

JC-2898 July, 6<sup>th</sup> 2021

## Research Protocol

- (1) Laying of medical behavior data set platform: purchase hardware equipment, complete the laying of medical behavior data set platform dedicated to the research group, install data desensitization platform and data annotation platform, and design special app for patients' lifestyle follow-up.
- (2) Construction of acquisition platform: according to the scene of medical behavior to be collected, the corresponding acquisition platform is selectively built, in which several rooms in the emergency room and comprehensive ward are selected to build a relatively fixed acquisition platform, including video acquisition device (equipped with vision, telephoto or wide-angle camera according to the scene installation and acquisition conditions), TOF camera, motion camera, etc, Audio collector and CDR system data acquisition device. The intensive care unit (ICU) is equipped with wearable acquisition equipment, audio acquisition device and CDR system data acquisition device. After discharge, the patients were followed up in outpatient department. They took the initiative to make follow-up calls and upload medical monitoring data through special app.
- (3) Collection of life style, medical behavior and medical outcome: as shown in Table 1, medical behavior is to collect all-round and multi-dimensional medical behavior of patients in emergency rescue room, Comprehensive Ward and intensive care unit, including video, audio, monitoring records, synchronous medical records, laboratory examination, imaging data of hospital his system and PACS system The data of critical care system, and the intervention of patients' life style during follow-up. The video machine is used for medical staff to operate panoramic scene and close range limb operation. It adopts manual or automatic acquisition, and the acquisition time is all the time of the patient in the area. It includes operation video and dialogue audio. For each

kind of medical behavior, the video machine is for medical staff to operate panoramic scene and close range limb operation, and the audio is recorded by close range microphone, which is collected manually or automatically. The collected data is uploaded to the medical intervention behavior data set platform through a special data transmission channel. The life style and final outcome of the patients were recorded in the late follow-up, through the combination of outpatient follow-up, active telephone follow-up and the use of special app by the patients themselves.

- (4) Data desensitization and data quality evaluation: desensitization of medical behavior video and audio data is carried out by using the de privacy algorithm of differential mechanism researched and developed by the research group. The researcher will evaluate the quality of video and audio data after desensitization, and the qualification criteria are: the whole process of recording medical behavior by video file is basically complete, the picture is clear, and it can be labeled and machine learning; The content of audio data conversation is clear, speech recognition is accurate, and the audio data after sensitive words are removed is basically consistent, which can establish association with video files. If the quality of video and audio desensitization does not cooperate, or the association cannot be established, it will be regarded as unqualified data and will be eliminated.
- (5) Data annotation and missing dimension completion: using the annotation platform for medical behavior data file coarse-grained annotation. According to the theory of multi-dimensional space-time collaborative representation, the missing dimensions of the initial data are supplemented by multi-dimensional active perception technology, and the medical device parameters or monitoring data, medical text records and outcome data collected or recorded in the process of medical behavior are collected. The de privacy algorithm of differential mechanism is used to desensitize the non video data of medical behavior.
- (6) Data representation and database construction: through the completeness and standardization representation method of medical behavior, the fine-grained formation characteristics such as body movements, movement patterns and fine behaviors of individuals related to medical behavior are extracted. According to the regulations of

the people's Republic of China on the management of human genetic resources, the open data specification of ethical compliance is formed, and the database is established, Build a large-scale open medical behavior data set.

3.4 selection and confirmation of main measurement indicators or outcome indicators Data volume of medical behavior data set: the open large-scale multimodal medical behavior data set constructed in this study, the labeled data volume of medical behavior (including follow-up data) of patients in intensive care unit and Comprehensive Ward reached 10TB, and the collection of medical behavior was stopped automatically. The medical behavior (including follow-up data) of the patients who were not hospitalized in the emergency room after labeling reached 2TB, and the collection of medical behavior was stopped automatically.