generalizability of clinical prediction models. Science (New York, N.Y.) 383, 164-167, doi:10.1126/science.adg8538 (2024). (4) Stroup, T. S. et al. The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project: schizophrenia trial design and protocol development. Schizophrenia bulletin 29, 15-31, doi:10.1093/oxfordjournals.schbul.a006986 (2003). (5) Fleischhacker, W. W., Keet, I. P. & Kahn, R. S. The European First Episode Schizophrenia Trial (EUFEST): rationale and design of the trial. Schizophr Res 78, 147-156, doi:10.1016/j.schres.2005.06.004 (2005).
No conflict of interest

https://doi.org/10.1016/j.nsa.2024.105181

P3534
NEUROSCIENCE APPLIED 3 (2024) 105180
AI-BASED MULTIMODAL TREATMENT RESPONSE PREDICTION IN SCHIZOPHRENIA: COMBINING INFLAMMATION AND IMAGING

P.A. Lalousis [1], A. Curzon [2], A. Malaviya [2], A. Egerton [1], J. MacCabe [1], V. Mondelli [3], R. Drake [4], J. Walters [5], S. Lawrie [6], S.L. Griffiths [2], J. Rogers [2], N. Barnes [2], B. Deakin [4], G. Khandaker [7], J. Suckling [8], R. Upthegrove [2]. [1] *King's College London, Psychosis Studies, London, United Kingdom;* [2] *University of Birmingham, Institute for Mental Health, Birmingham, United Kingdom;* [3] *King's College London, Institute of Psychiatry- Psychology and Neuroscience, London, United Kingdom;* [4] *University of Manchester, n/a, Manchester, United Kingdom;* [5] *Cardiff University, n/a, Cardiff, United Kingdom;* [6] *University of Edinburgh, n/a, Edinburgh, United Kingdom;* [7] *University of Bristol, n/a, Bristol, United Kingdom;* [8] *University of Cambridge, n/a, Cambridge, United Kingdom*

Background: The most dominant hypothesis for the cause of schizophrenia is excess of striatal dopamine, as suggested by the effectiveness of antipsychotic medications which act to block dopamine, mainly at the D2 receptor. However, approximately one third of patients fail to respond to D2 antagonists, whilst exhibiting similar antipsychotic dopamine receptor binding as responsive patients. Therefore, the non-responsive cohort may have illness driven by a different pathophysiological mechanism. One suggested theory is that non-responsive patients have more prominent immune system dysregulation than antipsychotic responders and that this may be related to more pronounced structural brain changes. The aim of this project was to develop Support Vector Machine (SVM) learning models of treatment response based on inflammatory markers and structural imaging data and examine their predictive utility in an external sample.
Method: We analysed data from the Schizophrenia: Treatment Resistance and Therapeutic Advances (STRATA) consortium and the Benefit of Minocycline on Negative Symptoms of Psychosis: Extent and Mechanism (BeneMin) datasets as part of the Psychosis Immune Stratified Medicine (PIMS) collaboration. Individuals were categorised as treatment responders if they met the following criteria: (i) they had been treated with a single antipsychotic medication since the onset of their condition, or any changes in medication were solely due to side effects rather than lack of effectiveness; (ii) they achieved a CGI-SCH score lower than 4; (iii) their PANSS total score was below 60; (iv) they had a compliance rating higher than 3. Conversely, the criteria for being classified as a treatment non-responder included (i) undergoing treatment with at least two different antipsychotics for over four weeks each, at doses above the minimum therapeutic level as specified by the British National Formulary; (ii) a CGI-SCH score of more than 3; (iii) a PANSS total score of 70 or higher; (iv) a compliance rating score higher than 3. An SVM model that classified treatment response vs non-treatment response was developed using the STRATA) dataset. The model was trained in a repeated nested pooled cross-validation framework with 5 outer CV2 permutations, 5 outer CV2 folds, 5 inner CV1 permutations, and 5 inner CV1 folds. The model was then externally validated to the BeneMin dataset in order to test its predictive utility.
Results: Combining inflammatory and imaging data the SVM model classified treatment response vs non-treatment response with a balanced accuracy of 70.8 and an area under the curve of 0.71. Patients classified in the treatment non-response group in the external validation sample showed no statistically significant differences at baseline. However, 12 months after baseline non-responders exhibited higher PANSS positive, PANSS general, and PANSS total scores.
Conclusion: This is, to our knowledge, the first study to date to incorporate inflammation and neuroimaging data to predict treatment response in schizophrenia. Our models may be able to aid the classification of treatment response in patients with schizophrenia. Our findings contribute to a growing body of evidence suggesting that immune system dysregulation may play a role in the pathophysiological mechanisms of treatment non-response in schizophrenia, with insight for targeted approaches in the future.
References
Conflict of interest
Paris Alexandros Lalousis has received honoraria for talks presented at educational meetings organised by Boehringer-Ingelheim outside of the presented work.

https://doi.org/10.1016/j.nsa.2024.105180

P3535
NEUROSCIENCE APPLIED 3 (2024) 105181
EXPLORING THE IMPACT OF CHATGPT IN PSYCHOEDUCATION FOR INDIVIDUALS WITH MOOD DISORDERS

F. Attanasio [1], G.A. Obumselu [2], C. Antonini [2], M. Prato [2], N. Lanzano [2], V. Fazio [2], L. Fregna [1], C. Colombo [2]. [1] *Sapienza University of Rome, Department of Human Neuroscience, Rome, Italy;* [2] *Vita-Salute San Raffaele University, Division of Clinical Neurosciences, Milan, Italy*

Introduction: Psychoeducation is a widely used intervention in psychiatry, based on the notion that providing patients with detailed information about illness and coping strategies can improve treatment outcome [1]. Recently, the rapid advancement of new technologies, such as artificial intelligence (AI), and chatbots, has made these tools readily available sources of information, even in the medical field, easily accessible to the general population [2]. According to recent literature, patients' requests for psychoeducational information from chatbots can serve as a support in the therapeutic process, but further studies are needed to confirm these tools' utility [3-4].
Aims: This study aims to evaluate the capability of ChatGPT to provide psychoeducational responses to mood disorders patients. The effectiveness of ChatGPT's responses will be compared and analyzed based on feedback from three distinct groups: patients, Psychiatric Rehabilitation Technicians (PRTs), and psychiatrists. The goal is to determine whether the responses from ChatGPT satisfactorily support the management of the condition according to mental health professionals' standards.
Methods: We recruited 28 subjects consecutively admitted to the Mood Disorder Unit of San Raffaele Hospital, Milan. Subjects were aged 18-75 years, diagnosed with Major Depressive Disorder or Bipolar Disorder, depressive episode, according to the DSM-5-TR, and provided informed consent. We excluded patients with schizophrenia and related disorders, intellectual disabilities and cognitive impairment. Two trained psychiatrists collected participants' sociodemographic and clinical data, and 5 open-ended questions formulated by each participant about their psychiatric condition that were submitted to OpenAI ChatGPT with GPT 4.0 architecture. ChatGPT's answers were examined with 5-point Likert Scales by participants, psychiatrists and PRT in blind, based on 5 parameters: relevance, comprehensibility, usefulness, empathy and acceptance. ANOVA were performed to evaluate differences in mean scores between groups.
Results: The analysis revealed significant differences in the total scores among the groups ($p<0.001$, ANOVA). Patients assigned the highest mean scores (22.46, SD 2.67), followed by educators (17.89, SD 3.35). Psychiatrists gave the lowest ratings, with an average score of 11.6 (SD 3.25). Comparisons for each category are presented in the table.
Conclusions: Evaluating ChatGPT's effectiveness in providing psychoeducational support reveals significant discrepancies between patients' and healthcare professionals' ratings. It is crucial to acknowledge this gap; patients may perceive responses as helpful but lack the necessary tools to assess their accuracy, potentially leading to harmful misinterpretations. Healthcare professionals demonstrate a more critical stance, highlighting deficiencies that may not be evident to patients. Despite patients' generally positive assessments in comprehensibility, relevance, and usefulness, professional training profoundly influences the perception of chatbot interactions' quality. These findings

underscore the need for cautious integration of such technologies, continuously evaluating their effectiveness and safety in collaboration with healthcare experts. While promising, chatbots should not replace the need for qualified human intervention, especially in sensitive fields like psychiatry.

	Patients	Psychiatrists	PRTs	p-value
Relevance, mean (SD)	4.48 (0.48)	2.75 (0.89	4.25 (0.93)	<0.001*
Comprehensibility, mean (SD)	4.71 (0.54)	2.93 (0.86)	4.18 (0.94)	<0.001*
Usefulness, mean (SD)	4.39 (0.83)	2.75 (1.00)	4.25 (0.89)	<0.001*
Empathy, mean (SD)	4.36 (0.87)	1.57 (0.50)	2.54 (0.88)	<0.001*
Acceptance, mean (SD)	4.39 (0.83)	1.61 (0.50)	2.68 (0.98)	<0.001*

* Anova

References

1. Colom, F., Vieta, E., Martínez-Arán, A., Reinares, M., Goikolea, J.M., Benabarre, A., Torrent, C., Comes, M., Corbella, B., Parramon, G., & Corominas, J., 2003. A randomized trial on the efficacy of group psychoeducation in the prophylaxis of recurrences in bipolar patients whose disease is in remission. Archives of General Psychiatry, 60(4), 402-407. 2. Xue, J., Zhang, B., Zhao, Y., Zhang, Q., Zheng, C., Jiang, J., Li, H., Liu, N., Li, Z., Fu, W., Peng, Y., Logan, J., Zhang, J., & Xiang, X., 2023. Evaluation of the Current State of Chatbots for Digital Health: Scoping Review. Journal of medical Internet research, 25, e47217. https://doi.org/10.2196/47217 3. Altamimi, I., Altamimi, A., Alhumimidi, A. S., Altamimi, A., & Temsah, M. H., 2023. Artificial Intelligence (AI) Chatbots in Medicine: A Supplement, Not a Substitute. Cureus, 15(6), e40922. https://doi.org/10.7759/cureus.40922 4. Pham, K. T., Nabizadeh, A., & Selek, S., 2022. Artificial Intelligence and Chatbots in Psychiatry. The Psychiatric quarterly, 93(1), 249–253. https://doi.org/10.1007/s11126-022-09973-8

No conflict of interest

https://doi.org/10.1016/j.nsa.2024.105181

P3536

NEUROSCIENCE APPLIED 3 (2024) 105403

THE USE OF COMPUTER-AIDED ANALYSIS OF FACIAL EXPRESSIONS BASED ON THE SUPPORT VECTOR MACHINE IN THE DIAGNOSIS OF ASD

K. Wilczyński [1], D. Pęszor [2], L. Cichoń [1], K. Wojciechowski [2], A. Kawalec [1], M. Janas-Kozik [1]. [1] *Medical University of Silesia, Department of Psychiatry and Psychotherapy of Developmental Age, Sosnowiec, Poland;* [2] *Polish-Japanese Academy of Information Technology, Research and development center, Warsaw, Poland*

Autism spectrum disorder (ASD) is a disorder characterized by communication deficits, especially in the field of non-verbal language i.a. in terms of facial expressivity. However, a shortage of proper, accurate methodology, caused a relatively small amount of research in this area. Available literature mostly agree, that patients with ASD manifest a dissociation between verbal and affective messages, and overall facial expression is either lowered or inappropriate. In the current literature, facial expressions in ASD have been rated primarily by extensively trained beholders what lowers the accuracy of assessment and may lead to i.a. confirmation biases. Our aim was to develop a model of acquisition and computer analysis of facial expressions and to verify its effectiveness in identifying people with an ASD diagnosis.

The study included 100 people divided into two groups in terms of ASD diagnosis. The study group consisted of 73 people, of which 51 (69.8%) were men, and the control group consisted of 27 people, including 16 (59.2%) men. The diagnosis of ASD was made by a specialist child and adolescent psychiatrist based on a developmental history and mental state examination and then confirmed using the ADOS-2 protocol. In the control group, the diagnosis of ASD was excluded based on the ADOS-2 protocol and developmental history. A statistically significant age difference was observed between the study group (14; 95% CI: 13.5-14.5 years) and the control group (16.3; 95%CI: 15.2-17.5 years) in the Mann-Whitney's U test.

All participants were subjected to a series of 3 exercises, including a semi-structured conversation based on a previously prepared scenario, recognizing facial expressions displayed on the screen and then imitating them based on the displayed images. Throughout the procedure, the participants' faces were recorded using 5 cameras placed in front of and to the sides of the test subject. Faces were then detected in the images using a sliding window algorithm in a multi-resolution representation of the Gaussian pyramid utilizing a linear classifier based on the Support Vector Machine (SVM) with a classical Histogram of Oriented Gradients (HOG) descriptor. A set of Ensemble of Regression Trees was used on the detected faces to fit the model of characteristic points on the face to each frame of the image. Based on characteristic points, anthropometric distances and their proportions were determined. On standardized anthropometric proportions, the SVM classifier was trained.

The obtained model was able to predict the diagnosis of ASD in the study population with almost 100% accuracy. The mean difference between the probability of the correct class and the probability of the incorrect class determined by the SVM on the test set was 56%.

Analysis of facial expressions using the trained SVM classifier seem to be a promising tool for ASD diagnosis. What's more, the methodology utilized allows it to be used f.e. on a smartphone. However, a full assessment of clinical suitability, especially in the context of non-standardised devices, requires further research. The obtained results also confirm the hypothesis that the presentation of facial expressions in ASD is significantly different from the neurotypical population.

References

Conflict of interest

Disclosure statement:

The study was financed by the Katowice City Hall as part of the Metropolitan Grant.

https://doi.org/10.1016/j.nsa.2024.105403

P3537

NEUROSCIENCE APPLIED 3 (2024) 105182

PREDICTION OF CLINICAL TOOLS SCORES USING TRANSFORMER-BASED MODELS IN PATIENTS WITH FIRST-EPISODE PSYCHOSIS

A. Periša [1], M. Arcan [2], D.P. Niland [3], F. Delahunty [3], D. Polšek [1], A. Savić [1,4], P. Brečić [1,4], J. Vukojević [1]. [1] *University Psychiatric Hospital Vrapče, Department for Diagnostics and Intensive Care, Zagreb, Croatia;* [2] *University of Galway, Insight SFI Research Centre for Data Analytics- Data Science Institute, Galway, Ireland;* [3] *Lua Health, Galway, Ireland;* [4] *University of Zagreb, School of Medicine, Zagreb, Croatia*

Introduction: Despite extensive research on applying machine learning to the prediction, diagnosis, and monitoring of mental disorders, clinical assessment remains the core tool in diagnosing and monitoring illness. Self-report instruments like the Patient Health Questionnaire (PHQ-9) for depression and the General Anxiety Disorder-7 (GAD-7) scale for anxiety disorders, as well as clinical instruments like the Positive and Negative Syndrome Scale (PANSS) for psychotic disorders and the Calgary Depression Scale for depressive symptoms in psychosis, are widely used in both research and everyday clinical practice. Early identification of risk and prevention of psychotic relapse has become a primary goal in the management of first-episode psychosis, particularly as signs of psychosis relapse might manifest nonspecifically, such as anxiety, mood swings, insomnia, or impairment in general functioning [1]. To date, linguistic features analyzed using natural language processing (NLP) have shown great predictive value in predicting depression and anxiety levels [2] or psychosis relapse, as it has been found that linguistic disturbances can be detected several weeks before relapse [3].

Aim: This study aimed to test whether our transformer-based model could predict outcome scores on clinical tools based on unstructured written diaries.

Method: Over the course of six months, patients experiencing a first episode of psychosis receiving treatment at the day hospital of the University Psychiatric Hospital Vrapče, Zagreb, were instructed to maintain weekly diaries of their subjective mental state and daily activities. Our dataset included 93 diaries from 21 inpatients, totaling 11,878 words with an average of 128 words per diary. The diaries were evaluated in relation to scores from PHQ-9, GAD-7, PANSS, UCLA Loneliness Scale, Global Assessment of Functioning (GAF) Scale, and Calgary Depression Scale. Weekly diaries were categorized based on PHQ-9 and GAD-7 questionnaire scores. Non-symptomatic and mild symptomatic cases were grouped into Class 0, while moderate to severe cases were in Class 1. The study employs Transformer models, particularly BERT and RoBERTa, renowned for their efficacy in NLP tasks, to capture contextual dependencies in input sequences and assess their performance. Since PANSS, UCLA, and GAF are continuous