



## Our Promise

We work towards including as many aspects as possible in our design to achieve a well-rounded examination, and to fully exploit the useful information contained in the valuable data provided by the patient.

We value our users' data privacy and guarantee a protection. We do not gather and apply the patients' data on any training without authorisation.

## KEY FEATURES



A webapp enabling the online automatic examination process and the display of statistical results and suggestions to the patient.



A visualisation of illness signifiers throughout the length of the conversation.

## MODEL STRUCTURE

**Sentiment Classifier** Interprets voice tone using a graph-type neural network along with a pretrained BERT model for language processing

**Illness-signifiers Recogniser** Various combinations of training methods and voice features considered:  
Spectrogram + Convolutional Neural Networks; Spectrogram / Mel-frequency cepstral coefficients (MFCCs) + Recurrent Neural Networks (LSTMs); Extracted spectral and prosodic features + Gaussian Mixture Models

**Text Analyser** A pretrained BERT model fine-tuned on a symptom-disease dataset for disease classification

Patient  
Voice AI

Patient Voice AI



# CONTACT INFO

This proposal is part of the larger MF Technology AI teleconsultation project, which aims to use audiovisual information from patient video calls to derive vital signs and additional health indicators through audio processing and deep learning. Specifically, this proposal focuses on evaluating the feasibility of an AI algorithm to assess patients' health by analyzing their voice and processing their speech. The goal is to develop an AI algorithm that can process existing audio recordings from a given dataset, interpret sound and voice tone, understand conversations, and analyze sentiment.

## OUR OBJECTIVES

Recent statistics indicate a significant decline in NHS Referral to Treatment (RTT) times, with the percentage of patients treated within 18 weeks dropping from about 94% in 2013/14 to around 58% in 2023/24, and median wait times increasing from 5.5-6.3 weeks to 13.8-15.0 weeks. The 92nd percentile wait times have also worsened from 15.8-17.1 weeks to 44.6-46.4 weeks. To address these issues, virtual teleconsultation is gaining popularity, enabling clinicians to remotely examine patients and capture vital health data through wearable devices. By incorporating video and audio cues during consultations, clinicians can gain further insights into patient health. Therefore, we believe that our audio-video consultation platform with multimodal AI capabilities will significantly enhance remote healthcare by providing valuable insights for better and timely clinical decisions.

# OUR MOTIVATION

We are a speech processing and machine learning technology research group consisting of six Imperial College London Electrical and Electronic Engineers dedicated to putting their knowledge acquired throughout their studies into practice to make a difference to the real world.