and behavioral levels. To formally test this idea, we propose a novel disease progression model of psychosis in which brain-level changes result in poor functional outcomes as mediated by specific cognitive and clinical symptoms.

At the meso-scale, graph theoretical measures of connectivity identify the hippocampus as a key convergence zone for cortical input (13). Graph theory as applied to neuroscience conceptualizes the brain as consisting of nodes, representing distinct brain regions, which are connected through edges, representing (structural or functional) connections (14, 15). This technique has further shed light on multiscale interactions between the hippocampus and manifestscale measures in psychosis. By grouping nodes of the hippocampus together into a hippocampal module, Makowski et al. (16) showed that longitudinal structural covariance-based connectivity between the hippocampal module and large-scale brain networks (17) is significantly reduced in FEP in comparison to healthy controls. At the manifest-scale, such a reduction in hippocampalcortical connectivity was associated with more severe negative symptoms, as mediated by impaired episodic memory (16). Using longitudinal data, Makowski et al. (16) thus

characterized a multiscale disease progression from hippocampal-cortical dysconnectivity to episodic memory and negative symptoms for the first time. The association between episodic memory and negative symptoms has been robustly established cross-sectionally and longitudinally across the psychosis spectrum, spanning individuals at clinical high-risk (18), FEP (19) and chronic psychosis (2). Other findings within the manifest-scale have shown that this relationship between episodic memory and negative symptoms is mediated through social cognition (i.e., emotion recognition and theory of mind) (20), providing a link from cognitive impairments to negative symptoms. At the social-scale, sex differences in episodic memory impairments (21) and impaired social cognitive abilities (22, 23) have been shown to predict functional outcome, with both relationships being mediated by negative symptoms. Negative symptoms are further directly associated with impaired functioning (24, 25) and drive symptomatic relapses, which predict poor functional outcome (26).

Based on these findings, we suggest a multiscale model in which hippocampal-cortical dysconnectivity leads to impaired episodic memory and social cognition, resulting in higher negative symptoms and, ultimately, poor functional outcome. However, taking into consideration the immensely heterogeneous nature of psychotic disorders, we recognize that this disease progression pattern may apply to only a sub-population of psychosis patients. Prior cluster analyses showed that the degree of impairment experienced by affected individuals can best be defined along a continuum. While there typically is a patient subgroup which shows significant impairments in episodic memory, social cognition and negative symptoms, there also is a second subgroup of patients showing normal-range performance on those measures (27-30). These normal-range performing patient groups are equally characterized by the presence of positive symptoms (29), justifying their diagnoses, yet the degree to which they express (socio)cognitive

and negative symptoms appears to differ. On these grounds, we i) hypothesize that there may exist at least two patient subtypes, one subtype showing normal-range performance on the components of the model and one subtype showing impairments. In the impaired subtype, we ii) hypothesize a disease progression pattern from hippocampal-cortical dysconnectivity to impaired episodic memory and social cognition, higher negative symptoms and poorer functional outcome. See Figure 1 for a visualization of these hypotheses.

To address the hypothesized disease progression in subtypes of the patient sample, we implemented a recently developed unsupervised machine-learning algorithm called Subtype and Stage Inference (SuStaIn) (31). SuStaIn uniquely combines the methodologies of disease progression modeling and clustering to infer longitudinal disease progressions from crosssectional data. By therefore clustering patient groups with shared common disease progression patterns, SuStaIn is excellently suited to address multiscale frameworks in samples as heterogeneous as psychosis. In the context of SuStaIn, the variables of our models are labeled *markers*, and SuStaIn chooses the disease progression of these markers independently, without the need to *a priori* define disease progression patterns. This allows identification of subgroups and their disease trajectories that fit the data most validly. So far, SuStaIn has been implemented in neurodegenerative diseases, such as Alzheimer's (31, 32), multiple sclerosis (33), frontotemporal dementia (34), Parkinson's disease (35) and once in the field of psychosis (36). In psychosis, SuStaIn identified two disease trajectories of brain atrophy of which one commenced in the hippocampus and another one in the Broca's area (36). To date, there is one other study to our knowledge that applied SuStaIn to a multiscale model (37), identifying four distinct temporal disease progression patterns of brain markers and cognitive functioning in disease subtypes of Alzheimer's. Building on this important

work, our study will, for the first time, implement SuStaIn in a multiscale model of psychiatric disorders.

One of the major clinical utilities of SuStaIn is that its staging approach provides a translational component to our data-driven model by directly allowing for patient stratification and disease prediction. This feature is of particular importance in the light of the current attention that clinical staging models, such as McGorry's clinical stages (38), have been given (39). McGorry et al. (38) hypothesize a unidirectional clinical development of severe mental disorders, in which new stages represent a significant change in clinical status (i.e., from individuals at clinical highrisk to FEP, relapse and persistent psychosis (38, 39)). The key purpose of the clinical staging model is to guide treatment selection based on symptom severity, cognition, and functioning (39). In this work, we sampled across clinical stages by combining data from two independent datasets on first- and multi-episode psychosis. Based on the qualitative differences between clinical stages regarding symptoms, cognition and functioning, it was of further interest to explore how SuStaIn's data-driven and biologically-informed staging would stratify individuals from distinct clinical stages. To this end, we employed z-score SuStaIn (31), which assesses linear disease progression in accordance with z-score deviations from the mean, resembling the assumed unidirectional progression of McGorry's clinical stages. Thus, applying z-score SuStaIn to two samples spanning the clinical staging model (first and multi-episode psychosis) uniquely allowed us to address whether the clinical thresholds by McGorry et al. (38) would correspond to the disease stages outlined by SuStaIn. To avoid any further ambiguity, we will from here onwards refer to McGorry's stages as *clinical* stages and to SuStaIn's stages as *disease* stages (as suggested by Young et al. (31)).

## Methods

### Participants

We sampled patient and control data from two independent datasets. Study 1 collected data in a FEP cohort and Study 2 collected data from a multi-episode psychosis (MEP) cohort (40). For Study 1, 100 patients and 60 non-clinical controls were recruited while Study 2 consisted of 166 patients and 81 non-clinical controls (see Supplement F1 for inclusion and exclusion criteria). Both patient samples were recruited from the Douglas Research Centre, Montréal, Canada. FEP patients were recruited after being admitted to the prevention and early intervention program for psychosis (PEPP-Montréal) (41), while MEP patients were patients of the outpatient and inpatient units. The control samples were recruited from the same catchment area. After excluding participants (see Supplement F2 for the exclusion process), Study 1 consisted of 57 patients and 52 controls and Study 2 consisted of 106 patients and 65 controls, rendering our total sample size 163 patients and 117 non-clinical controls. Ethical approval was granted by the Douglas Research Centre Ethics Board. Both studies were conducted in accordance with the Declaration of Helsinki and written informed consent was obtained prior to the study. Participants were compensated with monetary rewards.

Power calculations in the SuStaIn literature are still an active topic of debate, although previous work has offered pragmatic guidelines for sample size estimation tailored to SuStaIn (32). These guidelines propose that the spatial dimensions (number of markers in the model) and the temporal dimensions (z-scores deviations) are multiplied, resulting in a total number of features. For each of these features, 10-20 observations should be included in the analysis. As we chose 6 markers (left and right hippocampal-cortical connectivity, episodic memory, social

cognition, negative symptoms, functional outcome) and 3 levels of z-scores (1, 2, 3), a sample size between 180 - 360 should therefore be sufficient to detect meaningful subtypes differences and their disease progressions.

**Data Collection**

In both studies, data including sex, age, duration of illness and medication were assessed throughout an interview with trained research staff. Episodic memory and social cognition were measured via the Cogstate Schizophrenia Battery (42), including the Shopping List Test and Social-Emotional Cognition Test. Negative and positive symptoms were assessed using the Scale for the Assessment of Negative Symptoms (SANS) (43) and the Scale for the Assessment of Positive Symptoms (SAPS) (44), respectively, while functional outcome was measured through the Social and Occupational Functional Assessment Scale (SOFAS, (6)). See Supplement S1 for scale and scoring descriptions. As the SOFAS was not administered in MEP patients of Study 2, SOFAS scores were reconstructed (see Supplement S2 and F7 for the reconstruction procedure).

**MRI Acquisition**

MRI data were acquired with a 3T Siemens Magnetom Trio scanner, located at the Cerebral Imaging Centre of the Douglas Mental Health University Institute. For Study 1, a T1-weighted MPRAGE scan (voxel size = 1$mm^3$, field of view (FOV) = 256mm, repetition time (TR) = 2300ms, echo time (TE) = 2.98ms, flip angle = 9°) was obtained, followed by a high-resolution T2-weighted image (voxel size = 0.64$mm^3$, FOV = 206mm, TR = 2500ms, TE = 1.98ms). For the T2-weighted images a turbo spin echo sequence was used. Study 2 included a whole brain T1-weighted scan (voxel size = 1$mm^3$, TR = 18ms, TE =10ms, flip angle=30°) for which a flow-compensated 3D RF-spoiled GRE sequence was used.

## MRI Processing

We used the high-performance computing resources available via the Digital Research Alliance of Canada (formerly Compute Canada) to extract cortical thickness and hippocampal volumes. See Figure 2 for a visualization of the processing steps. First, cortical thickness values were derived through CIVET, Version 2.1.1 (45), as run on CBRAIN (46), which were subsequently parcellated into the 62 Desikan-Killiany-Tourville (DKT) regions (47). To extract hippocampal volumes, we submitted the acquired T1-informed T2-weighted (Study 1) or T1weighted (Study 2) MRI images to the multiple automatically generated templates (MAGeT)Brain algorithm (48, 49), resulting in nine hippocampal and adjacent white matter regions per hemisphere. Following data extraction, we excluded participants with poor quality control values on CIVET and/or MAGeT and further performed data harmonization via neuroCombat (50) to account for different scanner parameters between studies (see Supplement S3 for a detailed explanation of these processing steps).

## Hippocampal-Cortical Connectivity

### *Structural Covariance & Jackknife Bias Estimation Procedure*

To establish the structural covariance between the 80 brain regions (62 DKT and 18 hippocampal), we performed full-sample correlations for the patient and the control sample separately (14). To further provide information about subject-specific contributions to structural covariance we then performed the jack-knife bias estimation procedure (51, 52) on both matrices. The jack-knife bias estimation procedure recalculates the structural covariance matrix for the sample while leaving each participant out of the calculation once (leave-one-out procedure). By subtracting this leave-one-out structural covariance matrix from the full sample covariance matrix of the patient/control sample and taking absolute values, the individual's contribution to structural

covariance is computed, resulting in a subject-specific structural covariance matrix. The code has been made publicly available on GitHub (https://github.com/katielavigne/jackknife_connectivity).

### *Participation Coefficient*

The participation coefficient is a graph measure of centrality and evaluates the intermodular connectivity between nodes of one module and nodes of other modules (53). To establish the modules, we parcellated the cortical thickness values of the 62 DKT regions into the seven functional Yeo networks (17) while the volumes of the 18 hippocampal regions were grouped into one distinct hippocampal module, as first done by Makowski et al. (16). The participation coefficient was then calculated via the Brain Connectivity Toolbox in Python (bctpy; (53)). Higher participation coefficients indicate that the intermodular connectivity of a node is higher than the intramodular connectivity (53). For example, a high participation coefficient of a node within the hippocampal module (e.g., dentate gyrus region) would indicate that this region is more strongly connected to the cortical networks than to other regions of the hippocampus. For our analysis, the participation coefficients of the hippocampal regions were averaged for each hemisphere respectively; thus, higher values represented stronger lateralized hippocampal-cortical connectivity.

## Statistical Analysis

### *Data Preparation*

Demographic data were compared through chi-square tests and independent samples *t*tests. We then performed a linear regression with sex, age and total brain volume as regressors on the input data. For the variables hippocampal-cortical connectivity, episodic memory and social cognition, regression was performed in relation to the control values. The negative symptom and

functional outcome variables were not sampled in the control groups of either dataset. Thus, the data was first z-scored in comparison to previously reported control values of the SANS (54) and the SOFAS (55), after which the values were regressed out relative to the patient samples themselves. Z-scores for the entire sample were calculated based on these residual values. All zscores except for negative symptoms were multiplied by -1 to model ascending disease progression, as required for SuStaIn.

***SuStaIn***

Z-score SuStaIn is a type of SuStaIn which places subjects at specific disease stages depending on their z-score deviation from normality (31). Through an estimation of maximum likelihood, SuStaIn assesses the temporal sequence of marker deterioration and employs purely cross-sectional data to infer common patterns of disease progression via a stage progression model. Z-score SuStaIn (31) was performed in python (pySuStaIn (56)). We set the threshold of z-scores to 1, 2 and 3 deviations from the mean for each biomarker, while reaching a maximum threshold of 5 z-scores at the end of the disease progression as in previous work (31, 57). Considering the absence of symptom and functional outcome data in controls, we ran SuStaIn on the patient data only. We then validated the resulting subtypes using 10-fold Markov Chain Monte Carlo (MCMC) iteration cross validation. Log-likelihood tests and the Cross-Validation Information Criterion (CVIC, (31)) were used to obtain an indication of the optimal number of subtypes. The code for the steps of data preparation and SuStaIn has been made publicly available on GitHub (https://github.com/janatotzek/2023-multiscale-markers-psychosis).

***Follow-up Analyses***

Demographic data were compared between the resulting disease subtypes through chisquare and independent samples *t*-tests. To address how the identified subtypes differed on the

markers, a one-way ANOVA with Bonferroni-corrected independent sample *t*-tests were conducted as follow-up analyses. A significance threshold of alpha = 0.05 (two-tailed) was used. To establish marker deterioration, we compared the marker means for the resulting disease subtypes to the means of Subtype 0, which per definition included patients at disease stage 0 (31). We further compared the mean differences between disease subtypes to establish similarities and differences across disease subtypes. Stage inference was then performed by calculating at which disease stage the subtype mean reached *n* (1, 2, and 3) z-score deviations from the mean. These results were supported by visually examining the positional variance diagrams, indicating the likelihood of participants to deviate *n* z-scores from the mean at a specific disease stage of each disease subtype. A section on subtype and stage inference was further added to visualize the disease progression in each identified disease subtype based on mean values per subtype stage. Finally, to address distinctions between FEP and MEP, we split both samples into low versus high negative symptoms based on a mean split of z-scored negative symptoms which we calculated for each group (FEP and MEP) separately. Such a mean split allowed us to distinguish currently symptomatic (negative symptoms) vs. stable FEP and MEP patients and to address potential similarities and differences across both samples.

## Results

### Demographic Data

A summary of demographic characteristics can be seen in Table 1. See Supplement F3 for a visualization of age and illness duration in both datasets.

## Subtype Inference

SuStaIn sorts all individuals which are assigned to disease stage 0 into a separate Subtype 0, which is characterized by normal-range performance on the included markers. As our sample exclusively consists of psychosis patients, Subtype 0 thus consists of patients which, per definition, do not show z-score-based impairment on the included markers. To therefore differentiate Subtype 0 from other patient subtypes which do show impairment, we will from here on refer to all subtypes except Subtype 0 as *disease* subtypes. We ran SuStaIn on a maximum of two disease subtypes, which resulted in a CVIC of 3342.68 (log-likelihood of –169.09) for one disease subtype and a CVIC of 3342.80 (log-likelihood of – 169.16) for two disease subtypes. Typically, lower CVIC scores suggest a better model fit (31), yet differences between models which are < 6 are not considered to be meaningful in the SuStaIn literature (58). As our results therefore seem to suggest a suitable fit of both models, we report the results for a model with two disease subtypes below (see Supplement S4 for results of the model with one disease subtype). In the model with two disease subtypes, 56 participants were categorized as belonging to Subtype 0. Of those 56 participants, 26 were FEP patients and 30 were MEP patients. 86 participants were classified as belonging to Subtype 1, of whom 26 were FEP patients and 60 were MEP patients. 21 subjects were classified as belonging to Subtype 2, of whom 5 were FEP patients and 16 were MEP patients.

The one-way between-subject ANOVA showed a significant main effect of Subtype on left ($F(2, 160) = 65.42$, $p < .001$, $\eta_p^2 = 0.45$) and right hippocampal-cortical connectivity ($F(2, 160) = 35.57$, $p < .001$, $\eta_p^2 = 0.31$). Post-hoc independent samples *t*-tests revealed that Subtype 2 exhibited significantly lower bilateral hippocampal-cortical connectivity in comparison to Subtype 0 and Subtype 1. In contrast, Subtype 0 and Subtype 1 did not differ significantly on those measures, indicating that only patients in Subtype 2 were characterized by impaired hippocampal-cortical

connectivity when collapsing across stages. The main effect of Subtype on episodic memory was also significant ($F(2, 160) = 76.36$, $p < .001$, $\eta_p^2 = 0.49$) and follow-up analyses showed that Subtype 1 was characterized by significantly lower episodic memory performance than Subtype 0 and Subtype 2, while Subtype 0 and Subtype 2 did not significantly differ in episodic memory performance. The same pattern was observed for social cognition ($F(2, 160) = 15.12$, $p < .001$, $\eta_p^2 = 0.16$) and negative symptoms ($F(2, 160) = 20.09$, $p < .001$, $\eta_p^2 = 0.20$), suggesting that Subtype 1 was impaired on memory, social cognition and negative symptoms. Regarding functional outcome ($F(2, 160) = 7.22$, $p = .001$, $\eta_p^2 = 0.08$), Subtype 1 was significantly impaired in comparison to Subtype 0, yet did not differ from Subtype 2, indicating that Subtype 1 scored lowest on functional outcome. See Figure 3 for a visualization of the follow-up independent samples *t*tests and Table 2 for demographic characteristics and statistical comparisons of the subtypes. Beyond the markers of our model, all three Subtypes did not differ on medication and illness duration, while Subtype 1 was characterized by significantly higher positive symptoms than Subtype 0.

**Stage Inference**

As Subtype 0 per definition implies no disease progression, SuStaIn exclusively infers disease progressions for the disease subtypes. For Subtypes 1 and 2, data are provided up to disease stage 9 and 6 in our sample respectively. Longitudinal inference as performed through SuStaIn then infers biomarker progression up until disease stage 18 (see Figure 4). Subtype 1 shows an early deterioration of episodic memory, which deviates one z-score from the mean at disease stage 1, reaching two z-score deviations at disease stage 2 and three z-score deviations at disease stage 6. The second marker to deviate is social cognition, reaching one z-score deviation

from the mean at disease stage 3, two z-scores at disease stage 4 and three z-scores at disease stage 5. Negative symptoms and functional outcome then both reach one z-score deviation from the mean at disease stage 7, yet negative symptoms deteriorate faster by reaching two z-score deviations at disease stage 9. Negative symptoms and functional outcome then do not reach further z-score deviations in our sample. Left hippocampal-cortical connectivity reaches one zscore deviation from the mean at disease stage 9, while right hippocampal-cortical connectivity never substantially deviates from the mean in our sample. Nevertheless, the longitudinal inference of SuStaIn shows that it most likely starts deteriorating from disease stage 12 onwards, which is beyond the stages represented in our sample. Therefore, individuals in Subtype 1 progress from episodic memory to social cognition, negative symptoms, functional outcome, left hippocampalcortical connectivity and right hippocampal-cortical connectivity (see Figure 4a, Figure 5).

Subtype 2 (Figure 4b) shows a deterioration of bilateral hippocampal-cortical connectivity by one z-score deviation from the mean at disease stage 1, with left connectivity reaching two zscore deviations at disease stage 3 and right connectivity at disease stage 4. Neither deteriorate further in our sample. The z-score means per stage (Figure 5) indicate that social cognition reaches one z-score deviation at disease stage 3, but then the mean scores diminish back to around 0 in the following disease stages. Episodic memory starts deviating from the mean by one z-score at disease stage 4 and reaches 3 deviations at disease stage 6. Given that social cognition and episodic memory therefore deteriorate at around the same stage, yet episodic memory continues to deteriorate while social cognition scores decrease again, we mention episodic memory before social cognition in our disease progression. Negative symptoms and functional outcome do not deviate substantially from the mean in our sample, yet the longitudinal inference of SuStaIn

(Figure 4b) shows that functional outcome starts deviating earlier than negative symptoms. Altogether, Subtype 2 progresses from left hippocampal-cortical connectivity to right hippocampal-cortical connectivity, episodic memory and social cognition, functional outcome, and negative symptoms.

**Subtype and Stage Inference**

As visualized in Figure 5, SuStaIn identifies two disease subtypes in addition to Subtype 0. Subtype 1 is predominantly characterized by a deterioration of episodic memory, social cognition, functional outcome and higher negative symptoms and progresses from episodic memory to social cognition, negative symptoms, functional outcome, left hippocampal-cortical connectivity and right hippocampal-cortical connectivity. Subtype 2 is characterized by hippocampal-dysconnectivity and progresses from left hippocampal-cortical connectivity to right hippocampal-cortical connectivity, episodic memory and social cognition, functional outcome and negative symptoms.

**First- and Multi-Episode Psychosis**

When comparing demographic characteristics between both datasets, Table 1 shows that MEP patients are older and have significantly higher illness duration, medication dosage, total brain volume, negative and positive symptoms and significantly lower episodic memory and functional outcome than FEP. However, FEP and MEP patients are distributed equally across subtypes and disease stages (see Figure 5, right panels). When splitting the MEP and FEP sample into low versus high negative symptoms based the respective mean z-score of each sample, we see that FEP and MEP patients who score low on negative symptoms (FEP: n = 30; MEP: n = 53) do not differ significantly from each other on any of the markers (see Figure 6a). For FEP and MEP

with high negative symptoms (FEP: n = 27; MEP: n = 53), however, there are significant differences in functional outcome with MEP high showing significantly more impaired functional outcome than all other groups. In general, MEP patients with high negative symptoms show significantly lower episodic memory than FEP with low negative symptoms, and significantly lower functional outcome than any of the other groups. In addition, right hippocampal centrality is significantly higher in MEP with low negative symptoms than FEP with high negative symptoms. Particularly, the findings of episodic memory and functional outcome suggest that both markers are overall more affected in MEP versus FEP but that there are nevertheless MEP individuals who score comparably low to FEP. The distribution of high vs. low FEP and MEP across disease stages is seen in Figure 6b and 6c, visualizing that individuals with low negative symptoms tend to be grouped in earlier disease stages while individuals with high negative symptoms are distributed across all stages and dominate the later disease stages.

**Difference to One-Subtype Model**

In a model with one disease subtype, Subtype 1 is characterized by significantly lower episodic memory, social cognition, higher negative symptoms, and poorer functional outcome in comparison to Subtype 0 (see Supplements S4, T1, F5) and progresses from episodic memory to social cognition, negative symptoms, functional outcome, left hippocampal-cortical connectivity and right hippocampal-cortical connectivity (see Supplement F6). These findings are congruent with Subtype 1 as identified in a model with two disease subtypes (described above). When comparing the results of both models (see Supplement S4), Subtype 2 of the two-subtypes model is equally distributed between Subtype 0 and Subtype 1 of the one-subtype model (see Supplement F4). Considering that Subtype 2 of the two-subtypes model shows marked deterioration of bilateral hippocampal-cortical connectivity, which was seen in neither Subtype 1 of the one- or two-

subtypes models, these deficits appear to be masked in the one-subtype model. While both models show a statistically similar fit for our data (as described under "Subtype Inference"), the twosubtypes model seems to uncover the importance of hippocampal-cortical connectivity in a part of the patient sample to predict cognition and symptom development.

## Discussion

We set out to advance our understanding of poor functional outcomes of psychosis by proposing a multiscale model (spanning the brain, cognition, symptoms and functioning), and exploring heterogeneous disease trajectories across these markers. To these aims, we employed zscore SuStaIn to evaluate our multiscale model in a heterogeneous sample of first-episode and multi-episode psychosis patients. SuStaIn is a recently developed machine-learning algorithm which uniquely combines clustering and disease progression modeling (31). To our knowledge, this work is the first to implement SuStaIn in a multiscale model of psychiatric disorders. We identified two *disease* subtypes in addition to a patient subtype without impaired connectivity, (social) cognition, symptoms, or functioning. While both subtypes supported our proposed progression model in terms of episodic memory, social cognition, negative symptoms, and functional outcome, hippocampal-cortical connectivity only preceded these markers, consistent with our hypothesis, in Subtype 2. In Subtype 1, hippocampal-cortical connectivity deficits emerged *following* cognitive impairments and symptoms, rather than preceding them. This suggests that hippocampal-cortical dysconnectivity can but does not necessarily precede cognitive impairments. These results demonstrate how multiscale frameworks can offer important insights into the complex temporal nature of cross-scale interactions, while underlining the need to differentiate between patient subtypes in disorders as heterogeneous as psychosis.

**Longitudinal Inference in Psychosis**

In previous work, Makowski et al. (16) reported that reduced hippocampal-cortical connectivity predicts negative symptoms as mediated through episodic memory. Surprisingly, our results show that Subtype 1 exhibits higher negative symptoms and impairment in episodic memory, while Subtype 2 shows distinct deteriorations of hippocampal-cortical connectivity. However, none of the identified disease subtypes showed an impairment on all three markers simultaneously. A potential explanation for this discrepancy might lie in Makowski et al. (16) performing their analyses on the full sample while we made a specific effort to cluster disease subtypes. Performing analyses on the full sample might have resulted in obscured heterogeneity between the clusters, bearing the potential of pooling the impairments of hippocampal-cortical connectivity, episodic memory and negative symptoms across subtypes. Moreover, the previous study included longitudinal data, which is not yet possible with SuStaIn, and may have captured impairments at later timepoints. These findings are of particular importance when considering that the sample assessed in Makowski et al. (16) partially overlaps with the sample of Study 1.

An additional aspect to be addressed regarding the previous mediation model is the temporal disease progression from hippocampal-cortical connectivity to negative symptoms as mediated by episodic memory. On these grounds, we hypothesized hippocampal-cortical connectivity to precede impaired cognition and symptoms, which we found in Subtype 2. However, the identified disease trajectories of Subtype 1, in which hippocampal-cortical dysconnectivity *follows* deteriorations in memory and social cognition, hints towards the existence of other neural markers that drive socio-cognitive deficits in psychosis. In addition to other metrics of hippocampal-cortical connectivity (e.g., based on white matter structure or function), neural correlates of social cognition could be key, such as the amygdala (59), for which dorsomedial shape development has been shown to predict conversion to psychosis (60), while further being

associated with verbal memory performance (61) and low emotion recognition capacities (60). Beyond the association with social cognition, the amygdala and hippocampus have further been shown to be distinctly associated with negative symptoms (62). Thus, future work aiming to disentangle multiscale contributions to impaired social cognition and episodic memory may wish to consider the independent contributions of the amygdala and hippocampus as well as other measures of brain connectivity.

We further identified an almost identical temporal development of negative symptoms and functional outcome in both disease subtypes. One explanation for this finding might lie in the pooling of SOFAS social and occupational functioning scores (6), which are used estimate overall functional outcome. With regard to the social functioning component, premorbid social functioning in particular has been shown to be predictive of negative symptom severity in later disease stages (63, 64). While premorbid social functioning and social outcomes of psychosis are two distinct constructs, considering premorbid functioning as a potential confounder may help to further disentangle the relationship between negative symptoms and functional outcomes. At this point, we would like to point out that our key objective was to propose a multiscale perspective on functional outcomes in psychosis, providing a link from the brain to functional outcomes through socio-cognitive deficits and symptoms. We did not intend to provide an exhaustive account of all factors influencing functional outcomes of psychosis (i.e., premorbid social functioning; IQ (65), etc.) and recommend future research considers these aspects in addition to the components of our model.

Our findings further need to be discussed in the light of clinical staging models. McGorry et al. (38) hypothesize a unidirectional disease progression, in which new clinical stages represent a significant change in clinical status (i.e., from FEP to MEP; (38, 39)). By including patients from

two distinct clinical stages, FEP and MEP, and applying a machine-learning algorithm which models linear disease progression regardless of clinical staging, we were in the unique position to address whether these clinical stages can be mapped onto the disease stages identified by SuStaIn. As McGorry suggests, MEP, as a group, showed lower episodic memory and functioning and higher negative symptoms than FEP. Yet, when applying the longitudinal inference feature of SuStaIn, FEP and MEP participants were spread equally across disease subtypes and, most importantly, across disease stages. While this seems contradictory to McGorry's clinical stages at first sight, it is important to differentiate between clinical *stages* and disease *states*. As outlined in the introduction, clinical *stages* carry qualitative value about patient history and inform treatment decisions. Disease *states*, however, are the patient's scores on symptom severity, functioning and cognition at a specific point in time. SuStaIn therefore infers longitudinal disease stages from patient *states* which might explain why the clinical stages of McGorry and the disease progression as outlined by SuStaIn do not necessarily overlap. More precisely, z-score SuStaIn assesses z-score deviations from the mean. The algorithm thus places individuals with good *state* performance (i.e., low negative symptoms, high functioning, good episodic memory performance at assessment) in early disease stages and individuals with poor *state* performance (high negative symptoms, low functioning, poor episodic memory performance at assessment) in later disease stages. The assumption which underlies this modeling is the one of linearity, implying that individuals who currently have good state performance will later develop poorer state performance, with the degree of impairment and the order of deterioration differing between disease subtypes. By showing that MEP and FEP are placed in the same SuStaIn stages, we show that *states* across MEP and FEP are comparable, which demonstrates that states can vary significantly *within* clinical stages. Our results could therefore be considered a natural extension of McGorry's clinical stages, potentially

explaining clinical variation seen within clinical stages. In this sense, our findings might contribute to a refining of stage-based frameworks by allowing for a fusion between clinical staging models and biomarker-informed disease trajectories.

An additional point to consider is the cyclical nature of psychotic symptomatology. Z-score SuStaIn is well-suited to mimic the linearity of the clinical staging model of McGorry et al. (38); however, progression patterns might deviate across the scales of our model. When, for example, looking at markers such as negative symptoms, prior research has shown that symptom severity might not increase over the course of disease progression, but can fluctuate in a subsample of patients (66). The fact that low and high negative symptoms are associated with distinct degrees of functional impairment and episodic memory speaks in favor of this hypothesis, by showing that the markers might fluctuate depending on symptomatic vs. remitted illness. Linking this back to our findings, SuStaIn likely sorted MEP patients who were in remission into early disease stages, as these remitting patient scores had similar mean scores to patients in early clinical stages (FEP). To explore this further and to better capture intra-individual fluctuations within clinical stages, future work should evaluate other disease progression models, such as the kernel density estimation model in event-based SuStaIn (35, 37, 67), which do not assume linear disease progression but rather assess the presence or absence of symptoms. Such an approach might allow for a more reliable modeling of McGorry's clinical stages while acknowledging the variety of disease *states* within each clinical stage.

Our findings have important implications for multiscale perspectives on psychosis. To our knowledge, our work addresses a multiscale model of psychiatric symptoms using SuStaIn for the first time. Notably, our findings show that, even within multiscale frameworks, the components of

clustering in combination with disease progression modeling are essential to fully capture the complex nature of psychosis. By implementing a mechanistic approach towards psychosis disease progression that integrates multiple levels of measurement, we therefore address the criticism of the current diagnostic classification systems which predominantly focus on symptom-based diagnostics (1). As z-score SuStaIn assigns individuals to distinct disease subtypes and stages, our results also provide clinical utility and translational opportunities for patient stratification and diagnostics based on machine-learning. A good example of clinical utility beyond the implications for McGorry's clinical staging model (38) is the application of SuStaIn in psychosis by Jiang et al. (36), who found that one of their disease subtypes was related to higher antipsychotic medication efficacy. While we did not find any significant group differences in terms of medication, future studies in larger samples might benefit from this utility. Considering that hippocampal volume and episodic memory are also highly valuable in major depressive disorder (68), our findings might further allow for transdiagnostic applications predicting disease progression.

Our study has several limitations which should be considered. The SOFAS scores of Study 2 were reconstructed on the basis of the Personal and Social Performance (PSP) Scale (69), which is a novel procedure. While we have statistical support for this method (see Supplement S2 and F7), we nevertheless suggest the replication of this methodology in a larger sample. Further, the SOFAS (6) and SANS (43) were not sampled for the control group of both studies. For this reason, the SOFAS and the SANS data had to be z-scored relative to previously reported control data (54, 55) and our SuStaIn analysis had to focus on patient data only, which rendered Subtype 0 our reference for follow-up analyses even though Subtype 0 in itself also consisted of patients. Through the future inclusion of a control group, which would most likely be almost entirely subtyped as Subtype 0 (31), it would be possible to compare the deterioration of markers to a more reliable

subtype. Having a control group which consists of non-clinical controls would further allow us to determine whether patients in Subtype 0 do factually show normal-range performance on the markers of the model, or whether they do show deviations from the norm. If deviations from the norm were seen on a continuum (with non-clinical controls and Subtype 1 & 2 as the extremes and Subtype 0 with intermediate scores), our results might be interpreted as Subtype 0 individuals merging into Subtype 1 or Subtype 2 at later disease stages. Additionally, we need to acknowledge the relatively small sample of our study, which only consists of 163 patients. In accordance with the guidelines for SuStaIn power estimates described above and in (32), our results may therefore be slightly underpowered, and we suggest a replication of our findings in a larger sample.

To conclude, as the first investigation of multiscale disease progression in psychosis using SuStaIn, we identified two data-driven disease subtypes with distinct longitudinal disease trajectories. These findings are of particular importance for elucidating the complex and heterogeneous nature of psychosis even within multiscale models and further contribute to the neurobiological underpinnings of episodic memory in predicting poor functional outcomes. In addition, these are the first data-driven findings to complement traditional clinical staging models by inferring longitudinal disease trajectories across clinical stages. Further research should target broader measures of hippocampal-cortical connectivity through combining distinct neuroimaging modalities and implement machine-learning models which do not assume linear disease progression (e.g., event-based SuStaIn).

## Acknowledgements

We would like to thank all participants who took part in the two studies which contributed to our project. We would also like to thank PEPP-Montréal and the CRISP Research Lab of the Douglas Research Centre for participant recruitment and data collection. We would also like to thank Caroline Dakoure and Joshua Unrau for their support with data preparation, Karyne Anselmo for her help with reconstructing the SOFAS, and Alan C. Evans for early methodological guidance. We used the high-performance computing resources available via the Digital Research Alliance of Canada (https://alliancecan.ca/en).

## Funding

Study 1 was funded by the Canadian Institutes of Health Research in collaboration with the Otsuka Lundbeck Alliance (#141636), while Study 2 was funded by the Canadian Institutes of Health Research (#106434) and the Otsuka/Lundbeck Alliance (#20135257). The funding agencies did not influence the study design, data collection or writing of the manuscript.

## Author Contribution

Conceptualization: ML, KML, JFT. Supervision: ML, KML, DH. Writing – Original Draft: JFT. Formal analysis: JFT. Writing – Review & Editing: JFT, KML, RJ, AM, JLS, ML, DH, ALY, DR. Software: MMC, ALY. Methodology: ALY. Visualization: JFT. Funding Acquisition: MMC, RJ, AM, JLS, ML.

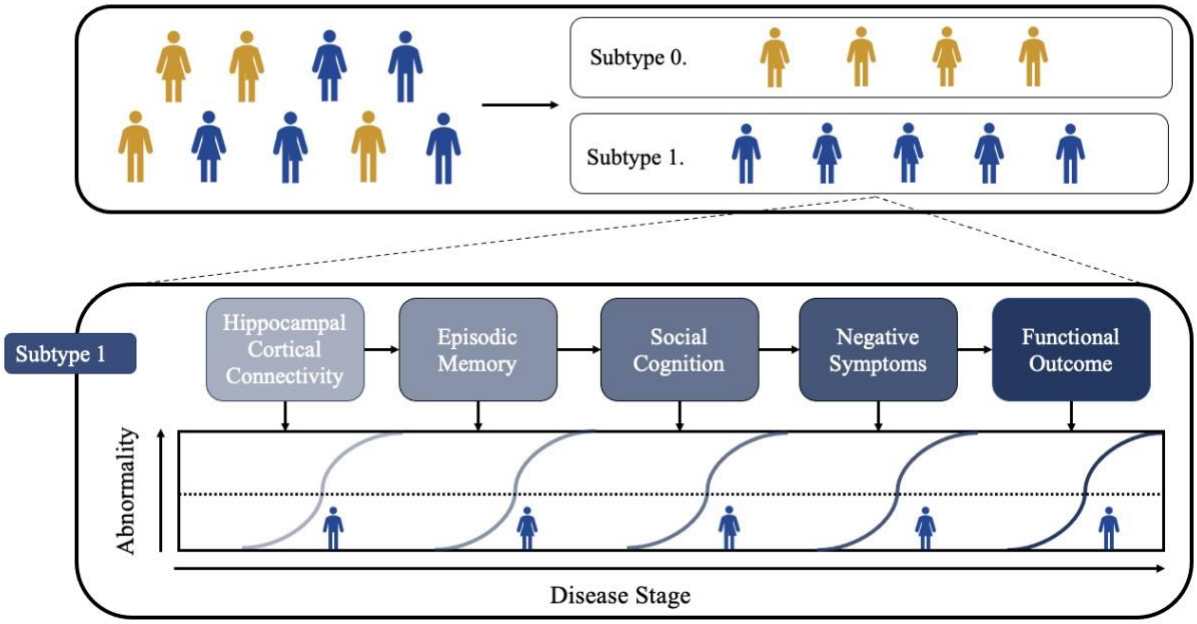
## Competing Interests

ML holds salary awards through the James McGill Professorship, the Canadian Institute of Health Research and the Fonds de recherche du Québec – Santé and reports grants from Otsuka Lundbeck Alliance, diaMentis, Hoffman-La Roche, personal fees from Lundbeck Canada, personal fees from

Otsuka Canada, grants and personal fees from Janssen and personal fees from MedAvanteProphase outside the submitted work. KML reports salary awards through the Mitacs Accelerate fellowship in partnership with Otsuka Canada and personal fees from Otsuka Canada and Lundbeck Canada. DH has received financial compensation as a consultant for P1vitalProducts Ltd. JLS holds a salary award from the Fonds de recherche du Québec – Santé. AM reports receipt of grants, fees or honoraria from Lundbeck and Otsuka and salary awards by the Canada Research Chairs program. MMC holds salary awards from the Fonds de recherche du Québec – Santé and reports funding from the Canadian Institute of Health Research, the Natural Sciences and Engineering Research Council of Canada, the Weston Brain Institute, Healthy Brains Healthy Lives and the Fonds de recherche du Québec – Santé. R.J. served as member of advisory board committees and speaker for Bristol Myers Squibb, Pfizer, Sunovian, Janssen, Myelin and Associates, Lundbeck, Otsuka, Shire, and Perdue, and received grants from Janssen, Otsuka, Lundbeck, Bristol Myers Squibb, Astra Zeneca, and HLS Therapeutics Inc. ALY was supported by a Skills Development Fellowship (MR/T027800/1) from the Medical Research Council and a Career Development Award from the Wellcome Trust [227341/Z/23/Z]. This research was funded in whole, or in part, by the Wellcome Trust [227341/Z/23/Z]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. All of these disclosures are unrelated to the present study. All other authors have nothing to disclose.

## Figure 1.

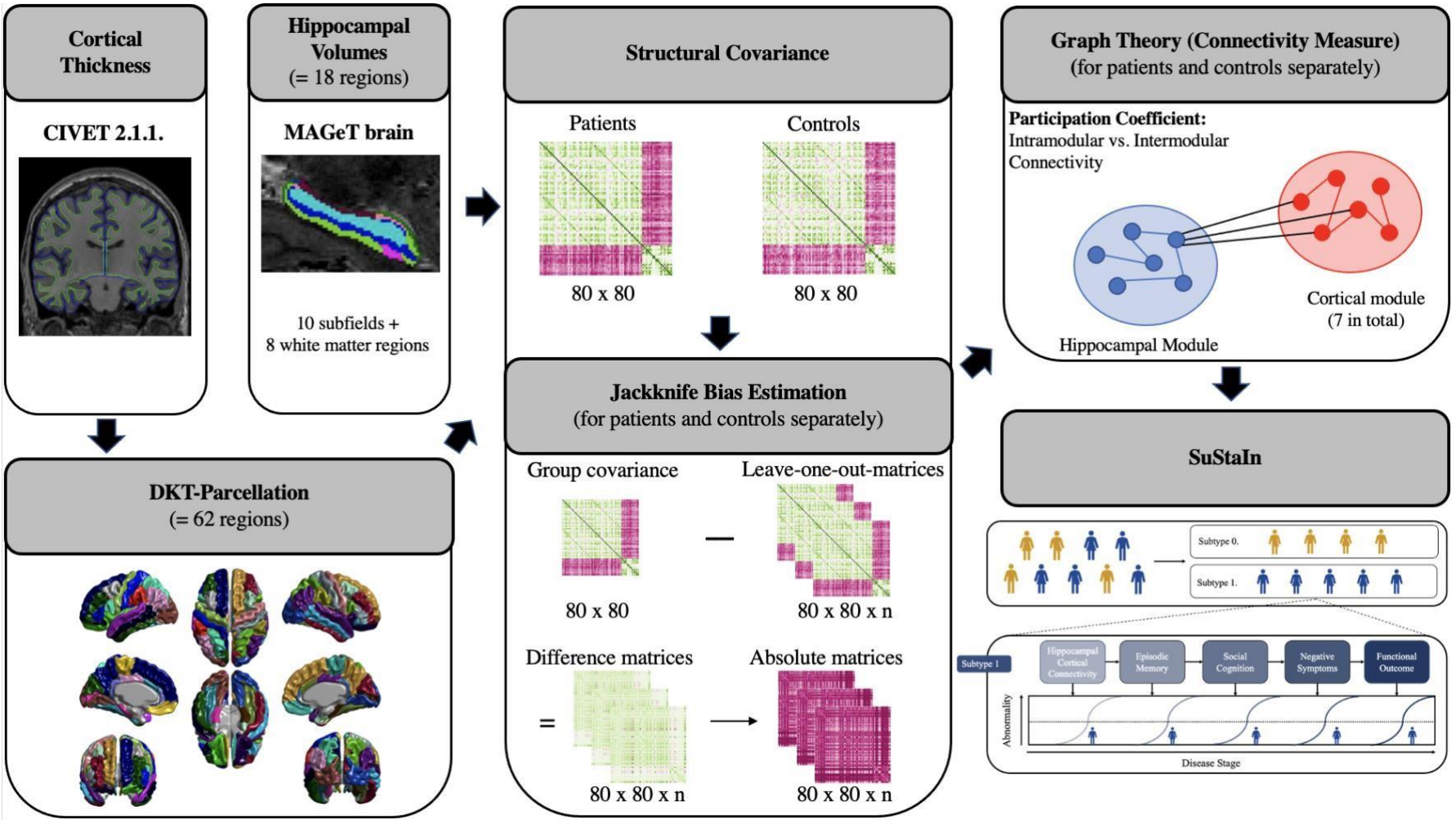
*Visualization of Hypothesized Multiscale Model*




*Note.* This figure displays the subtyping and staging of our proposed multiscale model. We expected to identify at least two subtypes in our sample. In the impaired subtype (Subtype 1), we hypothesized a disease progression pattern in accordance with the proposed multiscale model (hippocampal-cortical dysconnectivity to impaired episodic memory to social cognitive deficits to negative symptoms to poor functional outcome). The curves indicating the marker development serve as a visualization of disease progression and are not a representative delineation of each marker's individual development.

**Longitudinal Inference in Psychosis**
**Figure 2.**

*Methods Workflow*




*Note.* This figure displays the workflow of data processing, calculation of hippocampal-cortical connectivity and the application of SuStaIn.

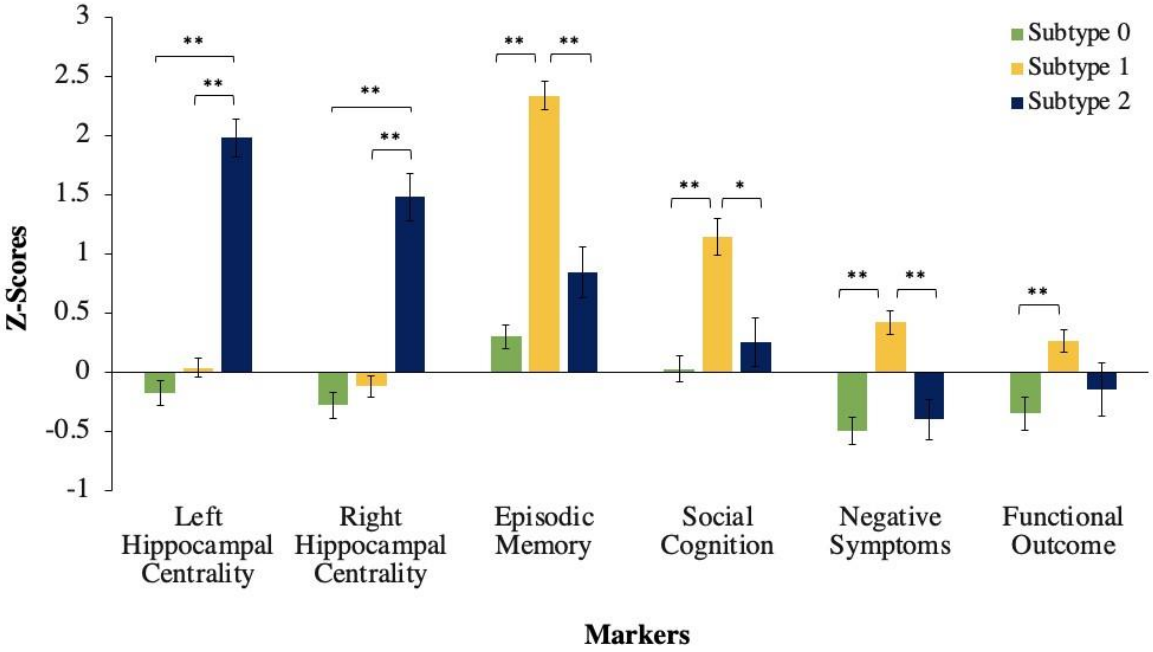
**Table 1.**
*Demographic Characteristics*

| | Study 1. | | Study 2. | | Full Dataset | |
|---|---|---|---|---|---|---|
| | **Contr. (n = 52)** | **FEP (n = 57)** | **Contr. (n = 65)** | **MEP (n = 106)** | **Contr. (n = 117)** | **Pat. (n = 163)** | | |
| **Sex** (male/female) | (34 / 18) | (36 / 21) | (45 / 20) | (80 / 26) | (79 / 38) | (116 / 47) |
| | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** |
| **Age** | 24.70 (4.30) | 25.02 (4.14) | 33.77 (8.93)$^{c***}$ | 35.24 (8.26)$^{c***}$ | 29.73 (8.52) | 31.66 (8.61) |
| **TBV**[a] | 1.32 (0.13) | 1.30 (0.12) | 1.54 (0.05)$^{c***}$ | 1.49 (0.06)$^{b***c***}$ | 1.44 (0.14) | 1.42 (0.12) |
| **Left Hippocampal Centrality** | 0.80 (0.04) | 0.80 (0.04) | 0.81 (0.03) | 0.79 (0.04)$^{b*}$ | 0.80 (0.04) | 0.79 (0.04) |
| **Right Hippocampal Centrality** | 0.79 (0.03) | 0.80 (0.03) | 0.80 (0.04) | 0.80 (0.03) | 0.80 (0.03) | 0.80 (0.03) |
| **Episodic Memory** | 19.00 (2.36) | 15.63 (3.12)$^{b***}$ | 17.65 (2.78)$^{c**}$ | 13.85 (3.52)$^{b***c**}$ | 18.25 (2.68) | 14.47 (3.48)$^{d***}$ |
| **Social Cognition** | 1.14 (0.17) | 1.05 (0.19)$^{b*}$ | 1.14 (0.14) | 1.03 (0.21)$^{b**}$ | 1.14 (0.15) | 1.04 (0.20)$^{d**}$ |
| **Negative Symptoms** | 2.58 (3.13)[e] | 12.74 (8.96)$^{b***}$ | 2.58 (3.13)[e] | 22.15 (9.88)$^{b***c***}$ | 2.58 (3.13)[e] | 18.86 (10.55)$^{d***}$ |
| **Functional Outcome** | 84.00 (5.16)[e] | 57.11 (17.76)$^{b***}$ | 84.00 (5.16)[e] | 45.69 (13.04)$^{b***c**}$ | 84.00 (5.16)[e] | 49.68 (15.78)$^{d***}$ |
| **Positive Symptoms**[f] | - | 12.39 (13.88) | - | 18.18 (17.71)$^{c*}$ | - | 16.18 (16.67) |
| **Illness Duration (years)**[f] | - | 0.73 (0.93) | - | 13.08 (7.82)$^{c***}$ | - | 9.64 (8.67) |
| **CPZ Adherence**[g] | - | 285.23 (236.30) | - | 636.91 (600.14)$^{c**}$ | - | 479.26 (502.92) |

*Note.* [a]TBV = Total Brain Volume, reported in million decimals, [b] $p < 0.05$ when comparing controls vs. patients within each dataset, [c] $p < 0.05$ when comparing patients between datasets and controls between datasets, [d] $p < 0.05$ when comparing controls and patients of both datasets combined, [e] Control mean and SD for negative symptoms are taken from Oruç et al. (54) and for functional outcome from Agid et al. (55), [f]for positive symptoms FEP (n = 56), MEP (n= 106); for duration of illness FEP (n = 41), MEP (n= 106); [g] CPZ adherence = chlorpromazine equivalent dose weighted by adherence, FEP (n = 52), MEP (n= 64); [*]Cohen's d => |0.2|, [**] Cohen's d => |0.5|, [***]Cohen's d => |0.8|

**Figure 3.**

*Comparison between Subtypes*




*Note.* This figure shows the mean comparisons between subtypes on the six markers of our model. ** $p < .001$, * $p < .05$

**Table 2.**

*Demographic Characteristics of Subtypes*

| | **Subtype 0. (N = 56)** | **Subtype 1. (N = 86)** | **Subtype 2. (N = 21)** |
|---|---|---|---|
| **Sex** (male/female) | (41 / 15) | (61 / 25) | (14 / 7) |
| | **Mean (SD)** | **Mean (SD)** | **Mean (SD)** |
| **Age** | 30.76 (8.60) | 31.59 (8.77) | 34.35 (7.76) |
| **TBV**[a] | 1.43 (0.12) | 1.41 (0.12) | 1.43 (0.13) |
| **Left Hippocampal Centrality** | 0.81 (0.03) | 0.80 (0.03) | 0.73 (0.03)[ce] |
| **Right Hippocampal Centrality** | 0.81 (0.03) | 0.80 (0.03) | 0.75 (0.03)[ce] |
| **Episodic Memory** | 17.24 (2.10) | 12.36 (2.90)[c] | 15.76 (2.84)[ce] |
| **Social Cognition** | 1.13 (0.13) | 0.96 (0.22)[c] | 1.09 (0.15)[cd] |
| **Negative Symptoms** | 13.93 (9.06) | 22.90 (10.33)[c] | 15.48 (8.48)[cd] |
| **Functional Outcome** | 55.21 (16.47) | 45.81 (13.89)[b] | 50.76 (17.36) |
| **Positive Symptoms**[f] | 12.16 (13.43) | 19.19 (18.76)[b] | 14.38 (13.02) |
| **Illness Duration (years)**[g] | 7.64 (8.39) | 10.17 (8.97) | 12.73 (7.17) |
| **CPZ Adherence**[h] | 403.76 (368.30) | 537.22 (581.96) | 451.21 (478.73) |

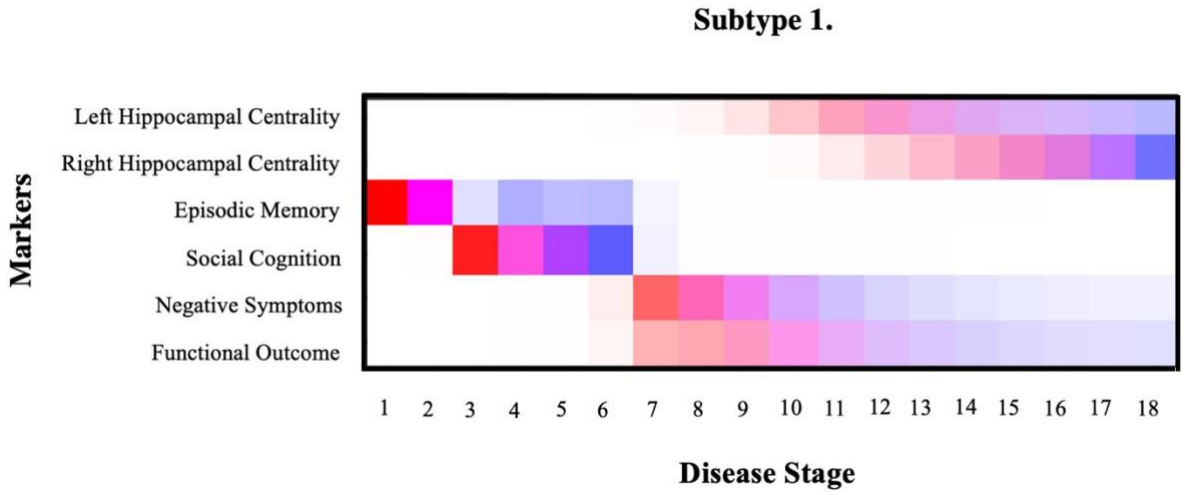
*Note.* [a]TBV = Total Brain Volume, reported in million decimals [b] $p < .05$, [c] $p < .001$ when compared to Subtype 0, [d] $p < .05$, [e] $p < .001$when comparing Subtype 1. with Subtype 2. [f] *n* positive symptoms Subtype 0 (FEP = 25, MEP = 30), Subtype 1 (FEP = 26, MEP = 60), Subtype 2 (FEP = 5, MEP = 16) [g] *n* duration of illness Subtype 0 (FEP = 20, MEP = 30), Subtype 1 (FEP = 18, MEP = 60), Subtype 2 (FEP = 3, MEP = 16)

[h] CPZ Adherence = chlorpromazine equivalent dose weighted by adherence, Subtype 0 ((FEP = 23, MEP = 19), Subtype 1 (FEP = 24, MEP = 37), Subtype 2 (FEP = 5, MEP = 8)

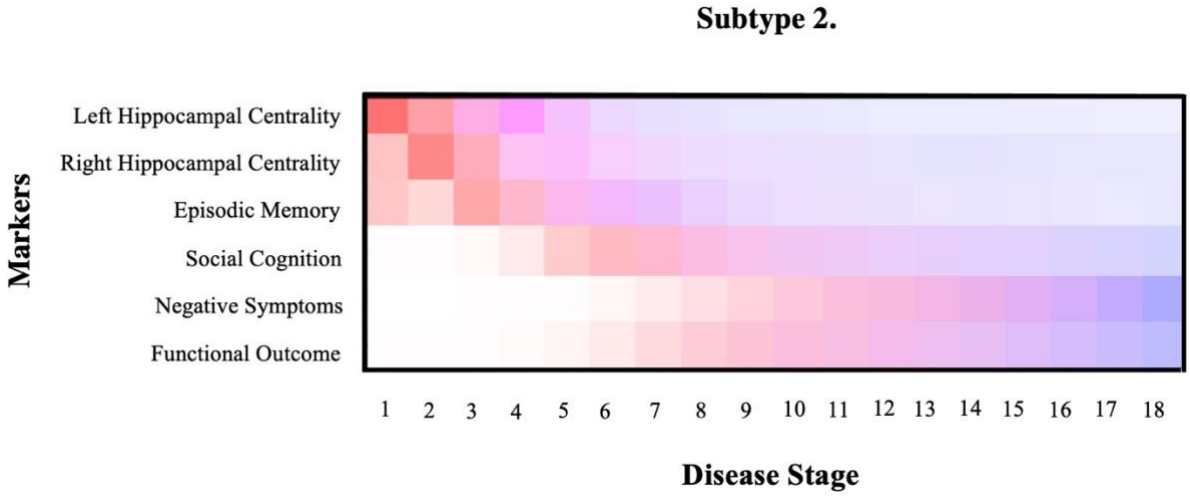
**Figure 4.**

*Disease Progression in both Disease Subtypes*

a)



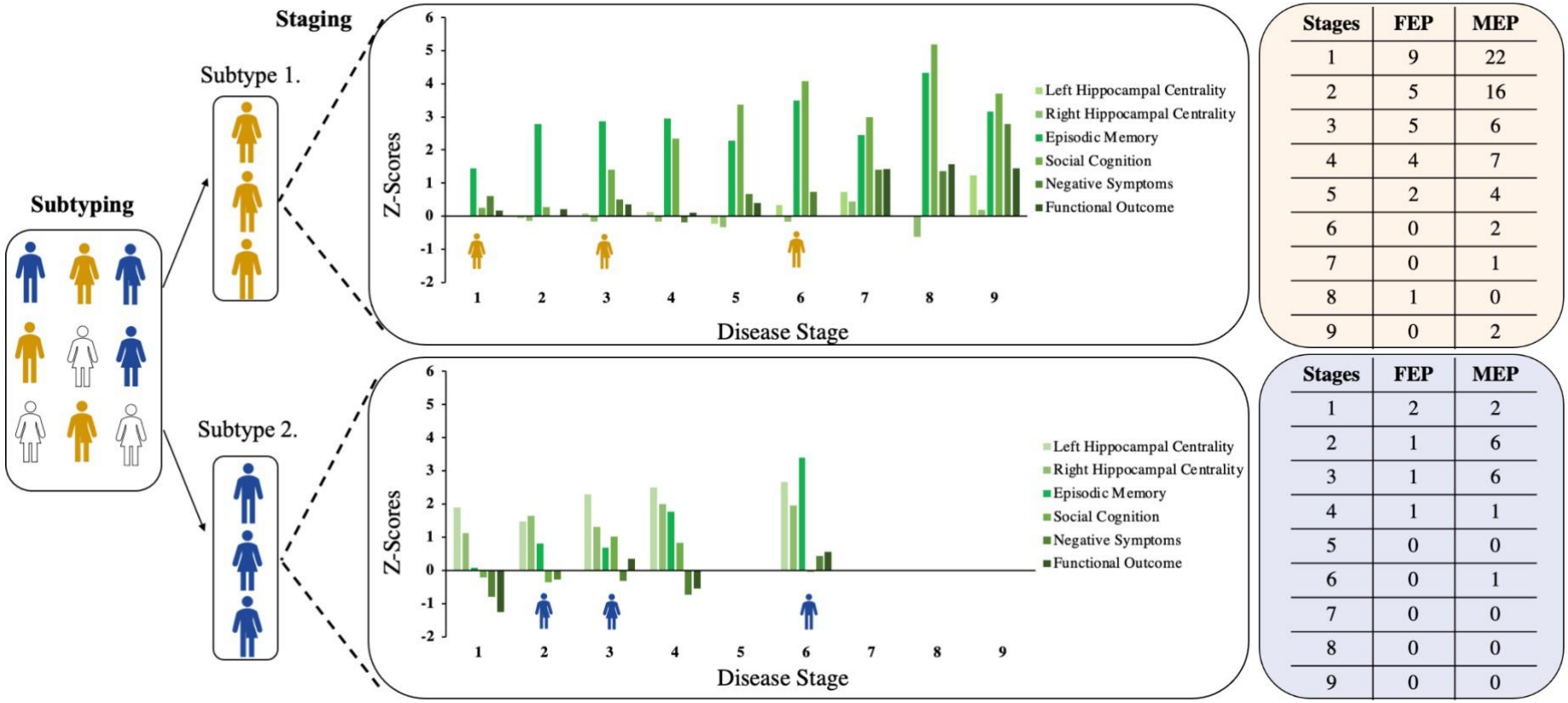

b)




*Note.* This figure shows the disease progression of the markers in both disease subtypes respectively. The color coding indicates a deviation from normality by one z-score in red, by two z- scores in pink and by three z-scores in blue. The stronger the color the higher the probability of individuals at that stage to deviate by one z-score, two z-scores etc.

## Figure 5.

*Subtype and Stage Inference*



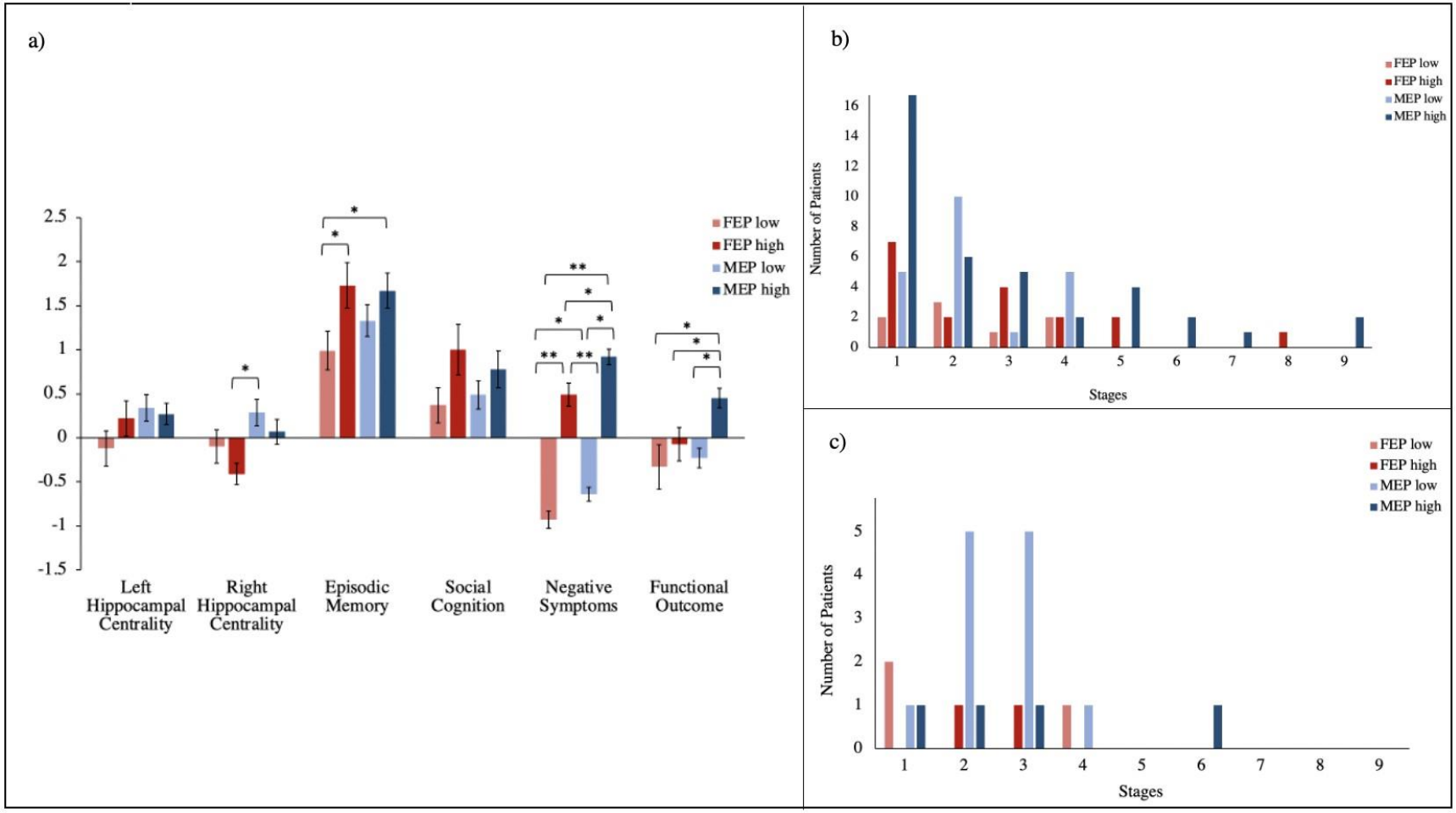

| Stages | FEP | MEP |
|---|---|---|
| 1 | 9 | 22 |
| 2 | 5 | 16 |
| 3 | 5 | 6 |
| 4 | 4 | 7 |
| 5 | 2 | 4 |
| 6 | 0 | 2 |
| 7 | 0 | 1 |
| 8 | 1 | 0 |
| 9 | 0 | 2 |

| Stages | FEP | MEP |
|---|---|---|
| 1 | 2 | 2 |
| 2 | 1 | 6 |
| 3 | 1 | 6 |
| 4 | 1 | 1 |
| 5 | 0 | 0 |
| 6 | 0 | 1 |
| 7 | 0 | 0 |
| 8 | 0 | 0 |
| 9 | 0 | 0 |

*Note.* This figure shows how patients are first subtyped into Subtype 0 (white), disease Subtype 1 (yellow), and disease Subtype 2 (blue) and how the disease progression of the respective disease subtypes is then inferred by placing the patients onto a disease stage of this subtype. The table on the right indicates the amount of FEP and MEP patients in each subtype.

**Figure 6.**

*FEP and MEP along low and high Negative Symptoms*




*Note.* This figure shows the comparison between FEP and MEP who score low and high on negative symptoms on the markers of our model (a) and their distribution across the stages of our disease subtype 1 (b) and 2 (c). ** $p < .001$, * $p < .05$

## References


1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5 (5th ed.). American Psychiatric Association. 2013.
2. Heinrichs RW, & Zakzanis, K. K. Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. Neuropsychology. 1998;12(3):426.
3. Sommer IE, Bearden CE, Van Dellen E, Breetvelt EJ, Duijff SN, Maijer K, et al. Early interventions in risk groups for schizophrenia: what are we waiting for? npj Schizophrenia. 2016;2(1):1-9.
4. Bora E, Lin A, Wood S, Yung A, McGorry P, Pantelis C. Cognitive deficits in youth with familial and clinical high risk to psychosis: a systematic review and meta-analysis. Acta Psychiatrica Scandinavica. 2014;130(1):1-15.
5. Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet. 2017;390(10100):1211-59.
6. Goldman HH, Skodol, A. E., & Lave, T. R. Revising axis V for DSM-IV: a review of measures of social functioning. Am J Psychiatry. 1992;149:9.
7. Lepage M, Bodnar, M., & Bowie, C. R. Neurocognition: clinical and functional outcomes in schizophrenia. The Canadian Journal of Psychiatry. 2014;59(1):5-12.
8. Chong HY, Teoh SL, Wu DB-C, Kotirum S, Chiou C-F, Chaiyakunapruk N. Global economic burden of schizophrenia: a systematic review. Neuropsychiatric disease and treatment. 2016;12:357.
9. Wu EQ, Birnbaum HG, Shi L, Ball DE, Kessler RC, Moulis M, et al. The economic burden of schizophrenia in the United States in 2002. Journal of Clinical Psychiatry. 2005;66(9):1122-9.
10. Marwaha S, Johnson S. Schizophrenia and employment. Social psychiatry and psychiatric epidemiology. 2004;39(5):337-49.
11. van den Heuvel MP, Scholtens LH, Kahn RS. Multiscale neuroscience of psychiatric disorders. Biological Psychiatry. 2019;86(7):512-22.
12. Lavigne KM, Kanagasabai K, Palaniyappan L. Ultra-high field neuroimaging in psychosis: A narrative review. Frontiers in Psychiatry. 2022:2688.
13. Mišić B, Goñi J, Betzel RF, Sporns O, McIntosh AR. A network convergence zone in the hippocampus. PLoS computational biology. 2014;10(12):e1003982.
14. Evans AC. Networks of anatomical covariance. Neuroimage. 2013;80:489-504.
15. Bullmore E, Sporns O. Complex brain networks: graph theoretical analysis of structural and functional systems. Nature reviews neuroscience. 2009;10(3):186-98.

16. Makowski C, Lewis JD, Khundrakpam B, Tardif CL, Palaniyappan L, Joober R, et al. Altered hippocampal centrality and dynamic anatomical covariance of intracortical microstructure in first episode psychosis. Hippocampus. 2020;30(10):1058-72.

17. Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology. 2011.

18. Valli I, Tognin, S., Fusar-Poli, P., & Mechelli, A. Episodic memory dysfunction in individuals at high-risk of psychosis: a systematic review of neuropsychological and neurofunctional studies. Current pharmaceutical design. 2012;18(4)(443-458).

19. Mesholam-Gately RI, Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. Neurocognition in first-episode schizophrenia: a meta-analytic review. Neuropsychology. 2009;23(3):315.

20. Raucher-Chéné D, Thibaudeau, E., Sauvé, G., Lavigne, K. M., & Lepage, M. Understanding others as a mediator between verbal memory and negative symptoms in schizophrenia-spectrum disorder. Journal of Psychiatric Research. 2021;143:429-35.

21. Buck G, Lavigne, K. M., Makowski, C., Joober, R., Malla, A., & Lepage, M. Sex differences in verbal memory predict functioning through negative symptoms in early psychosis. Schizophrenia bulletin. 2020 46(6):1587-95.

22. Lin CH, Huang, C. L., Chang, Y. C., Chen, P. W., Lin, C. Y., Tsai, G. E., & Lane, H. Y. Clinical symptoms, mainly negative symptoms, mediate the influence of neurocognition and social cognition on functional outcome of schizophrenia. Schizophrenia research. 2013;146(1-3):231-7.

23. Kharawala S, Hastedt C, Podhorna J, Shukla H, Kappelhoff B, Harvey PD. The relationship between cognition and functioning in schizophrenia: A semi-systematic review. Schizophrenia Research: Cognition. 2022;27:100217.

24. Milev P, Ho, B. C., Arndt, S., & Andreasen, N. C. Predictive values of neurocognition and negative symptoms on functional outcome in schizophrenia: a longitudinal first-episode study with 7-year follow-up. American Journal of Psychiatry. 2005;162(3):495-506.

25. Devoe DJ, Braun A, Seredynski T, Addington J. Negative symptoms and functioning in youth at risk of psychosis: a systematic review and meta-analysis. Harvard review of psychiatry. 2020;28(6):341.

26. Wunderink L, van Bebber, J., Sytema, S., Boonstra, N., Meijer, R. R., & Wigman, J. T. Reprint of: Negative symptoms predict high relapse rates and both predict less favorable functional outcome in first episode psychosis, independent of treatment strategy. Schizophrenia research. 2020;225:69-76.

27. Green MJ, Girshkin L, Kremerskothen K, Watkeys O, Quidé Y. A systematic review of studies reporting data-driven cognitive subtypes across the psychosis spectrum. Neuropsychology Review. 2020;30(4):446-60.

28. Rocca P, Galderisi S, Rossi A, Bertolino A, Rucci P, Gibertoni D, et al. Social cognition in people with schizophrenia: a cluster-analytic approach. Psychological Medicine. 2016;46(13):2717-29.

29. Strauss GP, Horan WP, Kirkpatrick B, Fischer BA, Keller WR, Miski P, et al. Deconstructing negative symptoms of schizophrenia: avolition–apathy and diminished expression clusters predict clinical presentation and functional outcome. Journal of psychiatric research. 2013;47(6):783-90.

30. Sauvé G, Malla A, Joober R, Brodeur MB, Lepage M. Comparing cognitive clusters across first-and multiple-episode of psychosis. Psychiatry Research. 2018;269:707-18.

31. Young AL, Marinescu, R. V., Oxtoby, N. P., Bocchetta, M., Yong, K., Firth, N. C., ... & Alexander, D. C. Uncovering the heterogeneity and temporal complexity of neurodegenerative diseases with Subtype and Stage Inference. Nature communications. 2018;9(1):1-16.

32. Vogel JW, Young AL, Oxtoby NP, Smith R, Ossenkoppele R, Strandberg OT, et al. Four distinct trajectories of tau deposition identified in Alzheimer's disease. Nature medicine. 2021;27(5):871-81.

33. Eshaghi A, Young AL, Wijeratne PA, Prados F, Arnold DL, Narayanan S, et al. Identifying multiple sclerosis subtypes using unsupervised machine learning and MRI data. Nature communications. 2021;12(1):1-12.

34. Young AL, Bocchetta M, Russell LL, Convery RS, Peakman G, Todd E, et al. Characterizing the clinical features and atrophy patterns of MAPT-related frontotemporal dementia with disease progression modeling. Neurology. 2021;97(9):e941-e52.

35. Oxtoby NP, Leyland L-A, Aksman LM, Thomas GE, Bunting EL, Wijeratne PA, et al. Sequence of clinical and neurodegeneration events in Parkinson's disease progression. Brain. 2021;144(3):975-88.

36. Jiang Y, Wang J, Zhou E, Palaniyappan L, Luo C, Ji G, et al. Neuroimaging biomarkers define neurophysiological subtypes with distinct trajectories in schizophrenia. Nature Mental Health. 2023;1(3):186-99.

37. Fiford CM, Sudre CH, Young AL, Macdougall A, Nicholas J, Manning EN, et al. Presumed small vessel disease, imaging and cognition markers in the Alzheimer's Disease Neuroimaging Initiative. Brain communications. 2021;3(4):fcab226.

38. McGorry PD, Purcell R, Hickie IB, Yung AR, Pantelis C, Jackson HJ. Clinical staging: a heuristic model for psychiatry and youth mental health. Medical Journal of Australia. 2007;187(S7):S40-S2.

39. Shah JL, Scott J, McGorry PD, Cross SP, Keshavan MS, Nelson B, et al. Transdiagnostic clinical staging in youth mental health: a first international consensus statement. World Psychiatry. 2020;19(2):233-42.

40. Béland S, Lepage M. The relative contributions of social cognition and self-reflectiveness to clinical insight in enduring schizophrenia. Psychiatry Research. 2017;258:116-23.

41. Iyer S, Jordan G, MacDonald K, Joober R, Malla A. Early intervention for psychosis: a Canadian perspective. The Journal of nervous and mental disease. 2015;203(5):356-64.

42. Pietrzak RH, Olver J, Norman T, Piskulic D, Maruff P, Snyder PJ. A comparison of the CogState Schizophrenia Battery and the Measurement and Treatment Research to Improve Cognition in Schizophrenia (MATRICS) Battery in assessing cognitive impairment in chronic schizophrenia. Journal of clinical and experimental neuropsychology. 2009;31(7):848-59.

43. Andreasen NC. The Scale for the Assessment of Negative Symptoms (SANS): conceptual and theoretical foundations. The British journal of psychiatry. 1983;155(S7):49-52.

44. Andreasen NC. The Scale for the Assessment of Positive Symptoms (SAPS). The British journal of psychiatry. 1984;155(S7):49-52.

45. Ad-Dab'bagh Y, Lyttelton O, Muehlboeck J, Lepage C, Einarson D, Mok K, et al., editors. The CIVET image-processing environment: a fully automated comprehensive pipeline for anatomical neuroimaging research. Proceedings of the 12th annual meeting of the organization for human brain mapping; 2006: Florence, Italy.

46. Sherif T, Rioux P, Rousseau M-E, Kassis N, Beck N, Adalat R, et al. CBRAIN: a webbased, distributed computing platform for collaborative neuroimaging research. Frontiers in neuroinformatics. 2014;8:54.

47. Klein A, Tourville J. 101 labeled brain images and a consistent human cortical labeling protocol. Frontiers in neuroscience. 2012;6:171.

48. Chakravarty MM, Steadman P, Van Eede MC, Calcott RD, Gu V, Shaw P, et al. Performing label-fusion-based segmentation using multiple automatically generated templates. Human brain mapping. 2013;34(10):2635-54.

49. Pipitone J, Park MTM, Winterburn J, Lett TA, Lerch JP, Pruessner JC, et al. Multi-atlas segmentation of the whole hippocampus and subfields using multiple automatically generated templates. Neuroimage. 2014;101:494-512.

50. Fortin J-P, Cullen N, Sheline YI, Taylor WD, Aselcioglu I, Cook PA, et al. Harmonization of cortical thickness measurements across scanners and sites. Neuroimage. 2018;167:10420.

51. Ajnakina O, Das T, Lally J, Di Forti M, Pariante CM, Marques TR, et al. Structural covariance of cortical gyrification at illness onset in treatment resistance: a longitudinal study of first-episode psychoses. Schizophrenia bulletin. 2021;47(6):1729-39.

52. Das T, Borgwardt S, Hauke DJ, Harrisberger F, Lang UE, Riecher-Rössler A, et al. Disorganized gyrification network properties during the transition to psychosis. JAMA psychiatry. 2018;75(6):613-22.

53. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. Neuroimage. 2010;52(3):1059-69.

54. Oruç S, Gülseren G, Kusbec O, Özbulut O. An evaluation of neuropsychiatric symptoms in Parkinson's disease patients. Nigerian Journal of Clinical Practice. 2017;20(7):900-4.

55. Agid O, McDonald K, Siu C, Tsoutsoulas C, Wass C, Zipursky RB, et al. Happiness in first-episode schizophrenia. Schizophrenia research. 2012;141(1):98-103.

56. Aksman LM, Wijeratne PA, Oxtoby NP, Eshaghi A, Shand C, Altmann A, et al. pySuStaIn:
a Python implementation of the Subtype and Stage Inference algorithm. SoftwareX. 2021;16:100811.

57. UCL-Pond. PySuStaIn/sustainworkshop.ipynb GitHub [updated Oct 22, 2021. Available from:
https://github.com/uclpond/pySuStaIn/blob/master/notebooks/SuStaInWorkshop.ipynb
.

58. UCL-Pond. SuStaIn tutorial using simulated data GitHub [updated Oct 22, 2021. Available from: https://github.com/ucl-pond/pySuStaIn/blob/master/notebooks/SuStaIn%20tutorial%20using%20simulated%20data.ipynb.

59. Jáni M, Kašpárek T. Emotion recognition and theory of mind in schizophrenia: a metaanalysis of neuroimaging studies. The World Journal of Biological Psychiatry. 2018;19(sup3):S86-S96.

60. Guimond S, Mothi SS, Makowski C, Chakravarty MM, Keshavan MS. Altered amygdala shape trajectories and emotion recognition in youth at familial high risk of schizophrenia who develop psychosis. Translational Psychiatry. 2022;12(1):1-8.

61. Killgore WD, Rosso IM, Gruber SA, Yurgelun-Todd DA. Amygdala volume and verbal memory performance in schizophrenia and bipolar disorder. Cognitive and Behavioral Neurology. 2009;22(1):28-37.

62. Makowski C, Bodnar M, Shenker J, Malla A, Joober R, Chakravarty M, et al. Linking persistent negative symptoms to amygdala–hippocampus structure in first-episode psychosis. Translational Psychiatry. 2017;7(8):e1195-e.

63. Addington J, Addington D. Patterns of premorbid functioning in first episode psychosis: relationship to 2-year outcome. Acta Psychiatrica Scandinavica. 2005;112(1):40-6.

64. Bucci P, Galderisi S, Mucci A, Rossi A, Rocca P, Bertolino A, et al. Premorbid academic and social functioning in patients with schizophrenia and its associations with negative symptoms and cognition. Acta Psychiatrica Scandinavica. 2018;138(3):253-66.

65. Leeson VC, Barnes TR, Hutton SB, Ron MA, Joyce EM. IQ as a predictor of functional outcome in schizophrenia: a longitudinal, four-year study of first-episode psychosis. Schizophrenia research. 2009;107(1):55-60.

66. Ventura J, Subotnik KL, Gitlin MJ, Gretchen-Doorly D, Ered A, Villa KF, et al. Negative symptoms and functioning during the first year after a recent onset of schizophrenia and 8 years later. Schizophrenia Research. 2015;161(2-3):407-13.

67. Firth NC, Primativo S, Brotherhood E, Young AL, Yong KX, Crutch SJ, et al. Sequences of cognitive decline in typical Alzheimer's disease and posterior cortical atrophy estimated using a novel event-based model of disease progression. Alzheimer's & Dementia. 2020;16(7):965-73.

68. Jayaweera H, Hickie I, Duffy S, Mowszowski L, Norrie L, Lagopoulos J, et al. Episodic memory in depression: the unique contribution of the anterior caudate and hippocampus. Psychological Medicine. 2016;46(10):2189-99.

69. Morosini PL, Magliano L, Brambilla La, Ugolini S, Pioli R. Development, reliability and acceptability of a new version of the DSM-IV Social and Occupational Functioning Assessment Scale (SOFAS) to assess routine social funtioning. Acta Psychiatrica Scandinavica. 2000;101(4):323-9.