Received: 15 March 2024 | Revised: 19 June 2024 | Accepted: 22 July 2024
DOI: 10.1111/acps.13743

ORIGINAL ARTICLE

Acta Psychiatrica Scandinavica WILEY

Using computer vision of facial expressions to assess symptom domains and treatment response in antipsychotic-naïve patients with first-episode psychosis

Karen S. Ambrosen[1] | Cecilie K. Lemvigh[1] | Mette Ø. Nielsen[1,2] | Birte Y. Glenthøj[1,2] | Warda T. Syeda[3,4] | Bjørn H. Ebdrup[1,2]

[1]Center for Neuropsychiatric Schizophrenia Research (CNSR) & Center for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS), Mental Health Center, Glostrup, Copenhagen University Hospital, Mental Health Services CPH, Copenhagen, Denmark

[2]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

[3]Melbourne Neuropsychiatry Center, Department of Psychiatry, The University of Melbourne, Melbourne, Victoria, Australia

[4]Melbourne Brain Centre Imaging Unit, Department of Radiology, The University of Melbourne, Melbourne, Victoria, Australia

Correspondence
Karen S. Ambrosen, Center for Neuropsychiatric Schizophrenia Research (CNSR) & Center for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS), Mental Health Center, Glostrup, Copenhagen University Hospital, Mental Health Services CPH, Nordstjernevej 41, 2600 Glostrup, Denmark.
Email: karen.marie.sandoe.ambrosen@regionh.dk

Funding information
Danish Agency for Science Technology and Innovation; Gerhard Linds legat; Lundbeck Foundation; Region Hovedstadens Psykiatri; Marie and Børge Kroghs Foundation

Abstract

Background: Facial expressions are a core aspect of non-verbal communication. Reduced emotional expressiveness of the face is a common negative symptom of schizophrenia, however, quantifying negative symptoms can be clinically challenging and involves a considerable element of rater subjectivity. We used computer vision to investigate if (i) automated assessment of facial expressions captures negative as well as positive and general symptom domains, and (ii) if automated assessments are associated with treatment response in initially antipsychotic-naïve patients with first-episode psychosis.

Method: We included 46 patients (mean age 25.4 (6.1); 65.2% males). Psychopathology was assessed at baseline and after 6 weeks of monotherapy with amisulpride using the Positive and Negative Syndrome Scale (PANSS). Baseline interview videos were recorded. Seventeen facial action units (AUs), that is, activation of muscles, from the Facial Action Coding System were extracted using OpenFace 2.0. A correlation matrix was calculated for each patient. Facial expressions were identified using spectral clustering at group-level. Associations between facial expressions and psychopathology were investigated using multiple linear regression.

Results: Three clusters of facial expressions were identified related to different locations of the face. Cluster 1 was associated with positive and general symptoms at baseline, Cluster 2 was associated with all symptom domains, showing the strongest association with the negative domain, and Cluster 3 was only associated with general symptoms. Cluster 1 was significantly associated with the clinically rated improvement in positive and general symptoms after treatment, and Cluster 2 was significantly associated with clinical improvement in all domains.

Conclusion: Using automated computer vision of facial expressions during PANSS interviews did not only capture negative symptoms but also combinations of the three overall domains of psychopathology. Moreover, automated assessments of facial expressions at baseline were associated with initial antipsychotic treatment response. The findings underscore the clinical relevance of facial expressions and motivate further investigations of computer vision in clinical psychiatry.

KEYWORDS

action units, drug-naïve, facial expressions, psychopathology, schizophrenia

1 | INTRODUCTION

Facial expressions are an integral part of non-verbal communication in humans. Blunted affect defined by diminished facial emotional expressivity is a key aspect of the negative symptom domain in schizophrenia.[1,2] Abnormal facial expressions have also been observed before the time of the diagnosis in clinical high-risk individuals, and these deficits have been shown to relate to a higher rate of conversion to psychosis.[3] Reliable assessment and quantification of negative symptoms is, however, notoriously challenging and involves a considerable element of rater subjectivity.[4,5] Evaluation of facial expressions is currently not used for systematic routine clinical assessment, although facial expressions may often be incorporated into the general clinical description of the patient's appearance.[1] Thus, an objective and reliable assessment of this symptom domain could potentially improve clinical practice. Automated computer vision techniques may be applied in support of clinical ratings, potentially enhancing the objectivity of the assessments. The Facial Action Coding System (FACS) is a framework for coding facial emotional expressions by identification of specific muscle movements called Action Units (AUs).[6] Recent developments have led to computer vision software for automated coding following this framework. Despite the rapid development of computer vision techniques, the application in studies of psychiatric populations remains sparse, and studies to date substantially vary both in terms of methodology and outcome measures.[7] Only a handful of studies have explored the associations between facial expressions and symptom severity in schizophrenia and have generally suggested an association between facial expressions and negative symptoms.[8–11] However, the included patients in these studies have primarily been selected based on a prespecified level of negative symptoms with less emphasis on positive and general symptom load. In addition, the previous studies were all conducted in medicated and chronically ill patients, rendering the question of whether these associations may be influenced by antipsychotic treatment, which is prone to induce extrapyramidal symptoms including blunted facial expressions. Further, other illness-related factors such as drug use or social isolation may have confounded previous studies. Finally, some videos used in the previous studies were obtained during unique interview settings instructed by the researchers limiting the comparability between studies.

Significant outcomes

- Data-driven facial expressions in antipsychotic-naïve patients with first-episode psychosis are related to clinical outcomes.
- Facial expressions are not only associated with negative symptoms but also with positive and general symptom domains.

Limitations

- The study is a Proof-of-Concept, and the findings should be validated in an independent sample.
- The inclusion of healthy participants would allow assessment of the normal variation of facial expressions.
- The Positive and Negative Syndrome Scale (PANSS) may be sub-optimal for clinical assessment of the negative symptom domain.

The present study aimed to investigate data-driven facial expressions obtained from video recordings of standard, semi-structured psychopathological interviews, and their relations to clinical outcomes in initially antipsychotic-naïve patients with first-episode psychosis. Based on the previous literature, we expected facial expressions to predominantly capture the negative symptom domain, but we also explored associations with positive, general, and total symptom domains. Secondly, because of the relationship between negative symptoms and prognosis,[12,13] we expected facial expressions to be associated with short-term treatment outcomes.

2 | MATERIALS AND METHODS

2.1 | Participants

We used available video material from antipsychotic-naïve patients with first-episode psychosis from the Pan European Collaboration on Antipsychotic-Naïve Schizophrenia cohort (PECANS, ClinicalTrials.gov Identifier: NCT01154829).[14] In the PECANS cohort, patients with a schizophrenia spectrum diagnosis aged 18–45 years were recruited from psychiatric hospitals and outpatient mental health centers in the Capital Region of Denmark. ICD-10 diagnoses of schizophrenia or schizoaffective psychoses were based on structured diagnostic interviews (Schedule of Clinical Assessment in Neuropsychiatry, SCAN, version 2.1). Full scale intelligence (FSIQ) was estimated using four tasks from The Wechsler Adult Intelligence Scale—Third edition (block design, matrix reasoning, similarities, and vocabulary).[15] Patients with a current diagnosis of drug dependency according to ICD-10 were excluded, but a previous diagnosis of drug dependency or current occasional use of drugs was accepted. Current drug use was screened by urine test (Rapid Response, Jepsen HealthCare, Tune, Denmark). Any previous exposure to antipsychotic medication was an exclusion criterion, as well as treatment with antidepressant medication within the last month. If necessary, benzodiazepines or sleeping medication were allowed during the study period. Further exclusion criteria for the patients were organic brain damage, previous impact-related unconsciousness, contraindications for treatment with amisulpride, and intellectual disability (IQ < 70). Following baseline assessments, patients were treated with flexible doses of amisulpride for 6 weeks and subsequently re-assessed. The project was approved by the Danish National Committee on Biomedical Research Ethics (H-D-2008-088) and conducted following the declaration of Helsinki II. All patients provided written informed consent.

2.2 | Psychopathology and treatment response

Psychopathology was assessed at baseline and after 6 weeks of treatment using the Positive and Negative Syndrome Scale (PANSS).[16] PANSS is a semi-structured clinical interview expected to last 35–40 min comprising 30 items. The items were originally divided into domains constituting seven positive symptoms (P1-P7), seven negative symptoms (N1-N7), and 16 general symptoms (G1-G16). Following the interviews, consensus ratings were performed by two trained clinicians. Each item is rated on an ordinal scale from 1 to 7 with higher scores representing more severe symptoms. Treatment response was calculated by the relative change in PANSS scores from baseline to 6-week follow-up as

$$\Delta\text{PANSS} = \frac{\text{PANSS}_{\text{Follow-up}} - \text{PANSS}_{\text{Baseline}}}{\text{PANSS}_{\text{Baseline}}}.$$

2.3 | Video recordings

For the entire duration of the PANSS interview, patients were video recorded with a sampling rate of 25 images/second. The videos were recorded under real-world clinical conditions, that is, varying backgrounds, recording angles, and light. The patient's face and upper part of the torso were visible in all videos. In this study, computer vision analyses were applied to the baseline PANSS video recordings.

2.4 | Estimation of facial action units

The Facial Action Coding System (FACS) was applied using OpenFace 2.0, an open-source machine-learning tool for facial behavior analysis.[17] The toolkit is capable of facial landmark detection, head pose estimation, facial action unit (AU) recognition and intensity estimation, and eye-gaze estimation. OpenFace 2.0 was applied to the video recordings to estimate the activation of AUs and their intensities.[18] The activation intensity of 17 AUs from FACS and confidence estimates were computed frame-wise for the entire duration of each recording. Subsequently, frames with a confidence level below 80% were removed. Participants with less than 90% frames remaining were excluded from the analyses.

2.5 | Estimation of facial expressions

An overview of the analysis pipeline is provided in Figure 1. For each participant, a facial correlation matrix

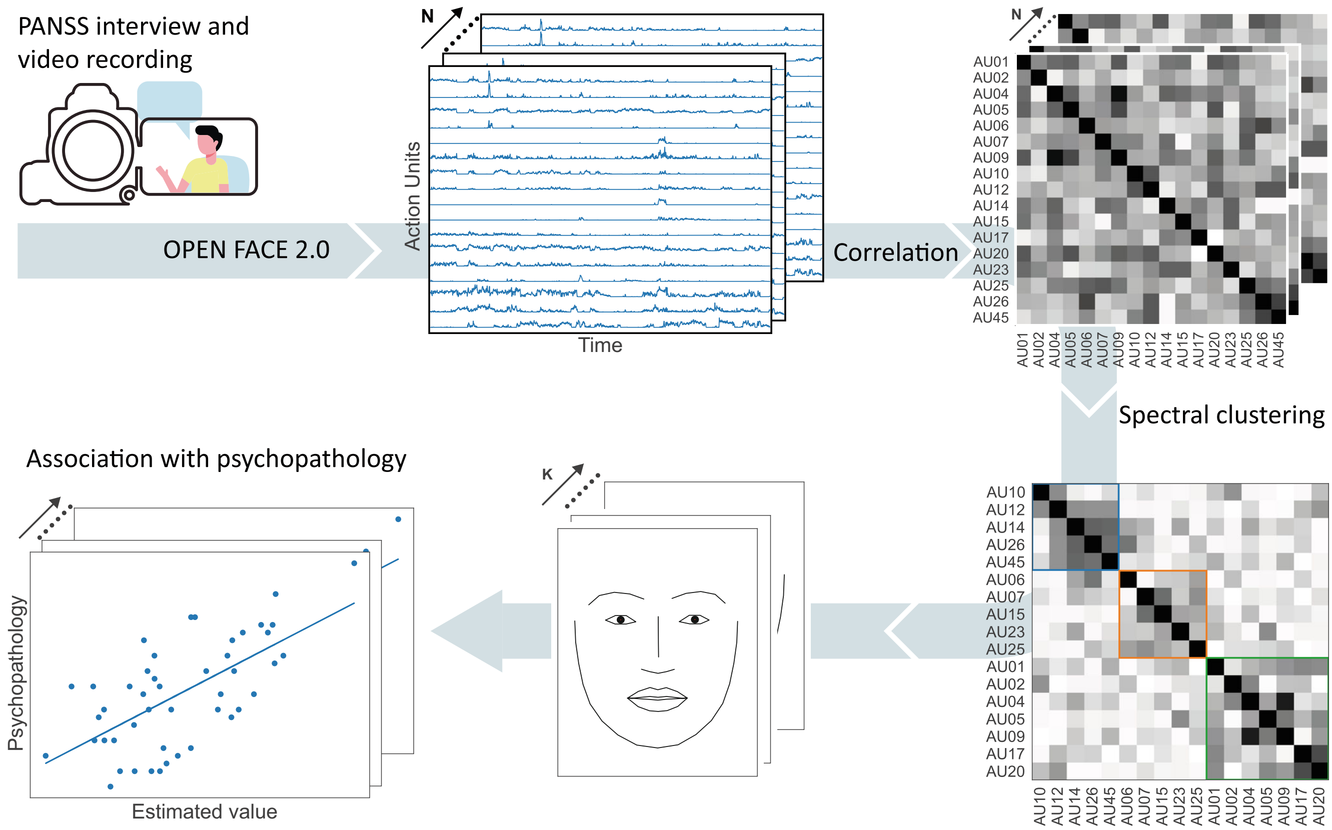

FIGURE 1 Overview of the analysis pipeline. PANSS interviews were video-recorded and Action Unit (AU) timeseries were extracted using OpenFace 2.0. The pair-wise co-activations of the AUs were calculated using Spearman's rank correlation and spectral clustering was performed on the group-level correlation matrix. The associations between the identified clusters, that is, facial expressions, and psychopathology were assessed using multiple linear regression.

was calculated by correlating the 17 AU timeseries using Spearman's rank correlation, and the Fisher r-to-z transformation was applied. The correlation matrices were averaged across all participants. To obtain distinct facial expressions, spectral clustering was applied. Spectral clustering is particularly useful for clustering high-dimensional datasets with many features as it clusters the data based on the eigenvectors of the graph Laplacian matrix. Furthermore, spectral clustering does not include assumptions about the cluster shapes and is therefore capable of clustering highly non-convex data. Spectral clustering divides the AUs, rather than the patients, into mutually exclusive clusters based on their similarity in activation pattern. Therefore, all clusters represent data from all patients. The clustering aims to partition the correlation matrix such that within-cluster correlations are maximized, and the between-cluster correlations are minimized.[19] The Scikit-learn[20] implementation of spectral clustering was applied, using the absolute values of the group-level correlation matrix as the affinity matrix. The *cluster_qr* method was applied to assign labels in the embedding space after the Laplacian embedding.[21] The optimal number of clusters was selected using the eigengap heuristic.[22] The eigengap is defined as the difference between successive eigenvalues when the eigenvalues are sorted in ascending order. The optimal clustering was applied to the correlation matrix of each patient.

2.6 | Associations with psychopathology and treatment response

The within-cluster correlations were used as features to investigate associations with psychopathology and treatment response. The associations were tested by multiple linear regression, separately for each cluster and PANSS subscale. Feature selection was performed using recursive feature elimination for each feature set of size one to P, where P was the maximum number of features in each cluster. The P resulting models were compared using the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and adjusted explained variance (R^2_{adj}). The model with the highest R^2_{adj} for each problem was fitted and the Spearman's rank correlation between the model-estimated scores and the clinically

TABLE 1 Characteristics of the patients.

	Baseline	Follow-up
n	46	34
Age, mean (SD)	25.4 (6.1)	25.2 (6.4)[a]
Sex, *n* (%)		
Female	16 (34.8)	12 (35.3)
Male	30 (65.2)	22 (64.7)
Parents socioeconomic status, *n* (%)[b]		
A	7 (16.3)	4 (11.8)
B	26 (60.5)	23 (67.6)
C	10 (23.3)	7 (20.6)
Full scale intelligence, mean (SD)[c]	97.7 (20.2)	
Years of education, mean (SD)[d]	12.1 (2.3)	
Amisulpride (mg), mean (SD)	0.0 (0.0)	241.2 (141.1)
PANSS total, mean (SD)	80.8 (14.1)	65.9 (14.6)
PANSS Pos, mean (SD)	19.5 (3.4)	14.2 (3.5)
PANSS Neg, mean (SD)	20.8 (7.5)	19.7 (6.0)
PANSS Gen, mean (SD)	40.5 (7.8)	32.0 (8.6)
Urine screening—THC, *n* (%)[e]		
Negative	41 (97.6)	
Positive	1 (2.4)	
Urine screening—BENZODIAZEPINES, *n* (%)		
Negative	41 (97.6)	
Positive	1 (2.4)	

[a]Lower mean age at follow-up due to fewer patients.
[b]Three patients have missing data on Parental Socioeconomic status.
[c]Six patients have missing data on full scale intelligence.
[d]Five patients have missing data on years of education.
[e]Four patients have missing data on urine screening.

rated baseline PANSS subscales and change in PANSS subscales, respectively, were calculated. The p-values of the correlations were corrected for multiple comparisons using Bonferroni correction.

3 | RESULTS

3.1 | Participants

Fifty patients had available video recordings from their baseline PANSS interview. Removing the frames with less than 80% confidence and subsequently excluding the patients with less than 90% of their video recording remaining resulted in a final sample of 46 patients. The duration of the resulting recordings was 22–97 min with an average duration of 48.7 minutes.

At baseline, the patients were moderately ill with an average PANSS total score of 80.8 (SD = 14.1).[23] The patients received monotherapy with flexible doses of amisulpride with an average dose of 241.2 mg (SD = 141.1 mg) after 6 weeks of treatment. The average PANSS total score was reduced after 6 weeks to 65.9 (SD = 14.6) with clinical improvement across all domains. One patient tested positive for THC and one patient tested positive for benzodiazepine use. Rerunning the analyses excluding these participants did not change the overall patterns reported below. One patient received sleeping medication PRN (Baldrian, non-prescription); however, intake was not allowed 24 hours before the assessment. Moreover, one patient had a previous diagnosis of drug dependency. Exclusion of this patient from the analyses did not change the overall pattern. More characteristics of the patients are provided in Table 1.

After 6 weeks of treatment, 12 patients were lost to follow-up. The reasons for dropping out were the need for change in medication type ($N = 3$), loss of contact ($N = 2$), without cause ($N = 3$), and other reasons ($N = 4$). The characteristics of the patients who stayed in the study and those who dropped out are provided in Supplementary Material Table S2. The basic demographics did not differ between the two groups but the patients who dropped out had more severe symptoms with an average total PANSS score of 84.3 (SD = 12.5), higher estimated full scale intelligence (mean = 102.8, SD = 23.2), and more years of education (mean = 12.8, SD = 2.6) compared to the patients who stayed in the study with an average total PANSS score of 76.9 (SD = 13.9), estimated full scale intelligence of 96.4 (SD = 19.2), and average years of education of 11.8 (SD = 2.1).

3.2 | Facial expressions

Three clusters of facial expressions were found to describe the data best (Figure 2). Cluster 1 comprised AUs related to muscles in the upper face, Cluster 2 primarily comprised AUs related to muscles in the lower face, and Cluster 3 comprised a mix.

The eigenvalues, sorted in ascending order, used as the criterion for cluster selection are visualized in Supplementary Figure S1. The number of clusters with the largest eigengap was selected. The averaged correlation matrix across all patients and the averaged correlation matrix sorted according to the clustering are shown in Supplementary Figure S2.

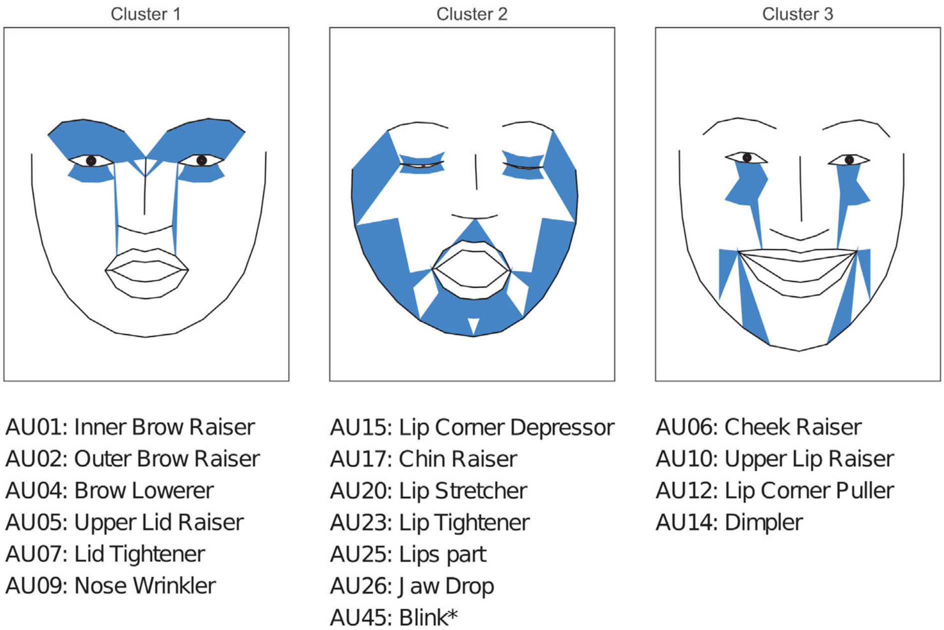

FIGURE 2 The three distinct facial expressions identified using spectral clustering. The visualization is made with PyFeat.[24] *PyFeat does not include AU45, but instead AU43 (Eyes closed) is visualized. Both AU43 and AU45 are characterized by the relaxation of levator palpebrae superioris, orbicularis oculi (pars palpebralis).

3.3 | Associations with psychopathology

The within-cluster correlation coefficients were used as features in a multiple linear regression model. The clusters contained six, seven, and four AUs, respectively, resulting in: 15 features in Cluster 1, 21 features in Cluster 2, and 6 features in Cluster 3. The information criteria scores (AIC and BIC) and the maximum R^2_{adj} used to find the optimal feature set are visualized in Supplementary Figure S3. The model for each combination was selected to minimize the information criteria scores and maximize the explained variance. For details on the specific features included in each model after feature selection, please see Supplementary Table S3. The presented results are uncorrected for basic demographics but including age and sex did not change the findings (data not shown). The correlations between the clinically rated and model-estimated PANSS scores are visualized in Figure 3.

3.4 | Baseline

The within-network temporal correlations of Cluster 1 showed significant associations with the positive ($\rho = 0.50$, $p_{Bonf} = 0.011$, $R^2_{adj} = 0.20$) and general ($\rho = 0.55$, $p_{Bonf} = 0.002$, $R^2_{adj} = 0.18$) PANSS domains. Cluster 2 showed significant associations with all domains, but the strongest association was with the negative domain (negative: $\rho = 0.66$, $p_{Bonf} < 0.0001$, $R^2_{adj} = 0.36$, positive: $\rho = 0.46$, $p_{Bonf} = 0.03$, $R^2_{adj} = 0.13$, general: $\rho = 0.60$, $p_{Bonf} = 0.0002$, $R^2_{adj} = 0.21$). Cluster 3 was only significant associated with the general domain ($\rho = 0.45$, $p_{Bonf} = 0.043$, $R^2_{adj} = 0.14$).

3.5 | Follow-up

Both Cluster 1 and Cluster 2 showed significant associations with treatment response after 6 weeks. For both clusters, the association was strongest for improvements in positive symptoms (Cluster 1: $\rho = 0.69$, $p_{Bonf} = 0.0001$, $R^2_{adj} = 0.24$, Cluster 2: $\rho = 0.86$, $p_{Bonf} < 0.0001$, $R^2_{adj} = 0.50$) and they both also showed associations with improvements in general symptoms (Cluster 1: $\rho = 0.61$, $p_{Bonf} = 0.003$, $R^2_{adj} = 0.21$, Cluster 2: $\rho = 0.62$, $p_{Bonf} = 0.002$, $R^2_{adj} = 0.21$). Moreover, Cluster 2 was significantly associated with reductions in negative symptoms ($\rho = 0.77$, $p_{Bonf} < 0.0001$, $R^2_{adj} = 0.62$). Cluster 3 did not show any significant correlations with treatment response. All correlations are provided in Supplementary Table S1 and scatterplots of the model-estimated and clinically rated PANSS scores are provided in Supplementary Figures S4-S6.

4 | DISCUSSION

Using computer vision, we investigated if data-driven facial expressions during PANSS interviews were related to baseline psychopathology and initial treatment response in antipsychotic-naïve patients with first-episode psychosis. Analyses of the AU co-activations showed that three clusters best captured the data, each representing a facial expression.

Cluster 1 comprised AUs in the upper face, related to brows, lids, and nose. Studies have shown that upper facial features are most important for identifying fear, anger, and sadness.[25–27] Moreover, Wegrzyn et al. showed

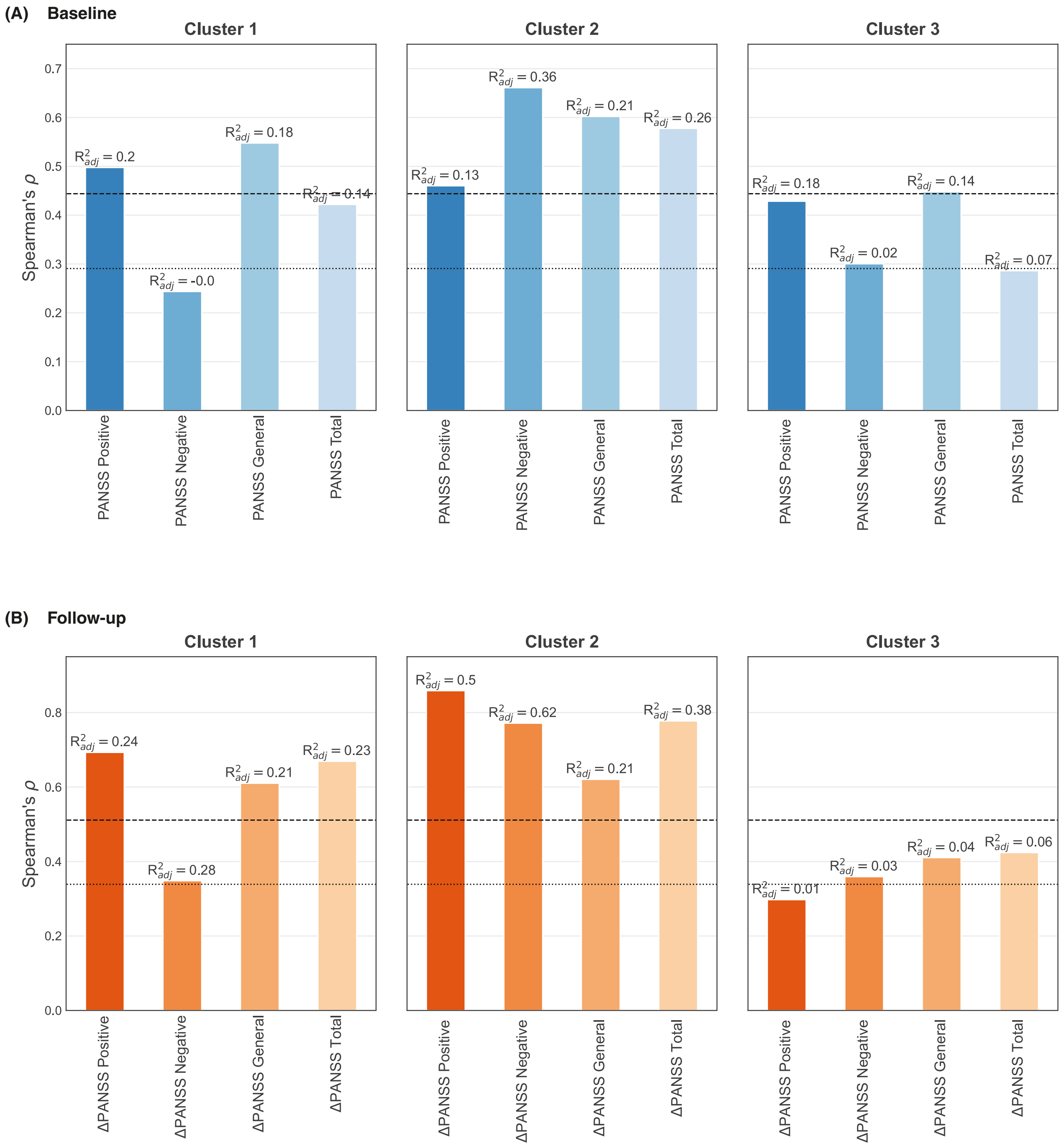

FIGURE 3 Spearman's rank correlations between the clinically rated and the model-estimated PANSS subscales (Panel A) and changes in PANSS subscales (Panel B) for the three clusters. The dotted and dashed lines mark the significance levels of the correlations, $\alpha = 0.05$ and $\alpha = 0.002$, respectively. The adjusted explained variance (R^2_{adj}) for each model is stated above the bars.

that upper lid raiser (AU5) was most important for identifying fear, lid tightener (AU7) for identifying anger, and inner brow raiser (AU1) and brow lowerer (AU4) for identifying sadness.[27] These four AUs were all assigned to Cluster 1. Cluster 2 primarily comprised AUs in the lower face, related to lips, jaw, and chin. Lower facial features are used to recognize emotions of happiness and disgust.[26,27] Cluster 2 comprised seven AUs, which are involved in all basic emotions, except for happiness. Clinically the overrepresentation of muscle activation in the lower face could be interpreted as the absence of movements in other parts of the face, suggesting that apart from activating the perioral muscles which are needed to speak, the patient is otherwise displaying flat affect. It should be noted that all the patients are included in all three clusters. Cluster 3 comprised AUs associated with expressions of happiness.[28] Happiness is the easiest emotion to identify,[27] which may explain why these AUs were assigned to a separate cluster. The identification of a cluster associated with happiness is in line with previous findings, although the methodology differs between the studies.[10,29]

We observed several associations between facial expressions and psychopathology. Cluster 1 was significantly associated with positive and general symptoms. Cluster 2 was significantly associated with psychopathology across all domains, but the strongest association was observed for negative symptoms. Cluster 3 was associated with general symptoms only. Surprisingly, the facial expressions were associated with all symptom domains and not only the negative domain as hypothesized. This may be explained by the patient group included, which was not selected based on their negative symptom severity and hence represents a broader spectrum of patients with schizophrenia compared to previous studies with one study specifically recruiting patients high on negative symptoms[8] and the studies by Tron et al. mainly included patients suffering from post-psychotic residual negative signs schizophrenia.[9,10] Our patients were moderately ill according to the average PANSS score with similar levels of positive and negative symptoms.[23]

Based on the association between negative symptoms and prognosis,[12,13] we also expected facial expressions at baseline to be associated with treatment response. In line with this hypothesis, we showed that Cluster 1 was associated with the observed improvement in positive and general symptoms after 6 weeks of treatment while Cluster 2 was associated with improvement in all domains, that is, the same pattern as observed at baseline. It should be noted that 12 patients dropped out before the follow-up visit and are therefore not included in the treatment response analyses. There was no clear pattern in the patients who dropped out; they had on average more severe symptoms at baseline, but unexpectedly higher estimated full-scale intelligence, and more years of education than those who stayed in the study. A common reason for dropping out was a need for change in medication, indicating the lack of treatment response, which could impact the results. Using computer vision to assess facial expressions at the time of diagnosis may therefore be used as an objective prognostic marker that could be easily incorporated into clinical practice.

Our study differs from the previous literature in the type of psychosis patients included. All previous studies included chronic and medicated patients with schizophrenia. In contrast, the patients included in this study were strictly antipsychotic-naïve patients experiencing their first psychotic episode and thus our findings cannot be attributed to side effects from medication or other illness-related factors. Treatment with antipsychotic medication may dampen the expressiveness of the face that otherwise may be associated with positive symptoms or bias the clinical ratings of negative symptoms.

For the video recordings, Tron et al.[9,10] used a 15-minute-long interview with one general and three evocative questions about current mood and recent emotional events. Similarly, Vijay et al.[11] used a 10–15-minute-long semi-structured naturalistic clinical interview. In contrast, like Bishay et al.,[8] we used a semi-structured psychopathological interview, which has a longer duration and is designed to evoke salient responses since the interview specifically revolves around present symptom severity.

Previous studies have used different temporal features of the AUs, such as activation ratio, activation level, activation length, change ratio, and fast change ratio.[8,9] Common for these features, the signal is compressed into a single measure for each AU, eliminating the temporal dynamics. Another applied approach is to cluster the AUs using k-means[10] or to extract spatiotemporal states using wavelet transforms and Hidden Markov Models.[29] The latter approach is interesting but with limited clinical applicability due to the requirement for recordings to be the same length. Keeping a set interview duration is feasible if a fixed naturalistic stimulus is used, but not for recordings that are only semi-structured with varying durations as in our study. To incorporate both temporal and spatial information, we used pairwise temporal correlations between AU and subsequently applied spectral clustering.

A possible limitation of the study is the use of PANSS to evaluate levels of psychopathology because it requires well-trained raters and poses a risk of rater subjectivity. All interviews were conducted by two trained medical professionals and consensus ratings based on the videos were conducted. Furthermore, PANSS is no longer the clinical gold standard for the assessment of negative symptoms. The negative domain of the PANSS scale only partly covers the current consensus definition of negative symptoms,[30,31] which are better assessed by more recent scales like the Brief Negative Symptom Scale (BNSS)[32] or the Clinical Assessment Interview for Negative Symptoms (CAINS).[33] These scales were not developed at the time of data collection in the PECANS cohort but may be considered in future studies. Although the PANSS interview is time-consuming and shows problems with validity, especially regarding the negative symptom domain, PANSS is still a widely used instrument to assess psychopathology in patients with psychosis. Another limitation is that the videos were not recorded to analyze facial expressions. Hence, detailed information for each session, such as the distance between the camera and the patient and the patients' positions, is unavailable. Moreover, the study is limited by a relatively small sample size and the drop-out rate. However, as the clusters are based on baseline data only, these are not affected by the drop-out rate. Also, the analyzed video recordings are quite long, improving the signal within each patient compensating for the relatively small sample size. Finally, our study is limited by the absence of a healthy control group, which may have provided relevant information on the normal variation in facial expressions.

A major strength of the study is the inclusion of antipsychotic-naïve first-episode patients, which are particularly difficult to recruit. In addition, the videos were obtained under real-world conditions using standard equipment that could be easily transferred to clinical practice, thereby adding evidence for the associations between facial expressions and clinical outcomes in patients suffering from psychosis. Hence our findings underscore the clinical relevance of automated assessment of facial expressions and motivate further investigations of computer vision in clinical psychiatry. Future work should include validation in a larger sample and explore the applicability of shorter videos.

FUNDING INFORMATION

This project was funded by the Lundbeck Foundation (ID: R13-A1349, R25-A2701), Danish Agency for Science Technology and Innovation (271–08-0690), Gerhard Linds legat, Marie and Børge Kroghs Foundation, and the Mental Health Services in the Capital Region of Denmark.

CONFLICT OF INTEREST STATEMENT

BYG has been the leader of a Lundbeck Foundation Centre of Excellence for Clinical Intervention and Neuropsychiatric Schizophrenia Research (CINS) (January 2009–December 2021), which was partially financed by an independent grant from the Lundbeck Foundation based on international review and partially financed by the Mental Health Services in the Capital Region of Denmark, the University of Copenhagen, and other foundations. All grants are the property of the Mental Health Services in the Capital Region of Denmark and administrated by them. BE is part of the Advisory Board of Eli Lilly Denmark A/S, Janssen-Cilag, Lundbeck Pharma A/S, and Takeda Pharmaceutical Company Ltd; and has received lecture fees from Bristol-Myers Squibb, Boehringer Ingelheim, Otsuka Pharma Scandinavia AB, Eli Lilly Company, and Lundbeck Pharma A/S. All other authors have nothing to disclose.

DATA AVAILABILITY STATEMENT

The data used in this study are not publicly available due to privacy or ethical restrictions.

ETHICS STATEMENT

This project was approved by the Danish National Committee on Biomedical Research Ethics (H-D-2008-088) and conducted following the declaration of Helsinki II.

PATIENT CONSENT STATEMENT

All patients provided written informed consent.

CLINICAL TRIAL REGISTRATION

The Pan European Collaboration on Antipsychotic-Naïve Schizophrenia cohort (PECANS, ClinicalTrials.gov Identifier: NCT01154829).

ORCID

Karen S. Ambrosen https://orcid.org/0000-0002-5638-2357
Mette Ø. Nielsen https://orcid.org/0000-0002-0780-7099

REFERENCES

1. Mandal MK, Habel U, Gur RC. Facial expression-based indicators of schizophrenia: evidence from recent research. *Schizophr Res*. 2023;252(January):335-344. doi:10.1016/j.schres.2023.01.016
2. Tandon R, Nasrallah HA, Keshavan MS. Schizophrenia, "just the facts" 4. Clinical features and conceptualization. *Schizophrenia Res*. 2009;110(1–3):1-23. doi:10.1016/j.schres.2009.03.005
3. Gupta T, Haase CM, Strauss GP, Cohen AS, Mittal VA. Alterations in facial expressivity in youth at clinical high-risk for psychosis. *J Abnorm Psychol*. 2019;128(4):341-351. doi:10.1037/abn0000413
4. Galderisi S, Kaiser S, Bitter I, et al. EPA guidance on treatment of negative symptoms in schizophrenia. *Eur Psychiatry*. 2021; 64(1):e21. doi:10.1192/j.eurpsy.2021.13
5. Norman RMG, Malla AK, Cortese L, Diaz F. A study of the interrelationship between and comparative interrater reliability of the SAPS, SANS and PANSS. *Schizophrenia Res*. 1996;19(1): 73-85. doi:10.1016/0920-9964(95)00055-0
6. Ekman P, Friesen WV. Facial action coding system: A technique for the measurement of facial movement. Consulting Psychologists Press; 1978.
7. Jiang Z, Luskus M, Seyedi S, et al. Utilizing computer vision for facial behavior analysis in schizophrenia studies: a systematic review. *PLoS ONE*. Vol 17. Public Library of Science; 2022: e0266828. doi:10.1371/journal.pone.0266828
8. Bishay M, Palasek P, Priebe S, Patras I. SchiNet: automatic estimation of symptoms of schizophrenia from facial behaviour analysis. *IEEE Trans Affect Comput*. 2021;12(4):949-961. doi:10.1109/TAFFC.2019.2907628
9. Tron T, Peled A, Grinsphoon A, Weinshall D. Automated facial expressions analysis in schizophrenia: a continuous dynamic

approach. *Commun Comput Inform Sci*. 2016a;604:72-81. doi:10.1007/978-3-319-32270-4_8
10. Tron T, Peled A, Grinsphoon A, Weinshall D. Facial expressions and flat affect in schizophrenia, automatic analysis from depth camera data. 3rd IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2016, 220–223. 2016b. doi:10.1109/BHI.2016.7455874
11. Vijay S, Baltrušaitis T, Pennant L, Baker JT, Morency L-P. Computational study of psychosis symptoms and facial expressions. *Computing and Mental Health Workshop at CHI (Vol 2.)*. 2016.
12. Correll CU, Schooler NR. Negative symptoms in schizophrenia: a review and clinical guide for recognition, assessment, and treatment. *Neuropsychiatr Dis Treat*. 2020;16:519-534. doi:10.2147/NDT.S225643
13. Sarkar S. Conceptualization and treatment of negative symptoms in schizophrenia. *World J Psych*. 2015;5(4):352-361. doi:10.5498/wjp.v5.i4.352
14. Nielsen MØ, Rostrup E, Wulff S, et al. Alterations of the brain reward system in antipsychotic naïve schizophrenia patients. *Biol Psychiatry*. 2012;71(10):898-905. doi:10.1016/j.biopsych.2012.02.007
15. Wechsler D. *Manual for the Wechsler Adult Intelligence Scale—Third Edition (WAIS-III)*. The Psychological Corporation; 1997. https://scholar.google.com/scholar_lookup?title=Manual+for+the+Wechsler+Adult+Intelligence+Scale+-+third+edition+(WAIS-III)&author=Wechsler+D.&publication+year=1997
16. Kay SR, Fiszbein A, Opler LA. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13(2):261-276. doi:10.1093/schbul/13.2.261
17. Baltrusaitis T, Zadeh A, Lim YC, Morency LP. OpenFace 2.0: facial behavior analysis toolkit. Proceedings—13th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2018, 59–66. 2018. doi:10.1109/FG.2018.00019
18. Baltrušaitis T, Mahmoud M, Robinson P. Cross-dataset learning and person-specific normalisation for automatic Action Unit detection. *2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition, FG 2015*. 2015. doi:10.1109/FG.2015.7284869
19. Von Luxburg U. A tutorial on spectral clustering. *Stat Comput*. 2007;17(4):395-416. doi:10.1007/s11222-007-9033-z
20. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. *J Mach Learn Res*. 2011;12(9):2825-2830.
21. Damle A, Minden V, Ying L. Simple, direct and efficient multiway spectral clustering. *Inform Inference*. 2019;8(1):181-203. doi:10.1093/imaiai/iay008
22. Zelnik-Manor L, Perona P. Self-tuning spectral clustering. *Adv Neural Inform Process Syst*. 2004;17:1601-1608.
23. Leucht S, Kane JM, Kissling W, Hamann J, Etschel E, Engel RR. What does the PANSS mean? *Schizophr Res*. 2005;79(2–3):231-238. doi:10.1016/j.schres.2005.04.008
24. Cheong JH, Jolly E, Xie T, Byrne S, Kenney M, Chang LJ. Py-feat: python facial expression analysis toolbox. *Affect Sci*. 2023;4(4):781-796. doi:10.1007/s42761-023-00191-4
25. Kim M, Cho Y, Kim SY. Effects of diagnostic regions on facial emotion recognition: the moving window technique. *Front Psychol*. 2022;13(Sept):1-12. doi:10.3389/fpsyg.2022.966623
26. Schurgin MW, Nelson J, Iida S, Ohira H, Chiao JY, Franconeri SL. Eye movements during emotion recognition in faces. *J vis*. 2014;14(13):1-16. doi:10.1167/14.13.14
27. Wegrzyn M, Vogt M, Kireclioglu B, Schneider J, Kissler J. Mapping the emotional face. How individual face parts contribute to successful emotion recognition. *PLoS One*. 2017;12(5):1-15. doi:10.1371/journal.pone.0177239
28. Friesen W, Ekman P. EMFACS-7: Emotional Facial Action Coding System, Version 7. Unpublished Manual. 1984.
29. Jeganathan J, Campbell M, Hyett M, Parker G, Breakspear M. Quantifying dynamic facial expressions under naturalistic conditions. *Elife*. 2022;11:1-20. doi:10.7554/eLife.79581
30. Daniel DG. Issues in selection of instruments to measure negative symptoms. *Schizophr Res*. 2013;150(2–3):343-345. doi:10.1016/j.schres.2013.07.005
31. Marder SR, Kirkpatrick B. Defining and measuring negative symptoms of schizophrenia in clinical trials. *Eur Neuropsychopharmacol*. 2014;24(5):737-743. doi:10.1016/j.euroneuro.2013.10.016
32. Kirkpatrick B, Strauss GP, Nguyen L, et al. The brief negative symptom scale: psychometric properties. *Schizophr Bull*. 2011;37(2):300-305. doi:10.1093/schbul/sbq059
33. Forbes C, Blanchard JJ, Bennett M, Horan WP, Kring A, Gur R. Initial development and preliminary validation of a new negative symptom measure: the clinical assessment interview for negative symptoms (CAINS). *Schizophr Res*. 2010;124(1–3):36-42. doi:10.1016/j.schres.2010.08.039

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Ambrosen KS, Lemvigh CK, Nielsen MØ, Glenthøj BY, Syeda WT, Ebdrup BH. Using computer vision of facial expressions to assess symptom domains and treatment response in antipsychotic-naïve patients with first-episode psychosis. *Acta Psychiatr Scand*. 2025;151(3):270-279. doi:10.1111/acps.13743