**Presentation**

* Overview of the project goals, stakeholders, and scope (Cailin)
* System specification including a summary of functional and non-functional requirements, examples of representative user stories, technical constraints, initial system design/architecture, and demo of the first prototype. (Kiran)
* Quality of work including testing plan, acceptance criteria /tests, adopted acceptance techniques and use an application of discipline knowledge and/or project discipline knowledge or technology (e.g., computer science, advanced computer science, software development/engineering, data science/analytics, robotics, drones) (Bronte, Andersen)
* Plan for remaining requirements/user stories, and potential project risks (zheyuan)
* Group processes including collaboration and roles, client interaction, and potential improvements (Benjamin)
* A video demo of the prototype that the group developed as per agreed with the client (to be included in the submission along with the presentation slides) (Zichen)

# Intro – Cailin

The project’s aim is to evaluate the product qXR. qXR is a highly reputable technology, with the ability to detect multiple abnormal findings on a chest X-Ray in less than one minute.

Primarily, this involves identifying the best computer-aided detection (CAD) commercial products that will be able to diagnose TB that is on the market for our clients to use with high confidence and research to support the decision.

In addition, the team also plans to have an evaluation on user experience by identifying factors affecting user acceptance and the product’s impact on the clinical workflow.

A final report will analyse the various AI technologies using the qXR product from the company Qure.ai as a benchmark.

At this stage we will evaluate if qXR is the best choice, or suggest an alternative product instead. Factors we will consider are its accuracy and the identification of factors affecting user acceptance. All of these must be taken into consideration by the client when deciding which AI technology to implement.

The key stakeholders are:

* Information System researchers, specifically our client, Simon Poon
* Medical health care professionals who do the screenings for TB, as they are the ones who will be using the technology.
* Patients who are screened using the technology.

# System specification

# To give a brief overview of the system specification I’ll just run through some functional requirements, we need to:

# - Prove AI is technically viable

# - Benchmark qXR against other AI products

# - Identify accuracy

# - Identify efficiency

# Non functional requirements include

# - covering all requirements from client

# - Meet time schedule

# - Use reliable research

# Tech constraints

# - Availability of articles

# - Validity

# - Usability of dashboard

# User story example:

As a radiologist, I want to ensure that qXR is the most appropriate AI tool for my needs so that I can treat my patients accordingly.

* Acceptance Criteria: qXR successfully benchmarked against competitive AI tools that are certified

We defined this as functional

# Quality of work

While each of our user stories have a different acceptance criteria, they share a common theme, a high-quality, detailed product evaluation that they can rely on. Based on this we developed a testing plan which has six areas.

1. First is article quality. To ensure this, we search for articles in academic databases, only use journal articles, and exclude articles that are funded by the company that makes the product due to the risk of bias.

2. For systematic review quality, we are following the PRISMA statement which is designed to improve the transparency and reporting of systematic reviews.

3. Third is meta-analysis. A second person will check the data extracted into a spreadsheet to reduce the risk of data entry errors. Also for the UX side, we will be using the Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology model as the basis for the factors we identify.

4. To ensure that our evaluation is of good quality, we are following the ÉCLAIR guidelines which are a practical set of questions created to evaluate commercial Ais in radiology.

5. For report quality, we will follow the structure approved by the client, use academic writing, and consistent referencing.

6. Finally is our presentation quality, where we will select the most important report content to present alongside informative and professional slides, and a demonstration to visually present the report findings.

A limitation of our testing plan is that due to the nature of the project, most of the testing criteria developed is qualitative, which means that it will likely have some subjectivity.

**Discipline**

In this project, we used a lot of data analysis and statistics knowledge. Because of time, I will introduce the two most important terms that may confuse you guys, "AUC" and "PRISMA" .

AUC is the projection of the ROC curve to the x-axis. The two values that constitute ROC are true positive and false positive ratios, so the area of AUC is positively correlated with AI performance.

PRISMA is an evidence-based minimum set of items for reporting in systematic reviews and meta-analyses. In this project, we also used prisma. Using prisma, we can focus more on the problem we want to explore. With the drop of low-relevance articles again and again, the conclusion we draw from mate-analysis will be more Approaching our problem.

# Plan for remaining requirements/user stories, and potential project risks

# Group Processes

First of all, our group structure is demonstrated by the slide. Based on the differences between our project objectives and traditional software development projects and after consulting with our clients, we believe that the team structure from extreme programming is not applicable. Instead, we split our team into two sub-teams, the technical team, and the user experience team. To ensure a proper management framework, Cailin is chosen to be the team manager for the whole team, she is in charge of task distribution, client communication, and other team internal matters. Kiran and Bronte are the team leader for the tech team and user experience team separately.

For each week, we will have two client meetings and one group meeting to make sure we are on the right track. Each of the team members is having a positive attitude toward taking responsibility and tasks, we are all dedicated to this project.

The support from our clients is huge, our clients introduced us to the systematic review and meta-analysis and supplied multiple articles to coach us through the PRISMA process.

There are certain limitations due to the unorthodox team structure and the change of project scope earlier. Before week5, we found that even though the two sub-teams were functioning well, the lack of communication between the two teams had a negative impact on our productivity. After we found this issue, we immediately enhanced communication between the two teams, and now it is our belief that the team is on the right track.

# Demo

Due to the nature of our project, since we don’t need any programming task, we will demonstrate our prototype through some analysis. In this chart, we compare the sensitivity and specificity of 6 certified products for detecting tuberculosis. The sensitivity in medical imaging refers to the test to correctly identify patients with a disease and specificity refers to the test to correctly identify patients without the disease. Due to the WHO TPP principle, the products with sensitivity greater than 0.9 and specificity greater than 0.7 are considered good for detecting tuberculosis. In this case, we can find qXR and CAD4TB satisfy the minimal requirements of TPP principle. And Insight cxr almost reaches the minimal requirements. By contrast, chesteye and inferread dr performed relatively bad in detecting tuberculosis. In addition, VUNO reaches an extremely high specificity score and also reaches a good sensitivity score. However, there are few articles related to VUNO in detecting tuberculosis, so the data for VUNO may be biased which needs further research. In conclusion, qXR and CAD4TB performed well in detecting TB. Insight cxr can be an alternative to detect TB. Chesteye and inferread dr will probably be left out. And VUNO needs further deep research. That’s all for the prototype demo.