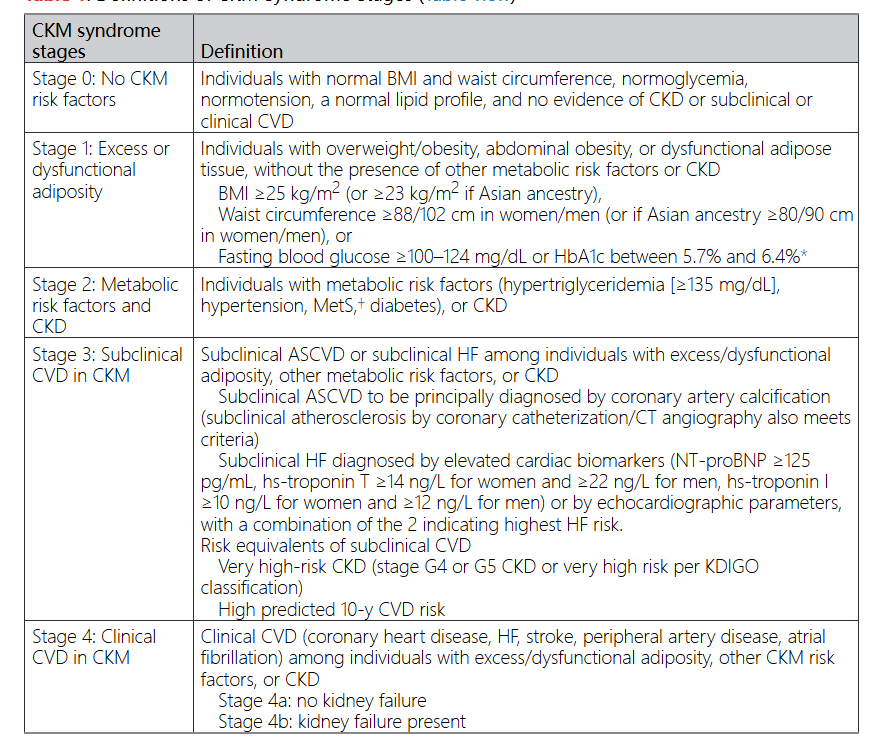
**方案概述（基于多模态大模型的CKM智能状态监测系统）**

**1. 目标：**

将来自UK Biobank的多模态数据（包括血检、尿检、Olink血浆蛋白、疾病基因风险数据、ECG、肾脏MRI、心脏超声数据等）转化为可解释的文本摘要报告。然后使用大语言模型（如ChatGPT），基于专业的CKM健康指南，动态地评估和更新患者的CKM状态。



**2. 输入数据处理及转化为文本摘要：**

* **数据转换**：为每个数据模态创建一套标准化的转换方法，将结构化或非结构化数据转化为关键文本描述或“报告摘要”。
  + **血液和尿液测试**：生成血糖、胆固醇、尿白蛋白等数值的描述性文本，突出偏离正常值的指标。
  + **基因风险评分**：用简明的语言描述疾病的遗传风险水平（低、中、高），作为初始CKM基线风险。
  + **Olink血浆蛋白**：根据蛋白质水平生成描述性摘要，特别关注与代谢、心脏和肾脏相关的标志物。
  + **ECG、MRI和超声数据**：提取关键的影像学信息，生成如“心脏结构无异常”或“左心室肥厚迹象”等简洁总结。
* **自动生成的报告模板**：
  + 每个阶段定义一套标准化的报告模板，以便输出一致的文本描述。例如：
    - **Stage 0**: 基因风险信息总结。
    - **Stage 1**: 血液、尿液和蛋白标志物的详细描述。
    - **Stage 2**: 增加心电图数据的解读。
    - **Stage 3**: 添加MRI和超声心动图的总结性描述。
    - **Stage 4**：添加临床诊断文本信息。（临床文本，UKB自带ICD code）

**3. 基于语言模型的状态评估与更新：**

使用ChatGPT或类似的大语言模型，结合CKM健康指南，自动读取各阶段生成的文本摘要并给出患者的CKM状态更新。具体过程如下：

* **状态分类**：模型分析当前的文本摘要并根据指南的标准将患者状态分为stage 0到stage4：（叶朝阳）
  + **这里是串联判别，stage0与stage1，stage1与stage2，stage2与stage3，stage3与stag4.（每一层级需要新增对应需要的新数据类型）**
* **动态状态更新**：语言模型依据最新的模态数据摘要和历史状态，更新患者的CKM阶段。随着每阶段的数据输入增加，模型逐步综合分析多个来源的信息，以便精确反映病人当前的健康状态。
* **基于指南的个性化建议**：模型还可提供建议，例如“定期检查血糖水平”或“开始进行超声检查以监测心脏结构”，以帮助制定下一步的诊疗计划。

**4. 输出：**

* **CKM状态总结**：模型输出当前的CKM阶段（如“Stage 2”），并附带简明的状态总结。
* **动态风险评分**：文本描述中附加了个体风险评分，评估其进展到下一个CKM阶段的可能性。
* **治疗与预防建议**：依据指南生成的干预措施（如“建议减重控制代谢风险”或“定期检测肾功能”）。

# 5.任务安排

**代码上传：https://github.com/labxscut/GWAS-Genetic-consult 正阳已是管理员**

**任务a1-1. SNP，人口，生化，图像报告模态数据的提取；交付：SNP，人口、生化的表格tsv；图像报告的文本txt（原文）；代码（python 命令行可运行）=》正阳 / 巴雷**

**任务a1-2. 对ukb样本的各模态数据按时间戳整理，交付：每个人一个文件夹，每次检查一个文件（tsv/txt)，文件名 ukbid\_yyyymmddhhmmss\_模态名.txt， ukbid\_yyyymmddhhmmss\_icd码.txt，代码（python 命令行可运行）=》正阳 / 巴雷**

**任务a2-1. 训练langchain + 大语言模型chatgpt，能够输入（疾病指南和CKM标准的综述pdf）作为上下文，根据提示准确判断分期； 交付：ukbid\_分期.tsv；代码（python 命令行可运行） =》任伯桢**

**任务a2-2. 每个stage随机取100例代表性结果的人工验证。交付图表：正确率，AUC =》任伯桢 / 莫老师**

**任务a3. 应用模型对ukb单人ckm演化分布进行统计分析，形成论文；等任务1和2初步有形后启动**

**任务b1. 对ukb图像模态数据的处理，从图像产生关键指标图像报告；交付：图像报告的文本txt；交付图表：正确率，AUC ；代码（python 命令行可运行）=》李嘉华 / 叶朝阳**

**任务b2. 把SNP数据融入模型，并提高模型性能**

**任务b3. 把olink数据融入模型，并提高模型性能参考文献**

[**https://www.ahajournals.org/doi/epub/10.1161/CIR.0000000000001184**](https://www.ahajournals.org/doi/epub/10.1161/CIR.0000000000001184)