Objectives: This pilot study is part of a new project "RESOUT - Smartphone app for RESident physicians' burnOUT monitoring and prevention" (FCT 2023.13877.PEX) and aims to explore levels of residents' burnout and to identify its individual and contextual correlates, in order to empirically inform the assessment and intervention strategies to be included in the app and the secondary outcomes to considered when analysing its effectiveness.
Methods: To date, 300 Pt residents (75% women) completed questionnaires to evaluate Burnout, Psychological distress/PD (depression/anxiety/stress) symptoms and attributions, Perfectionism, Worry and rumination/WR, Self-compassion/SC, Perceived Life Satisfaction/PLS and Occupational Satisfaction/POS, and Perceived-self-efficacy/PSE, -empathy/PE and -medical errors/PME.
Results: More than half of the sample had moderate (35.2%) / high (21.6%) burnout levels. Among these, 38.9% report thinking about leaving residency often/very, being the most mentioned reasons to do it: high workload (72%), lack of personal/family time (73.6%) and rewards (51%) and difficulty dealing with stress (48.3%).
Perfectionism, WR and low SC were predictors of burnout (betas>.40). SC mediates the relationship between the other predictors and burnout. Both perfectionism and SC moderate the relationship between have mentioned those reasons and burnout.
Burnout (or at least one of its dimensions) predicted low PLS (-.589), low PRS (-.524), PD (.677), frequency of leaving residency thoughts (.583), low PSE (-.324), PE (.665) and PME (.224) (all $p \leq .001$).
Conclusions: These findings confirm that provide residents with SC practices may help them manage perfectionism and reduce burnout symptoms and its negative impact on them and their patients. Considering they attribute their distress to the high job demands they face, digital interventions – accessible, flexible, AI-powered personalized - can promote their adherence.

Disclosure of Interest: None Declared

EPP222

Testing Empathy in Machines: A Comparative Study of Large Language Models and Motivational Interviewing

J. B. Rodrigues*, J. P. Azenha, P. Trindade and C. Adão

Departamento de Psiquiatria e Saúde Mental, Unidade Local de Saúde de Lisboa Ocidental, Lisboa, Portugal
*Corresponding author.
doi: 10.1192/j.eurpsy.2026.12058

Introduction: Large language models (LLMs) are increasingly proposed as tools for behavioral healthcare, with the potential to augment or even automate aspects of psychotherapy. Among therapeutic frameworks, motivational interviewing (MI) is especially relevant as it is well-validated for eliciting behavior change and improving adherence. MI's structured, empathetic style provides a rigorous benchmark for assessing whether LLMs can replicate therapeutic dialogue. While LLMs promise scalability and accessibility, their application in psychotherapy remains high-stakes, raising concerns of safety, fidelity, and ethical deployment.
Objectives: This study aims to evaluate whether current leading LLMs reproduce the principles of MI, focusing on empathy, change talk, rolling with resistance, and support for autonomy.
Methods: A standardized MI script addressing ambivalence toward alcohol reduction was applied across five leading LLMs: GPT-4, GPT-3.5, Claude, Gemini, and Llama-2. Outputs were coded against MI core processes (engaging, focusing, evoking, planning), MI principles (empathy, discrepancy, resistance management, non-confrontation, self-efficacy), and common techniques (open questions, reflections, affirmations, eliciting change talk, avoidance of directive advice).
Results: GPT-4 most consistently embodied MI: it engaged and focused effectively, elicited change talk, and demonstrated empathy through accurate reflections. It avoided confrontation and supported self-efficacy, applying all five principles with notable fidelity. Claude showed similar strengths, especially in empathy and discrepancy development, but was less reliable in rolling with resistance. Gemini balanced empathy with structure, producing clear open questions and affirmations, but defaulted to directive recommendations, limiting adherence to non-confrontation and autonomy support. GPT-3.5 generated open questions but frequently lapsed into advice-giving, undermining MI principles and the planning process. Llama-2 showed the lowest fidelity, often offering prescriptive advice rather than reflections, with minimal engagement in evoking change talk or consolidating commitment. Across models, nuanced techniques such as managing resistance and strengthening commitment were inconsistently applied, with GPT-4 and Claude demonstrating the highest overall fidelity to MI principles and techniques.
Conclusions: LLMs can reproduce surface features of MI and show promise for scaling access to supportive dialogue. However, consistent therapeutic stance and deeper MI techniques remain beyond current capacity. These findings echo concerns that while LLMs may simulate skills, they do not yet deliver the clinical sensitivity required for psychotherapy. Responsible development should prioritize domain-specific training, interdisciplinary oversight, and hybrid models where clinician supervision ensures therapeutic integrity.

Disclosure of Interest: None Declared

EPP223

Impact of Remote Patient Monitoring On Major Depressive episode And Bipolar Disorder Treatment Outcomes: EC-102 Randomized Controlled Trial

A. Yrondi[1], A. Bourla[2,3,4,5], R. Belzeaux[6,7], E. Olié[8,9,10], R. DAVID[11,12], J. Mallet[13,14,15], A. Amad[16], A. Sauvaget[17], M. Polosan[18,19], J. Holtzmann[20], F. Chevrier[21], E. Haffen[22,23], D. Bennabi[22,23], F. Stéphan[24], E. Artiges[25], A. Amrandi-Chickh[25], S. Bulteau[26,27], W. El-Hage[28,29], V. Rousseau[30], N. Lele[31], C. Martelli[32], L. Mekaoui[33], D. Jacon[34], G. Chabridon[35], L. Dormegny Jeanjean[36,37], F. Hermelin[38], E. Touré Cuq[39]* and P.-M. Llorca[40,41]

[1]Service de Psychiatrie et de Psychologie Médicale de l'adulte, Centre Expert Dépression Résistante FondaMental, CHU de Toulouse, Hôpital Purpan, ToNIC Toulouse NeuroImaging Centre, Université de Toulouse, INSERM, UPS, Toulouse; [2]Department of Psychiatry, Saint-Antoine Hospital, Sorbonne University; [3]Assistance publique Hôpitaux de Pari; [4]Medical Strategy and Innovation Department, Clariane; [5]NeuroStim Psychiatry Practice, Paris; [6]IGF, Univ. Montpellier, CNRS, INSERM; [7]Service de psychiatrie; [8]Service de Psychiatrie d'Urgence et de Soins Aigus, CHU Montpellier; [9]Université Montpellier; [10]IGF, CNRS, INSERM, Montpellier; [11]CHU de Nice; [12]Université Côte d'Azur, Nice; [13]Service de psychiatrie adulte, Hôpital Louis Mourier, AP-HP, Paris; [14]Service de psychiatrie adulte, EPSM du Loiret, Centre Hospitalier Universitaire d'Orléans; [15]B-CLINE, Laboratoire Interdisciplinaire pour l'Innovation et la Recherche en Santé d'Orléans (LI²RSO), Université d'Orléans, Orléans; [16]U1172 - LilNCog - Lille Neuroscience & Cognition, F-59000, CHU Lille, Lille; [17]EPSM de Vendée, Centre Hospitalier Georges Mazurelle, La-Roche-sur-Yon; [18]Inserm U1216, Univ. Grenoble Alpes; [19]Grenoble Institut de Neurosciences, CHU de Grenoble; [20]Service de Psychiatrie Adulte, CHU Grenoble Alpes,