NEW CHANGES

After consulting Dr Petrina, to reduce scope of work to help the group focus on the quality of work, we suggest you focus on one certified AI product “qXR”. The goal is to benchmark qXR against all other competing AI for TB diagnosis. Group 1 - DTA evaluation group would need to look for any evidence (develop, clinical studies) related to qXR and also any comparison studies qXR test against in the past (before and after CE certification), I believe it was 2018.

Project Background:

We are consultants hired by a health services company. Their radiologists regularly screen for and diagnose TB by performing X-rays and interpreting them. The company has heard of an AI technology called qXR which analyses x-rays to diagnose tuberculosis. They are considering introducing this technology into their clinical workflow with goals to improve accuracy, provide a second opinion and increase efficiency. We have been hired to investigate whether this technology is worth investing in, with a focus on its DTA compared to all certified competitors on the market and its user experience, specifically how it will impact the clinical workflow of the radiologists.

Project Output:

We are using the ECLAIR guidelines as a structure for our investigation and report. These are leading guidelines for assessing commercial AI solutions in radiology. It has 5 sections, and while we will cover all of them, our main focus is 2 (performance and validation) and 3 (usability and integration).

Project Report Division of tasks

1. Executive Summary and introduction (Cailin)
2. Overview of system from user perspective, appendix 1, add file to bitbucket (user stories) (Kiran)
3. Evaluation (Bronte)
4. System Structure (evaluation/final report structure)(zheyuan)
5. Tools to build the system (Zichen)
6. Information search/research and discipline knowledge use and application (Benjamin+Andersen)
7. Group processes, reflections and conclusions (Benjamin+Andersen)

**ALL to complete individual contribution section**

**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)

Task Priorities:

1. Project status report **due Monday**
2. Technical/ UX systematic review **search queries complete and at least some articles identified**
3. Additional ECLAIR guidelines task

## Relevance (Technical team)

* Complete by all
* Develop a deeper understanding of qXR

1.1 What problem is the application intended to solve, and who is the application designed for?

* Define the scope of application
* End-users
* research vs. clinical use
* usage as double reader, triage, other
* outputs (diagnosis, prognosis, quantitative data, other), indications and contra-indications

1.2 .What are the potential benefits, and for whom?

* Consider benefits for patients, radiologists/referring clinicians, institut

1.3. What are the risks associated with the use of the AI system?

* Consider risks of misdiagnosis (including legal costs), of negative impact on workflow, of negative impact on quality of training

## Performance and validation

* Complete by technical team

Questions to briefly cover:

* Are the algorithm’s design specifications clear?
* How was the algorithm trained?
* How has performance been evaluated?
* Have the developers identified and accounted for potential sources of bias in their algorithm?
* Is the algorithm fixed or adapting as new data comes in?

Systematic Review and meta-analysis:

* Identify all certified, on the market competitors.
* Use PRISMA to systematically find comparison DTA studies (make note if a comparison between qXR and a competitor is not found).
* Extract data from the selected studies
* Analyse the data
* Perform meta-analysis
* Discussion and conclusion of comparison of qXR to its competitors in regard to DTA.

## Usability and integration

* Complete by UX team

Questions to cover from the systematic review and meta-analysis:

* How can the application be integrated into your clinical workflow?
* How exactly does the application impact the workflow?
* What are the requirements in terms of information technology (IT) infrastructure?
* Interoperability - How can the data be exported for research and other purposes?
* Will the data be accessible to non-radiologists (referring physicians, patients)?
* Are the AI model’s results interpretable?

Systematic review and meta-analysis:

* Use PRISMA to find qXR UX studies, then use PRISMA to find UX studies on its competitors, and then if needed use PRISMA to find UX studies on similar technologies.
* Extract data
* Perform meta-analysis
* Discuss above questions and limitations

## Regulatory and legal aspects (Bronte)

* Does the AI application comply with the local medical device regulations?
* Does the AI application comply with the data protection regulations?

## Financial and support services considerations (Cailin, Benjamin)

* What is the licensing model?
* How are user training and follow-up handled?
* How is the maintenance of the product ensured?
* How will potential malfunctions or erroneous results be handled?

## Recommendation and Conclusion

* Whether we recommend they purchase the product with justification.
* Suggest next steps, e.g., trial

Technical Team

1. Select products for benchmarking against qXR

* Identify all other certified products on the website: <https://www.ai4hlth.org/>
* Establish criteria for inclusion/exclusion
* Include/exclude based on the above criteria
* Present options to Simon
* Final selection

1. Systematic review (PRISMA) on diagnostic test accuracy (DTA) of each product

* Formulate the review question
* Define inclusion and exclusion criteria
* Develop search strategy and locate studies
* Select studies
* Extract data
* Assess study quality
* Analyse and interpret results
* Disseminate findings

1. Compare DTA of each CAD product against qXR

* Search for comparison studies
* Before and after certification

NEW CHANGES

The scope for Group 2 – UX evaluation group remains the same. First look for any user studies on qXR, if not enough then expand to other similar AI products competing with qXR, and perhaps expand to other studies using any AI using x-ray for non TB diagnosis.

User experience team

1. Select articles about user experience.
   1. qXR
   2. Other Commercial AIs
   3. (If not enough) find all the tools for TB detection.
   4. (If not enough), find tools for detecting all lung diseases.
   5. (If not enough), find tools generally for medical imaging.
2. Systematic Review (PRISMA)

* Create search query
* Identify which databases to search
* Exclude based on title
* Exclude based on Abstract
* Exclude based on full text

1. Compare different models used for evaluating user experience.

* How do they modify TAM differently?
* The correlation between different constructions.

1. Identify factors affecting user acceptance and evaluate user experience
2. Form a precise conclusion.
3. Data visualisation
4. Prepare for presentation.