**Summary of Background Literature**

In 2019, Hibbert and Umeakunne published a manually calculated recommended daily allowance table of nutrients for people with SCD. [3] However, there is still a gap, as there is no A.I system to help track those specific nutrient intakes or to examine meals for possible interaction with sickle cell disease medication.

The Personalized Diet Recommendation System by Hussain et al., aims to assist cancer patients in planning their daily diets. It employs Case-based Reasoning, Rule-based Reasoning, and genetic algorithms to create customized diet menus based on individual health information. Users input data such as cancer type, treatment stage, activity level, food preferences, allergies, ethnicity, and side effects to build their profiles. The outcome is a system that recommends a personalized daily dietary allowance for energy, carbohydrates, protein, calcium, thiamin, niacin, riboflavin, vitamin A, vitamin C, vitamin D, and vitamin E. Based on these recommendations, the system suggests breakfast, lunch, and dinner menus tailored to the user's specific needs.

**The rationale for the proposed study, including specific aims and hypothesis.**

The aim is to demonstrate that A. I can be used to provide nutrition therapy to Ghanaians with sickle cell disease. Hence, a generalization can be made that A.I can be used to provide nutrition therapy for all individuals with sickle cell disease. It also aims to contrast and compare two dominant approaches to providing feedback reports to humans. The first is Large Language Models, and the second is rule-based expert systems.

**Hypothesis:**

* *Do Knowledge Graph Augmented Large Language models Perform Better than traditional rule-based expert systems in the Nutrition Therapy of Ghanaians with Sickle Cell Disease?*
* *Can a Rule-Based Expert system generate similar feedback reports to a Knowledge Graph Augmented Large Language Model in the Nutrition Therapy of Ghanaians with sickle cell disease?*

Human Participants, Budget and Scheduling

There would be no human participants in the experiment, as age, sex, and stomach content are all randomly generated test data for the experiment. All other data, such as food images and nutrient information, are obtained from publicly available data repositories. This project has no foreseeable financial expenditure, as the data, API, and computing resources to be used are free. The timeline for this research is pegged at 5 months.

**A detailed description of the study methodology**

**Data Collection**

5000 images of various African meals and Ghanaian meals were downloaded from RoboFlow and annotated using CVAT, which stands for Computer Vision Annotation Tool. Nutrient intake allowances for persons with SCD for protein, carbohydrates, omega-3 fatty acids, iron, vitamins B6, vitamin B12, vitamin A, vitamin C, vitamin D, vitamin E, folate, magnesium, zinc, and selenium. (Sub-divided by age and gender) was obtained from the only publicly available recommended daily intake table for sickle cell disease, published by Hibbert and Umeakunne in 2019.

Sickle Cell Disease medication names, compounds, and their possible interactions with certain foods were obtained from the online databases of the American Center for Disease Control, papers from the American National Library of Medicine, and information on approved treatment medications for sickle cell disease from the sickle cell anemia news platform. Chemical compounds present in various foods were obtained from FooDB. FooDB is an openly available internet-based repository of foods and their compounds and nutrients. The Database contains information about approximately 900,000 foods.

Nutrients found in Ghanaian foods per serving of 100 grams were obtained from Kaggle, My Fitness Pal Database, and numerous food blog websites. User information such as stomach content, weight, height, age, and sex are randomly assigned upon the program's start.

**Experiment Setup**

Using YOLO V8, a CNN-ResNet50 image recognition model was built and trained on 5000 images of Ghanaian and Ghanaian foods. The model was trained with epoch=100 to perform instance segmentation and image classification. The food image data was split into training and testing sets, and after the model was trained, the pre-labeled testing data was fed to the image classification model. Accuracy, Precision, Recall, and F-1 scores were calculated for the test data. The next step in the computer vision pipeline was to pass the segmented images into a portion estimator. The segmented foods' portion was estimated using Python's Zoe Depth Estimation model.

**Model Configuration**

**Traditional Expert System:** Using a Python programming language, we designed a rule-based system with a set of predefined rules and logical inferences that uses data collected in the data collection stage to advise a user on what can be safely consumed or what needs to be removed from their plate, to adhere to Recommended Dietary Intakes.

**Knowledge Graph-Augmented LLM:** We utilized a pre-trained LLM, specifically Open-AI's Davinci Text GPT, and augmented it with a knowledge graph representing domain-specific relationships, such as the relationship between foods, a user's stomach, a user's medication, and chemical compounds found in food.

We assessed five meals using the A.I system, and in each case, a report was generated for the user in natural English language. This report critically discussed how eating the meal could impact the user and what steps they could take regarding portioning. The result of the report was then contrasted and compared with that written by the Knowledge Graph Augmented LLM after being fed the same information via a prompt. The portioning and meal names information is fed into the prompt string using data returned by the computer vision model.

An Example of such a Prompt is:

*“I am a 37 Year Old Female with Sickle Cell Disease. I have eaten 2 Plates of rice and 2 Chicken drumsticks today. I am about to eat 20g of Aprapransa and 10g of Meat for dinner, should I eat it?”*

Python's spaCy library was utilized to determine the similarity between the reports generated by the Traditional Expert System and those generated by the Knowledge-Graph Augmented Large Language Model.