Project title: Can digital communication messages and clinical data be used to create a human digital twin for patients with psychotic disorders?

Project reference: DT4H_13_2024

1st supervisor: Sarah Morgan - Biomedical Engineering and Imaging Sciences

2nd supervisor: Hector Menendez Benito - Informatics

3rd supervisor: Kelly Diederen - Institute of Psychiatry, Psychology and Neuroscience

4th supervisor: Thomas Spencer - Institute of Psychiatry, Psychology and Neuroscience

Aim of the Project

This project aims to create a human digital twin (Lin, 2024) for patients with psychotic disorders, from both clinical data extracted from patients' Electronic Health Records and Natural Language Processing (NLP) features extracted from their digital communication (e.g. Whatsapp messages).

The first part of the project will focus on developing an app to extract summary NLP features from previously sent text or Whatsapp messages, which the student will then use to collect timeseries of NLP features from patients with psychotic disorders. The student will then create a human digital twin, combining NLP features and clinical data to build a detailed, multimodal representation of the patient. Finally, we will assess whether the digital twin can be used to predict psychotic episodes and hospitalisations that occurred at a later timepoint. Ultimately, predicting individual trajectories for patients with psychotic disorders ahead of time could provide new opportunities for preventive interventions.

Project Description

There is pressing clinical demand for tools that can predict the course of psychotic disorders at the individual patient level, enabling earlier, preventive interventions, yet such tools have proved elusive (Coutts *et al.*, 2023). Our prior work showed that Natural Language Processing (NLP) markers of altered language use such as reduced semantic coherence have significant predictive power for psychosis (Morgan *et al.*, 2021; Spencer *et al.*, 2021). However, most studies focus on group-level differences in NLP metrics, rather than modelling longitudinal changes within individual patients.

This project aims to bridge that gap by building digital twins of patients that integrate language and clinical data to predict significant shifts in clinical state ahead of time. To that end, we will combine anonymised NLP features from Whatsapp messages and clinical records, creating human behavioural models that profile the temporal evolution of patients (Ferdousi *et al.*, 2021). The ultimate goals of the digital twin model are to predict future disease trajectories and identify timepoints when clinical interventions might be most effective.

Our first aim is to develop an app that extracts summary NLP features, such as semantic coherence, from previously sent Whatsapp messages on patients' mobile devices. We will use this app to collect retrospective NLP data from N=30 patients with psychotic disorders and N=30 matched healthy control subjects, and link the resulting data to Electronic Health Record (EHR) data on a secure server.

We will then build a human digital twin behavioural model, combining social interaction data (longitudinal NLP measures) and clinical data. The digital twin will leverage Dr Menendez's prior work on profile-based modelling (Rodriguez *et al.*, 2017), to track multi-modal data over time. Validation will involve testing the model's ability to forecast future language features and detect differences between patients and controls. The longitudinal data will allow us to model the directionality of relationships between NLP measures and clinical state, which could hint at causal links. Finally, we will assess the digital twin's ability to predict psychotic episodes and hospitalisations in advance.

People with lived experience of psychosis will be involved throughout and the student will receive training in co-creation of Digital Health research. The supervisors have substantial experience of patient and public involvement in similar projects, e.g. co-designing an app to record speech data from psychosis patients.

If successful, this line of research could produce new prediction models for psychosis, enabling patients and clinicians to take earlier preventive action. The tools developed here could also be adapted for other brain disorders, e.g. depression.

Requirements:

This project would suit a student with good programming skills, a first degree in e.g. Computer Science, Engineering, Mathematics or Physics, and an interest in AI for mental health applications.

Suggested reading:

Coutts *et al.*, Nature Reviews Neurology (2023), 19, 221-234.

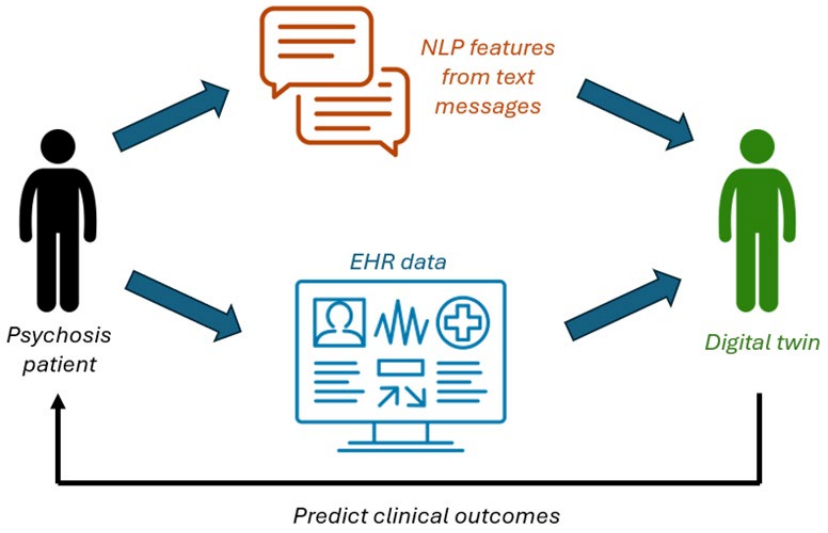
Ferdousi *et al.*, IEEE Globecom Workshops (2021), pp. 1-6.

Lin *et al.* Journal of Cloud Computing (2024), 13, 131.

Morgan *et al.*, Translational Psychiatry (2021), 11, 630.

Spencer *et al.*, Schizophrenia Research (2021), 228, 493-501.

Rodríguez-Fernández, Menéndez *et al.*, Expert Systems with Applications (2017), 70, 103-118.


NLP features from text messages
EHR data
Psychosis patient
Digital twin
Predict clinical outcomes