**P56: Benchmark Commercial AI product for Medical Imaging Services Provider**

**Client: Simon Poon and Neysa Petrina**

**ISYS3888\_TU15\_02\_Group2**

**Tutor: Yoki**

**Group members details**

* Bronte Lee 490420056
* Cailin Feldman 480439062
* Kiran Papalkar 490411205
* Zichen Li 500174041 (Second group: SOFT3888\_W16\_02, project: P25 - Travel plan and diary application)
* Jingbang Men 500401903
* Zheyuan Zhang 500023392
* Baoxi Liu 480022439 (Second group: SOFT3888\_TU12\_01, project: P08 - WeBike)

**Submission Date**

**24th October 2022**



**School of Information Technologies**

**Faculty of Engineering & IT**

**ASSIGNMENT/PROJECT COVER SHEET - GROUP ASSESSMENT**

**Unit of Study: ISYS3888**

**Assignment name: Final Group Report**

**Tutorial time: Tuesday 15:00 Tutor name: Yoki**

**DECLARATION**

**We the undersigned declare that we have read and understood the *University of Sydney Student Plagiarism: Coursework Policy and Procedure*, and except where specifically acknowledged, the work contained in this assignment/project is our own work, and has not been copied from other sources or been previously submitted for award or assessment.**

**We understand that failure to comply with the *Student Plagiarism: Coursework Policy and Procedure* can lead to severe penalties as outlined under Chapter 8 of the *University of Sydney By-Law 1999* (as amended). These penalties may be imposed in cases where any significant portion of my submitted work has been copied without proper acknowledgement from other sources, including published works, the internet, existing programs, the work of other students, or work previously submitted for other awards or assessments.**

**We realise that we may be asked to identify those portions of the work contributed by each of us and required to demonstrate our individual knowledge of the relevant material by answering oral questions or by undertaking supplementary work, either written or in the laboratory, in order to arrive at the final assessment mark.**

| **Project team members** | | | | |
| --- | --- | --- | --- | --- |
| **Student name** | **Student ID** | **Participated** | **Agree to share** | **Signature** |
| **Bronte Lee** | **490420056** | **Yes** | **Yes** | **Bronte Lee** |
| **Zichen Li** | **500174041** | **Yes** | **Yes** | **Zichen Li** |
| **Cailin Feldman** | **480439062** | **Yes** | **Yes** | **Cailin** |
| **Jingbang Men** | **500401903** | **Yes** | **Yes** | **Jingbang Men** |
| **Kiran Papalkar** | **490411205** | **Yes** | **Yes** |  |
| **Baoxi Liu** | **480022439** | **Yes** | **Yes** |  |
| **Zheyuan Zhang** | **500023392** | **Yes** | **Yes** | **Zheyuan Zhang** |

**Level 2, SIT Building, J12 T +61 2 9351 3423 ABN 15 211 513 464**

**The University of Sydney F +61 2 9351 3838 CRICOS 00026A**

**NSW 2006 Australia E** [**sit.info@sydney.edu.au**](mailto:sit.info@sydney.edu.au)

**Table of Contents**

[1 Introduction](#_heading=h.njgp2uyxxd67)

[1.1 Background](#_heading=h.i24zkn27r802)

[1.2 Purpose](#_heading=h.bcazs3e5waw6)

[1.3 Project Vision](#_heading=h.mu0na23ej7lh)

[1.4 Project Goals](#_heading=h.eiv1as7v4o5v)

[1.5 Project Objectives](#_heading=h.op3x73c40dyi)

[1.6 Project Achievements](#_heading=h.50jwk6j28sev)

[1.6.1 Short-Term Deliverables](#_heading=h.u6wb5hvmqq59)

[1.6.2 Long-Time Expectations](#_heading=h.wtdbc9ekzgul)

[1.7 Key stakeholders and interactions](#_heading=h.ccyn05im05ey)

[1.8 Resources and Risks Involved in the Project](#_heading=h.2kqqyasc3jei)

[1.8.1 Resources](#_heading=h.y50pw22wmu5c)

[`1.8.2 Risks](#_heading=h.sjlszip1i5n4)

[2 System Specification and Architecture Design](#_heading=h.jatox6ydbaq9)

[2.1 User Stories](#_heading=h.leghe36c926s)

[2.2 User Story Demonstration](#_heading=h.buc1hv66bgqb)

[2.3 Technical Constraints](#_heading=h.r8milgbunfdp)

[2.3.1 Consistency](#_heading=h.q8fnvn7dgr9a)

[2.3.2 Bias](#_heading=h.18qo5tf75ah0)

[2.3.3 Data Availability](#_heading=h.qgobdks8bvag)

[2.4 Key Changes to Scope](#_heading=h.9zlr6ux5effw)

[2.5 System Architecture and Design](#_heading=h.w4snzmgjjugt)

[2.5.1 System Architecture](#_heading=h.y4i10cxzl9cs)

[2.5.2 System Design and Implementation](#_heading=h.8u56m4fxevf3)

[2.5.3 Components of the System and their Interactions](#_heading=h.of48x6x8h5v5)

[2.5.4 Design Patterns and Off-the-Shelf Frameworks](#_heading=h.p5mtiskp682j)

[3 Evaluation](#_heading=h.53hv1yy1skhx)

[3.1 Testing Plan](#_heading=h.au57fw96fg0q)

[3.1.1 Testing Types](#_heading=h.dw1svc1khty4)

[3.1.2 Test Coverage](#_heading=h.4ebmailf0sfw)

[3.1.3 Significance and Limitations of Tests](#_heading=h.jhkaviogo1m5)

[3.2 Quality Assurance Testing Criteria](#_heading=h.bmwqx9bbv7v8)

[3.2.1 Article Quality](#_heading=h.yrhcdinoz7pd)

[3.2.2 Systematic Review Quality](#_heading=h.y74p444vleeb)

[3.2.3 Meta-Analysis Quality](#_heading=h.ivx88lixg4sv)

[3.2.4 Thematic Analysis Quality](#_heading=h.wnwjxzrz3r85)

[3.2.5 Evaluation Quality](#_heading=h.8sdn87xy2v22)

[3.2.6 Report Quality](#_heading=h.1mwrysvc2x1y)

[3.2.7 Presentation Quality](#_heading=h.ihhqocvwsfvu)

[3.3. Quality and Acceptance Tests](#_heading=h.f155rw241bdu)

[3.3.1 Article Testing](#_heading=h.ptlhtiei4us3)

[3.3.2 Systematic Review Testing](#_heading=h.af133t72wtek)

[3.3.3 Meta-Analysis Testing](#_heading=h.1ryyciz8fnkm)

[3.3.4 Thematic Analysis Testing](#_heading=h.jbi5cryl5t46)

[3.3.5 Evaluation Testing](#_heading=h.gvbw6o11qivv)

[3.3.6 Report Testing](#_heading=h.demruul1ddi3)

[3.3.7 Presentation Testing](#_heading=h.qqq30u1pn4hz)

[3.4 Quality Constraints](#_heading=h.6nb8n9fv1aaf)

[3.4.1 Technical Knowledge](#_heading=h.yasv6lsci3dq)

[3.4.2 Client Expectations](#_heading=h.qtspnraqfzl)

[3.4.3 Time](#_heading=h.rqvxshdd5ii4)

[4 Discipline knowledge and tools](#_heading=h.q2t3kltbdp32)

[4.1 Research discipline](#_heading=h.utl389lmi8sl)

[4.1.1 Academic background](#_heading=h.eb7nza697z4b)

[4.1.2 PRISMA](#_heading=h.9v5slt6bgy4y)

[4.1.3 Thematic analysis](#_heading=h.toavrqg6srs3)

[4.1.4 The ECLAIR guidelines](#_heading=h.l2yzl9seup3u)

[4.1.5 Diagnostic test accuracy](#_heading=h.ae70k9ey4e76)

[4.1.6 Meta-analysis](#_heading=h.b5cu50tqzf1g)

[4.2 Tools](#_heading=h.x2gnqd9pf58q)

[4.2.1 Team communication](#_heading=h.8sp7i0et72xb)

[4.2.2 File management](#_heading=h.sjxgz3sko8g9)

[4.2.3 Data analysis](#_heading=h.ja0rvwg7lep5)

[4.2.4 Article Collection](#_heading=h.mo9twgz8pbry)

[5 Group Processes](#_heading=h.jtg1kyb1wxtq)

[5.1 Application setup](#_heading=h.hf88cexa9xde)

[5.2 XP Roles](#_heading=h.actw6jt79hsn)

[5.3 Team Structure](#_heading=h.p5fvd1nhsp3w)

[5.4 Tech Team](#_heading=h.g43dtorvfk9g)

[5.5 UX Team](#_heading=h.5m6hx9k5hufb)

[5.6 Team Communication](#_heading=h.63mgj52woe0k)

[5.7 Client Communication](#_heading=h.c5eo4r5qgp3p)

[5.8 Issues and Solutions](#_heading=h.ks3ecj966orl)

[6 Reflections and Conclusions](#_heading=h.sd705oembn5k)

[6.1 Challenge and Risk Analysis](#_heading=h.pzvsbydpzd20)

[6.2 Limitations](#_heading=h.hlqoi44v29f2)

[6.3 Primary Strengths](#_heading=h.7344ubevqdxl)

[6.4 Programming Practices](#_heading=h.pc5viraq1r8f)

[7 Individual Contributions](#_heading=h.ic267ab2e7lb)

[Appendix](#_heading=h.lgn0bycjed68)

[A1. Location of Meeting Minutes and Client Communication](#_heading=h.1wt5p02xsv6k)

[A2. Involved articles](#_heading=h.5kj89etlmk3n)

[A3. Systematic Review Protocols](#_heading=h.u0a4mrs2x1dz)

[A4. Academic Websites](#_heading=h.nlhddo7fu3n3)

[A5. Meta Analysis Tools](#_heading=h.swo3zj32jqfd)

[A6. Forest Plots](#_heading=h.t0o5keckbtzi)

[A7. SROC Graph](#_heading=h.erijf5j3bh6c)

[References](#_heading=h.z18d94b2ebco)

# 

# 1 Introduction

## 1.1 Background

Tuberculosis (TB) approximately infects 10 million people a year and kills 1.5 million. If TB is diagnosed and treated early, most people are cured within six months of diagnosis. The WHO End TB Strategy aims to reduce TB deaths by 90% and to cut new cases by 80% between 2015 and 2030. To achieve this goal, missing cases must be identified and all cases diagnosed earlier. This can be achieved through more widespread and early screening of people at greater epidemiological risk than the general population (who are assumed to be disease-free). The first line of prevention is detecting people who may be infected and who require further investigation and treatment. The major techniques for screening include screening for symptoms (such as persistent cough and fever) by clinicians, low-cost rapid molecular assays (mWRD) and examination for radiological abnormalities indicative of TB. The latter has been performed mostly by expert human readers, such as radiographers, radiologists, and trained health care practitioners. However, an emerging technology is computer-aided detection (CAD) commercial products which diagnose TB using artificial intelligence algorithms.

At the present moment, there are nine certified computer-aided detection (CAD) products that may be used for TB detection. The ongoing development of these products reveals encouraging results and they have begun being integrated into the clinical environment. These products promote benefits such as workflow optimisation, supporting radiologists as a second reader, and improving patient care. However, two issues arise, first, do these products achieve their promoted benefits, and second, which product is most fit for purpose?

## 1.2 Purpose

The purpose of this project is to evaluate the commercial AI product qXR and provide recommendations for the client to inform their decision whether or not to invest in qXR. The client is a representative of the health services company ABC, who is intending to integrate an AI into their radiology workflow and is considering qXR. The client is seeking more information about the product qXR, particularly its diagnostic test accuracy compared to its competitors and user experience, so they know whether to invest in qXR, why, and any important considerations when investing.

## 1.3 Project Vision

We have been engaged to evaluate qXR and recommend whether to invest in it. Our vision for the project is that we provide a comprehensive evaluation of qXR and use that evaluation to inform a recommendation of whether or not to invest and explain what the client needs to consider.

## 1.4 Project Goals

The overall project goal is to create a report that evaluates qXR and recommends whether to invest in it. It will include analysis of 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. We will also use the ECLAIR guidelines to evaluate qXR from a range of angles, including legality, finance, risks and benefits, and its infrastructure.

## 

## 1.5 Project Objectives

ABC company is seeking more information about the product qXR, particularly its diagnostic test accuracy and user experience compared to competitor products. Specific objectives are:

* A systematic review of the literature providing evidence of diagnostic test accuracy and user experience.
* A meta-analysis to compare and benchmark qXR against commercial AIs used for medical imaging in TB diagnoses.
* Conducting a thematic analysis to identify factors affecting user acceptance of AIs used for medical imaging.
* Make evidence-based recommendations on a specific product solution for ABC company
* To ensure the quality of the content, use guidelines including PRISMA (for our source finding) and ÉCLAIR (for the structure of our evaluation) to ensure high-quality systematic reviews and evaluation.

## 1.6 Project Achievements

### 1.6.1 Short-Term Deliverables

A final report was made which analysed the various AI technologies using the qXR product from the company Qure.ai as a benchmark and with a conclusion of our product suggestion based on their accuracy as well as highlighting factors that can impact the clinical workflow once the CAD is decided on.

We met frequently as a group and divided into sub-groups technology and user experience. From there we followed the systematic review protocol alongside our final report to present our research and findings.

### 1.6.2 **Long-Time Expectations**

Our project goals are on track to being achieved. Using qXR as a benchmark and undertaking in-depth analysis of a majority of the CAD products, clear findings were made. We identify the most technically accurate product which will allow ABC company to select a product with confidence. The report also highlights aspects of the integration process that should be focused on when implementing, to limit the impact on workflow

## 1.7 Key stakeholders and interactions

The key stakeholders are:

* Information System researchers, specifically our client, Simon Poon, as they conduct research in search of the highest ranking technology, investigating what are the most impactful features that are used to rank the various technologies.
* ABC Company. as they are a nation-wide medical imaging services provider offering a wider range of medical imaging services who want to know what the best artificial intelligence (AI) product is currently available to produce promising results in image processing and image analysis for medical diagnoses, ABC would like to prepare for integrating these technological advances to enhance and support radiologists in optimizing their workflow efficiency and improve patient care.
* Medical health care professionals as they are the people who do the screenings for TB and will be using the CAD technology so they want to ensure they are using an accurate product that will be extremely easy to integrate and use.
* Program directors who authorise the use of technology in the radiology practice as they want to ensure their purchase will be creating efficient and effective work practices.
* Software and technical managers who ensure the technology they will be in charge of operates as expected and has minimal maintenance requirements
* Patients who are being screened with the CAD technology. Patients are important to keep in mind as their refusal with the technology will result in the purchase in the CAD product to be unused which will be a waste in clinical finances and meaning there will therefore be no improvements in the workflow

The project’s report will be delivered to Neysa Petrina as a representative of ABC Company and Simon Poon as an information system researcher. The decision made by ABC Company will directly impact on the other stakeholders. For medical healthcare professionals, program directors, and software and technical managers, if ABC chooses to adopt qXR, they would have to learn how to use it and feel confident that qXR is accurate and safe to use. This can be assisted by ABC Company providing their employees with a copy of the report. For patients, it would be intended that adopting qXR would have a positive impact by earlier detection of disease and shorter wait times. However, if qXR is not well-integrated or not as accurate as expected, it could have the opposite impact.

## 1.8 Resources and Risks Involved in the Project

### 1.8.1 Resources

Human Resources

Our team consists of seven members who are completing ISYS3888 this semester. This subject is the capstone for the information systems major, therefore, all members have completed a range of information systems subjects giving them considerable knowledge in this area. The academic background of the team is further elaborated on in [section 4.1.1](#_heading=h.eb7nza697z4b).

Time Resources

The project started on the 9th August 2022 and finishes on the 26th October 2022. This means that the total time to complete the project is eleven weeks and one day. It is expected that each team member would spend 10 hours every week working on the project.

Material Resources

The resources searched to collect articles for the systematic reviews are listed as such:

1. PubMed. A highly rated database from the National Library of Medicine and the National Center for Biotechnology Information of the United States of America. PubMed is a well-known database that contains over 34 million citations for biomedical literature from multiple medical journals. It includes Medline which is a premier database for medicine and health services. - Ben
2. ProQuest. ProQuest is also a platform built specifically for researchers. The platform holds a great academic reputation among the scientific community, 98% of the top 400 universities rely on ProQuest when it comes to research.
3. Embase. A premier biomedical and health care database that is European-oriented. This database was recommended by the University Library website when searching for medicine and health related databases.
4. ScienceDirect. A science database which includes many scientific areas and specifically has good coverage of health sciences and medical research. This was chosen because it has a broader view to capture studies that may not have been included in medicine-specific journals.
5. ‘The Lancet’. ‘The Lancet’ is a medical journal dedicated to publishing cutting-edge research regarding medical science. All the articles published in that journal were carefully reviewed and read by doctors and scientists.
6. The client of this review. Simon is a representative of the client of this research. In order to guide the review team which consisted of first-time participants to the systematic review, Simon provided around a dozen of valuable articles for the team to review.
7. Qure.ai official website. Qury.ai is the company of the targeted product, qXR. Though the evidence provided by the product’s official website could be potentially biased, the website still can be a valid source when it comes to gathering evidence and user cases after close examination.

For more information on our systematic review process please refer to our appendix A2, A3, A4 which will lead you to the documentation with our systematic review protocols, selected articles and final data collection.

### 1.8.2 Risks

This project faces a range of risks:

* A risk identified by the client was that our UX evaluation is unlikely to be specific to qXR. In this situation, the research is pooled data of multiple products and the report makes the assumption that the UX experience of qXR is similar. However, there is a risk that qXR’s user experience is different to the AI’s we study.
* Insufficient evidence to benchmark the diagnostic test accuracy of all of qXR’s competitors. For instance, a lack of articles or articles that cannot be used because they lack the appropriate data or are written in a language other than English.
* Poor communication of information. Good communication is essential for any project. However with the difficulty of hybrid meetings, confusion of the project it can lead to team member’s and client’s to remain silent instead of speaking and asking for clarification on project scopes, addressing issues faced etc.
* Differing time zones of members. The different time zones lead to the risk of not having to spend adjusting and rescheduling meetings and delay in response to urgent messages sent on slack
* Scope changes from our client. This could lead to confusion with our team as there would be multiple versions of the scope recorded and may require restructuring how the project is completed.
* Limited time with the research process and lack of support from academic staff (due to us not being postgraduate/honours students). The zero support and small time can result in having potential data quality issues with data collection for our final report as our team is unable to dedicate time to checking the values before analysis due to the tight deadline provided.

# 2 System Specification and Architecture Design

## 2.1 User Stories

Due to the nature of this project as a benchmarking report the user stories are limited to what the client and radiologists require from the report, that is, the knowledge they expect to gain through seeking consultants to benchmark these products as well as their general requirements for quality of the report The user stories from this project are also limited because they can only be “completed” when the report itself is completed. As the user stories were informed by the sections of the report as requested by the client, none were excluded and all acceptance criteria were met in the final report.

| # | User Story | Acceptance Criteria | Requirement | Status |
| --- | --- | --- | --- | --- |
| 1 | As a medical imaging provider, I want a comprehensive report so that I can justify my use of qXR. | Report covers all topics that client has requested | Non- Functional | Done |
| 2 | As a radiologist, I want to have an accurate AI product to detect TB so that I can give the correct diagnosis. | Most accurate AI product identified in comparison to others through analysis of AUC | Functional | Incomplete |
| 3 | As a radiologist, I want to have an efficient AI product to detect TB so that I can identify issues and assist my patients in recovering as soon as possible. | Efficiency of qXR is elaborated on. | Functional | Done |
| 4 | 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. | qXR successfully benchmarked against competitive AI tools that are certified, usability of qXR is identified, qXR is successfully argued to be best choice for client | Functional | Done |
| 5 | As a medical imaging provider, I want a reliable report so that my decision is well informed. | Report is well researched and thorough | Non- Functional | Done |
| 6 | As a medical imaging provider, I want to have the report by my required date so that I can make decisions in a timely manner | Report is submitted by the due date | Non- Functional | Done |
| 7 | As a radiologist, I want a product that will not interrupt my workflow so that I can maintain my focus on treating patients. | Identify the impact qXR has on the workflow and whether it helps or hinders the clinical workflow of radiologists when diagnosing TB | Functional | Done |
| 8 | As a radiologist, I want to know what problem qXR is solving in my worklife so that I can better understand the product | Adequately explain the problem and define users | Functional | Done |
| 9 | As a medical imaging provider, I want to have a knowledge of what qXR can bring us and what kind of patients it aims for. | Show that qXR’s benefits and its targeted user | Non- Functional | Done |
| 10 | As a medical imaging provider, I want to know if there are any risks that may occur in using qXR so that I can be aware of and mitigate them | Any risks present in the system are identified | Functional | Done |
| 11 | As a medical imaging provider, I want the performance and validation of qXR to be wholly covered in the report so that I can evaluate its benefit to me | Performance and validation is completed | Functional | Done |
| 12 | As a radiologist I want the specifications of the algorithm to be clear so that I can better understand how it functions | Clearly state design specifications of qXR | Functional | Done |
| 13 | As a radiologist, I want to know how qXR’s algorithm be trained so that I can better understand how it works | Clearly explain the algorithm trained | Functional | Done |
| 14 | As a medical imaging provider, I require knowledge of how qXR’s performance has been evaluated so that I can ensure it is the best option | An evaluation for qXR’s performance | Functional | Done |
| 15 | As a medical imaging provider, I want to ensure that the selected product is unbiased so that I can accept it as a valid product | Potential biases of qXR are assessed | Functional | Done |
| 16 | As a medical imaging provider, I need to know whether the algorithm is adapted to new data so that I can long term use this algorithm | Find whether the algorithm is fixed or adapting to new data | Functional | Done |
| 17 | As a radiologist, I want to ensure that qXR can be easily integrated into my workflow so that I am not unnecessarily hindered by using it | State the integration process | Functional | Done |
| 18 | As an owner of a medical imaging practice, i want to know what are the IT requirements to implement qXR | Identify IT requirements | Functional | Done |
| 19 | As a medical imaging provider, i want to know what the data output by qXR be used for | Identify what data is being used | Functional | Done |
| 20 | As a medical imaging provider, I want to know if the output of qXR will be given to other people instead of radiologists. | Identify who can get access to the data | Functional | Done |
| 21 | As a radiologist I want to know if the result from qXR is interpretable or not. | Identify the clarity and user feedback on qXR’s results and is there any difficulty for interpreting the result generated from qXR. | Functional | Done |
| 22 | As a medical imaging provider I want to be aware of the legal status of qXR in Australia. | Identify the laws that apply to medical devices in Australia, determine whether qXR can be legally used, and if not, the pathway for qXR to become legal in Australia | Functional | Done |
| 23 | As a radiologist I want to know the licensing model for qXR. | Identify the licensing model for qXR. | Functional | Done |
| 24 | As a medical imaging team, I want to know what the training process involved to use qXR | training process involved to use qXR | Functional | Done |
| 25 | As a radiologist I want to know how much maintenance is required once qXR has been implemented into the imaging system because I need to know how much resources it will require | Identify maintenance costs and frequency | Functional | Done |
| 26 | As a radiologist I want to know how the potential malfunctions and errors from qXR can be resolved because I need to know that any issue that arises will have instructions to solve the error | Potential malfunctions and errors identified | Functional | Done |

## 2.2 User Story Demonstration

Each user story has been demonstrated in the final report and below is a comprehensive list with pointers to the relevant section that addresses the user story, the final report can be found on bitbucket using this [link](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report%20Documents/User%20Stories.pdf)

1. This is demonstrated through the covering of ECLAIR guidelines as requested by the client in the report, further to this all requests by the client have been thoroughly researched and answered.
2. This user story was widely neglected in favour of identifying sensitivity and specificity and paying less attention to accuracy as defined by AUC, data was still extracted for each article, however, comparing specificity and sensitivity of the products was more comprehensive to address our client’s needs. This was the only user story that was incomplete.
3. Section 2.2 elaborated on the efficiency of qXR finding it takes less than one minute to diagnose.
4. Section 7 of the report has our final conclusion in which we suggest that qXR is the most effective tool out of the certified AI products to meet the client’s needs, citing our research into both the DTA as well as usability and workflow integration.
5. By following the PRISMA process we have ensured that the research is well done and thorough, evidence of this can be found on the bitbucket at [link](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Technical%20Documents/Technical%20Systematic%20Review.pdf) and [link](https://docs.google.com/document/u/0/d/1fvetkO7wzLRR7s-9ukJXNDwvOyJ5Wf3UZVyRBuy1Jz0/edit)
6. This user story will be demonstrated on Tuesday the 25th when the final product is submitted to the client, currently it is the only user story in progress.
7. Section 4.2 of the report outlines the systematic review and thematic analysis in order to determine the product’s impact on workflow.
8. Section 2.1 of the report addresses this user story by defining the problem faced by ABC and identifying the users of the product.
9. Section 2.2 of the report identifies the benefits of qXR for ABC
10. Section 2.3 of the report finds the risks of qXR, this was largely general risks with AI products as no specific risk analyses of qXR could be found. We identified the risk of automation bias in that radiologists may misdiagnose due to a fault in the AI as well as cybersecurity risks.
11. Section 3 of the report indicated the performance of qXR through the systematic review and meta-analysis and further validated the product by investigating its training and design specifications as well as potential biases and their mitigation.
12. Section 3.1 of the report identified the design specifications of qXR
13. Section 3.2 of the report went in depth on how qXR was trained via deep neural networks and identified what symptoms it was found to identify
14. Section 3.3 of the report identified through meta-analysis and systematic review that qXR performed significantly better than most alternatives other than CAD4TB which it only slightly outperformed. Graphs that were made in this stage can be found in the appendix, forest plots are located in [A6](#_heading=h.t0o5keckbtzi) and the SROC graph is found in [A7](#_heading=h.erijf5j3bh6c), these display how we came to our conclusions. The forest plots assisted in identifying consistency of the studies and our SROC graph gave a more visual representation of performance.
15. Section 3.4 of the report identified that qXR mitigated selection bias by using algorithmically generated datasets as well CT scans, it also identified an unmentioned bias of underdiagnosis for communities who do not have significant data.
16. Section 3.5 of the report informs the user that the algorithm does indeed adapt as new data comes in.
17. Section 4.1 of the report states both the process of integration into workflow as well as the options that a company has in terms of integration
18. Section 4.3 of the report details the communicated information from Qure, the provider of qXR, that the best option for integration is the use of Qure’s cloud-based hosting system. It also found it can be deployed on premise, however this requires further hardware and systems.
19. Section 4.4 of the report addressed this user story by investigating interoperability of the system.
20. Section 4.5 of the report identifies who will have access to the information that is accessed by the system.
21. Section 4.6 of the report looked at whether the results would be interpretable by human readers and distinguished any difficulties that would arise.
22. Section 5.1 and 5.2 of the report define the laws which apply to the qXR and whether qXR is able to be used as a valid product.
23. Section 6.1 of the report identifies the licensing model of qXR.
24. Section 6.2 of the report further elaborates on the user training of the system.
25. Section 6.3 of the report identifies the needs of the system in terms of maintenance and ensures that the product is of peak quality.
26. Section 6.4 of the report identifies the potential malfunctions that may arise and the handling of these errors.

## 2.3 Technical Constraints

### 2.3.1 Consistency

It was difficult to attain consistent results for each product in the comparison of DTA, this was because the articles were collected from a range of years, thus the products underwent many updates and changes, further to this, it was found that the accuracy of these products is weaker when the patient has had a history of being treated for TB, this means that as time goes on accuracy will decrease as it becomes easier to diagnose and treat people for TB. This constraint was addressed through tests of heterogeneity to ensure that the articles should be accepted and used together in a meta-analysis.

### 2.3.2 Bias

Due to the relatively specific nature of our topic of research, appropriate articles to analyse were limited with many identified articles coming from the company that owns the AI product. We found these articles to have inconsistent data that was incongruent with other findings, after identifying that these articles held bias towards the product the owner had a vested interest in, we decided to exclude these articles in favour of those outputted by less biased groups. While this did weaken the breadth of our analysis as some products could not be benchmarked due to a lack of data we believe it allows our results to be more trusted and valid for the client.

### 2.3.3 Data Availability

Due to the relatively novel aspect of many of these AI products, we found it particularly difficult to find peer-reviewed articles with relevant data, the UX team had significant challenges in this area after their initial systematic review for qXR’s usability yielded no appropriate articles, in this case the constraint was alleviated by broadening the search to include a general analysis of usability of AI diagnosis tools. The technical team faced challenges of locating articles for many certified AI products with some having no research into DTA at all, due to this they were forced to narrow their scope from the original 8 products to benchmark into 4, this meant there was less coverage of tools that the client might consider, however, it allowed the recommendation to be more concise and data-driven.

## 2.4 Key Changes to Scope

In the early stages of the project, the scope had multiple changes which are explained below.

| Date | Scope |
| --- | --- |
| 7/8/2022: Initial project scope provided to the group. | 1. Conduct a systematic review and meta-analysis to benchmark the diagnostic test accuracy of all certified AI products that diagnose tuberculosis. 2. Present your recommendations with detailed analyses based on your investigation and extensive benchmarking effort. |
| 11/8/2022: first client meeting scope | 1. Technical: conduct a systematic review and meta-analysis to benchmark the diagnostic test accuracy of all certified AI products that diagnose tuberculosis. 2. User experience: systematic review and meta-analysis to measure factors affecting User Acceptance of AI used for medical diagnosis (preferably AI imaging solutions for medical diagnosis, and even better for TB diagnosis). |
| 19/8/2022: email received from the client with a scope change | 1. Technical: focus on one certified AI product “qXR” and the goal is to benchmark qXR against all other competing AI for TB diagnosis, using systematic review and meta-analysis. 2. User experience: scope was the same, except to look for user experience studies specific to qXR as well. 3. The ÉCLAIR guidelines were included in the email. |
| 24/8/2022: client meeting to clarify new scope | 1. Context: the company has chosen a particular AI product they are interested in and have hired the group as consultants to prepare a report evaluating the product and provide recommendations for what to do next. 2. Confirmed that technical and user experience sections remain the primary focus of the project. 3. ÉCLAIR guidelines are the method of evaluation, with a focus on diagnostic test accuracy and user experience. 4. Need to provide a recommendation of whether it is worth investing in qXR. |

## 2.5 System Architecture and Design

### 2.5.1 System Architecture

The overall architecture of the project is conveyed through the structure of our report, there are seven sections in total that make up the report and each of these work in conjunction to provide a recommendation at the client’s request. The architecture is thus represented by our table of contents that informs what each section aims to answer for the client.

Below is the table of contents as it appears in the report for the client:

Executive Summary

Problem Articulation

Overview

Technical Findings

UX Findings

Recommendation

1 Introduction

1.1 Problem Statement

1.2 Purpose

1.3 Objectives

1.4 Stakeholders

2 Relevance

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

2.2 What are the potential benefits, and for whom?

2.3 What are the risks associated with the use of the AI system?

3 Performance and Validation

3.1 Are the algorithm’s design specifications clear?

3.2 How was the algorithm trained?

3.3 How has performance been evaluated?

3.4 Have the developers identified and accounted for potential sources of bias in their algorithm?

3.5 Is the algorithm fixed or adapting as new data comes in?

4 Usability and Integration

4.1 How can the product be integrated into your clinical workflow?

4.2 How exactly does the product impact the workflow?

4.3 What are the requirements in terms of IT infrastructure?

4.4 Interoperability - How can the data be exported for research and other purposes?

4.5 Will the data be accessible to non-radiologists?

4.6 Are the AI model’s results interpretable?

5 Regulatory and Legal Aspects

5.1 Does the AI application comply with the local medical device regulations?

5.2 Does the AI application comply with the data protection regulations?

6 Financial and Support Services Considerations

6.1 What is the licensing model?

6.2 How are user training and follow-up handled?

6.3 How is the maintenance of the product ensured?

6.4 How will potential malfunctions or erroneous results be handled?

7 Conclusion

7.1 Summary of Evaluation

7.2 Recommendation

The executive summary allows the client a succinct overview of what is found in the report and summarises the results and recommendation.

The introduction identifies the reason for undertaking the report as well as the key objectives of the report and stakeholders.

Relevance is a more detailed definition of qXR and displays its purpose, risks and benefits.

Performance and validation involves investigating the performance of qXR from a technical point of view, the most important part of this section is the evaluation of qXR’s performance, as this involves the systematic review and meta-analysis that was undertaken in order ot benchmark qXR against other AI products.

Usability and integration alleviates client concerns about integration into their workflow as well as key requirements for use of the product and how it is used. 4.2 is the systematic review and thematic analysis to determine the exact impact of AI assisted diagnosis on the client.

The regulatory and legal aspects section ensures that qXR is legally available to the client and further investigates compliance of the product with regulations to reassure the client that it is a sound product that is worth pursuing.

Financial and support services considerations inform the client of how to pay for the product as well as considerations about its continued use and options for using the software.

Finally, the conclusion includes a summary of the entire evaluation with core recommendations elicited from each section as well as our final recommendation to the client on whether they should employ the use of qXR within their organisation.

### 2.5.2 System Design and Implementation

The design of our report was largely informed by the ECLAIR guidelines, this was used as a framework to properly evaluate qXR as a medical imaging solution, this allowed a comprehensive report that would address all the client’s needs. Systematic review was used to gather the articles that would make up the basis for our core analysis of qXR, PRISMA was used at this stage to heighten the depth of the systematic review and allow the gathered articles to be trusted. In the analysis stage both meta-analysis and thematic analysis were used to form conclusions about the DTA of qXR as well as the impact of its integration into medical imaging workflow, this allowed for more informed conclusions and fully realised answers to our key questions.

These aspects of design were all used together in order to both make recommendations in specific areas of the report as well as cohesively form a final recommendation to the client in terms of their potential usage of qXR.

### 2.5.3 Components of the System and their Interactions

Articles

The smallest component of the project is the individual articles collected through our systematic review and through searches relating to other sections of the project. The articles form the evidence for answering the ECLAIR questions and are collected by the systematic review.

Systematic Review

Two systematic reviews were undertaken in this project for the purpose of systematically collecting articles relevant to the two primary focus questions, how has performance been evaluated and how exactly does the product impact on the workflow. A systematic review protocol was created to explain the entire systematic review process, including how the articles were being collected and an inclusion criteria to determine which specific articles can be included for analysis to synthesise the data.

Data Synthesis

From the included articles identified from the systematic review, their data was extracted and then synthesised. The technical systematic review used meta-analysis to synthesise their data and the user experience systematic review used thematic analysis. These analysis methods combine the data extracted from the multiple included articles to reveal insights that are used to answer the two ECLAIR questions that were the project’s primary focus.

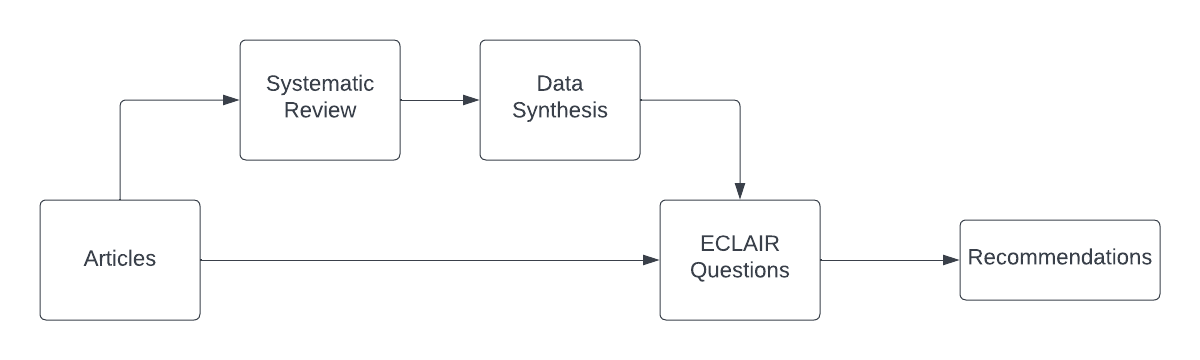
ECLAIR Questions

THe ECLAIR guidelines have 20 questions which cover a wide range of topic areas that require consideration and all were answered to provide a comprehensive evaluation of qXR. This project focused on two questions, how has performance been evaluated and how exactly does the product impact on the workflow. For these, a detailed response was provided which involved collecting articles through a systematic review and synthesising their data. For all other questions, a higher-level response was provided, although they were still thoroughly researched.

Recommendations

The overall purpose of this report is to provide a recommendation about whether or not to invest in qXR, and if so, provide recommendations about what action to take and issues to consider when investing. As the ECLAIR guidelines was the evaluation method used, the responses to the ECLAIR questions informed our recommendations.

System Components Diagram



### 2.5.4 Design Patterns and Off-the-Shelf Frameworks

Since our project does not require us to develop software, we did not use design patterns nor off-the-shelf frameworks in the software engineering sense, however, we did use established solutions for our evaluation and systematic review processes. At our client’s request, we used the Evaluating Commercial AI Solutions in Radiology (ECLAIR) guidelines as the method of evaluation. This is an established framework for evaluating AI products, such as the one we are evaluating for this project which meant that the team did not have to create their own method of evaluating qXR. The ECLAIR guidelines are elaborated on in [section 4.1.4](#_heading=h.l2yzl9seup3u). Additionally, to conduct the systematic reviews, the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) process was followed at the client’s request. PRISMA is an academically accepted standard for improving the transparency and quality of systematic reviews. Therefore, it provided a clear structure for the team to follow in writing systematic review protocols, conducting the systematic review, creating a diagram of the systematic review, and writing up the systematic review. The PRISMA process is elaborated on in [section 4.1.2](#_heading=h.9v5slt6bgy4y).

# 3 Evaluation

## 3.1 Testing Plan

The client will use the final report and presentation of this project to inform their decision about whether to move forward with the product qXR. Therefore, it is essential that the project is of a high quality so the client is confident they can rely upon the evaluation of qXR and recommendation of next steps. To ensure this, a testing plan was developed.

### 3.1.1 Testing Types

To test the project, two types of testing were used, quality assurance testing and acceptance testing. These are the most appropriate options for our project because it is an evaluation of a product and does not contain programming nor the creation of a system.

The quality assurance tests are divided into seven sections, each which covers a specific component of the project. A criteria for each section was designed at the beginning of the project with the goal being for the team to follow the criteria for each section to prevent quality issues. The project was also tested against this criteria multiple times to assess its quality. Testing was conducted by manually reviewing the project work. The complete quality assurance tests are available in [section 3.2](#_heading=h.bmwqx9bbv7v8).

Acceptance testing was undertaken to ensure that the project completed its scope. 26 user stories were created based on the project scope and an acceptance test was designed for each one to determine whether the user story was actually implemented. This was conducted at the end of the project to assess whether all of the user stories were completed. The testing was conducted manually by reviewing the project documents.

Since this project does not involve the creation of a system, therefore, there is no coding, which means a range of code testing techniques are not appropriate. We believe that the testing methods selected are sufficient for testing the project. The quality assurance tests begin with the smallest component of our project, the articles we collected, and progressively move through the project to the final report as a whole, therefore, as each component of the project is added, the project is tested to determine its quality. The acceptance tests use the user stories which were created to cover the entire project scope, therefore, assessing whether they are implemented sufficiently determines whether the project scope was fulfilled. Overall, these testing types ensure that the project scope was completed to an acceptable quality.

### 3.1.2 Test Coverage

These tests aim to achieve complete test coverage of the project. The quality tests cover all components of the project, from the individual articles to the overall report. The acceptance tests cover all of the evaluation questions, as well as the overall report.

Testing was also conducted multiple times over the course of the project to assess its ongoing quality. To check that the project achieved the project scope, broad and specific user stories were developed and used as acceptance tests.

### 3.1.3 Significance and Limitations of Tests

The testing plan has been methodically developed with a strong focus on ensuring academic rigour. As mentioned above, these tests are intended to achieve complete test coverage and assess the quality of the project from its individual components to the report. However, due to the nature of the project, much of the criteria developed is qualitative and manually tested which means that the quality of the project will likely have some level of subjectivity. To mitigate this, accepted processes, guidelines, and models have been incorporated where possible to raise the quality of the project. Additionally, a possible way to reduce the limitations of these tests would have been to employ an external person to test the report, similar to a peer-review. However, this was not able to be completed within the limited time to complete the project.

## 3.2 Quality Assurance Testing Criteria

### 

### 

### 3.2.1 Article Quality

The articles collected are the foundation of this project, they are the evidence used for evaluation and if they are of a poor quality the evaluation would be unreliable. The criteria to ensure article quality is as follows:

1. Use databases and journals to search for articles. Another acceptable source is the website of the product and government resources.
2. The systematic review articles must be limited to journal articles.
3. A company who makes a product evaluated in a study undertaken must not have funded the study undertaken in the article. Excluding these reduces bias, as companies typically make their own products look better.

The quality of the articles collected will be assessed by reviewing each article to check that it satisfies the above criteria. This process occurs during the article collection stage of the project and any articles which do not satisfy this objective criteria are excluded.

### 3.2.2 Systematic Review Quality

It is important to ensure the systematic review is completed thoroughly because the quality of it can significantly drop if an irrelevant article is included or an important article is left out. The PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) process is followed to ensure that the systematic review is of a high quality. PRISMA was designed to improve transparency and reporting of systematic reviews (Page et al., 2021) and forms the basis for the measures listed below to ensure the systematic review quality:

1. Design search queries with keywords relevant to the research question.
2. Include all results of search queries to ensure nothing is missed.
3. Define a clear inclusion criteria relevant to the question.
4. Screen the title and abstract of articles using the inclusion criteria and provide justification for the decision of each article.
5. Read the full text of the articles and use the inclusion criteria to select articles to be included and provide justification for the decision of each article.

The quality of the systematic reviews will be assessed by following the measures above and completing a PRISMA flow diagram.

### 3.2.3 Meta-Analysis Quality

Meta-analysis is the process used to synthesise the data extracted from the included studies identified from the systematic review of diagnostic test accuracy. The criteria to assess the quality of the meta-analyses is a methodical process developed with the intent of providing a transparent and reliable analysis.

1. Spreadsheet design: spreadsheets create a uniform and systematic way of recording the data extracted from the articles.
2. Data extraction: data will be extracted from the articles manually and entered into a spreadsheet. The data will be checked by a second person to minimise the risk of data errors.
3. Analysis: compare sensitivity and specificity and use ROC and AUC to visualise the comparison.
4. Transparency: all analysis conducted will be recorded and any exclusion of data will be clearly justified.

### 3.2.4 Thematic Analysis Quality

Thematic analysis is the process used to synthesise the data extracted from the included studies identified from the systematic review of user experience of AI in radiology. The criteria to assess the quality of the thematic analysis is developed with the intent of providing a logical and clearly explained analysis.

1. Spreadsheet design: spreadsheets create a uniform and systematic way of recording the data extracted from the articles. A spreadsheet will be designed to record data from user experience articles and data from all the articles will be recorded in the same spreadsheet.
2. Data extraction: data will be extracted from the articles manually and entered into a spreadsheet. The data will be checked by a second person to minimise the risk of data errors.
3. Analysis: UTAUT is to be used as the conceptual framework for creating the codebook and initial coding is to be done by members independently to avoid bias from seeing other’s perspectives. Members will then discuss to reach a consensus for codes, themes, and impact of the themes.
4. Transparency: all analysis conducted will be recorded and any exclusion of data will be clearly justified.

### 3.2.5 Evaluation Quality

The evaluation of qXR will follow the ECLAIR (Evaluating Commercial AI solutions in Radiology) guidelines because the guidelines are specific to the type of product qXR is. The guidelines are a set of questions which provide a practical guide to evaluate commercial AI solutions (Omoumi et al., 2021).

1. The ECLAIR guidelines have been followed by providing a response to all questions.
2. The client expects a high-level overview of the relevance, regulatory and legal aspects, and financial support services considerations questions. To provide this, high-quality sources will be located and an answer containing all key points and some analysis will be written for each question.
3. For performance and validation, a much deeper level of analysis is expected by the client. All questions will require a more detailed response, however, specifically for question 2.3, “How has performance been evaluated?”, a systematic review and meta-analysis is required to compare the diagnostic test accuracy of qXR to its competitors and a human to provide benchmarking.
4. The client also expects a deeper level of analysis for usability and integration. All questions will require a more detailed response, however, specifically for question 3.2, “How exactly does the application impact the workflow?”, requires a systematic review and meta-analysis to assess user experience to identify the product’s impact on the clinical workflow and the factors affecting user acceptance.

The quality of the evaluation completed will be assessed by checking that every question in the ECLAIR guidelines has been answered to the extent that the client expects and the expectations are outlined above.

### 3.2.6 Report Quality

The report is the document provided to the client and it must be of a high quality so the client can confidently rely on it to inform their decision about the product qXR. In addition to ensuring the content of the report is of a high quality, which is described above, the report itself has the following criteria:

1. Structure the report in a clear and logical manner. The structure of the final report was approved by the clients and this is contained and discussed further in section 2.5.
2. Use academic writing. This will be assessed by proofreading the document to ensure there are no spelling mistakes, grammatical errors, or informal language. Use clear and concise language, have logically structured paragraphs, and correctly use a consistent referencing style.
3. Report is provided to the client by the date determined by them.

### 3.2.7 Presentation Quality

The purpose of the presentation is to present a summary of the final report. Areas to assess quality are:

1. Content: the presentation will summarise the ECLAIR guidelines sections, with the majority of time spent on performance and validation and usability and integration. It will also have a strong focus on the outcome of the evaluation, recommendation, and limitations.
2. Slides: the slides are informative without being text heavy and have a visually pleasing and professional look.
3. Demonstration: visual aids, such as a dashboard or set of graphs and diagrams, are clear and support understanding the evaluation of qXR.
4. Presenters: all team members present in a professional manner, including projecting their voice clearly and making eye contact.
5. Q&A: presenters are knowledgeable in the sections of the report they completed so they are prepared to answer any questions the client has.
6. The presentation is presented to the client by the date determined by them.

## 3.3. Quality and Acceptance Tests

### 3.3.1 Article Testing

Quality of Systematic Reviews

| When Tested | Test Results | Improvements Made |
| --- | --- | --- |
| Week Seven: after first attempt of systematic review was complete | UX   1. Six articles were identified through Google Search which is not a valid identification method. 2. Some search results were not journal articles. 3. Successful   Technical   1. Successful 2. Some results were not journal articles. 3. Successful | 1. It was confirmed to group members that Google Search was not a valid identification method. 2. To improve article quality Google scholar was removed as an identification method because its searches are not repeatable and can mis-identify a source as academic. This would remove a significant number of articles from both systematic reviews, so it was decided to start the article collection process again. 3. Use the “journal article” or “article” feature when possible to limit search results to journal articles. |
| Week Ten: after second attempt of systematic review was complete | UX   1. Successful 2. One result was not a journal article. 3. Successful   Technical   1. Two articles were not identified through valid identification methods. 2. Successful 3. Successful | 1. The article that was not a journal article was removed. 2. For the two articles that were invalidly identified, one was re-identified through Qure.ai’s website and the other was found to be a pre-print duplicate of an already included published article. |
| Week Ten: after making the improvements from the previous test | UX   1. Successful 2. Successful 3. Successful   Technical   1. Successful 2. Successful 3. Successful |  |

Quality of Other ECLAIR Questions

| When Tested | Outcome |
| --- | --- |
| Week Nine: Draft report that was emailed to client | 1. Successful 2. N/A 3. Successful |
| Week Eleven: final draft |  |

### 

### 3.3.2 Systematic Review Testing

Quality Tests

| When Tested | Test Results | Improvements Made |
| --- | --- | --- |
| Week Seven: after first attempt of systematic review was complete | UX   1. Successful 2. Successful 3. Inclusion criteria only contained 3 points which were very brief. 4. Successful 5. Successful   Technical   1. Successful 2. Successful 3. No written inclusion criteria identified 4. Successful 5. Successful | 1. While only criteria 3 was unsuccessful for both reviews, the overall quality could be improved. Therefore, the groups developed systematic review protocols following PRISMA-P to provide a detailed explanation of the process they are following. This included a more detailed inclusion criteria. |
| Week Ten: after second attempt of systematic review was complete | UX   1. Successful 2. Successful 3. Successful 4. Successful 5. Some articles were incorrectly included and excluded   Technical   1. Successful 2. Successful 3. Successful 4. Successful 5. Successful | 1. When incorrect classifications were discovered, all articles were re-assessed. This limited the number of included studies so a slight alteration to the inclusion criteria was made to include people that work in radiology, rather than just radiologists. |
| Week Ten: after making the improvements from the previous test | UX   1. Successful 2. Successful 3. Successful 4. Successful 5. Successful   Technical   1. Successful 2. Successful 3. Successful 4. Successful 5. Successful |  |

### 3.3.3 Meta-Analysis Testing

Quality Tests

| When Tested | Test Results | Improvements Made |
| --- | --- | --- |
| Week Eleven: final draft of report | 1. Successful 2. Successful 3. ROC and AUC not completed 4. Missing analysis not explained | 1. Discussed with the technical team the incomplete analysis and they determined that they were unable to complete it due to data challenges and a lack of time. 2. The missing analysis was explained as a limitation of the systematic review and meta-analysis. |
| Week Eleven: final report | 1. Successful 2. Successful 3. ROC and AUC not completed 4. Successful |  |

Acceptance Test

| # | User Story | Acceptance Tests | Status |
| --- | --- | --- | --- |
| 2 | As a radiologist, I want to have an accurate AI product to detect TB so that I can give the correct diagnosis. | 1. Have multiple diagnostic test accuracy studies been identified? 2. Was meta-analysis performed to benchmark the AI products? 3. Was the most accurate AI product identified? | 1. Yes, [section 3.3.3](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.uavxmyajwsw6) 2. Partially, sensitivity and specificity and SROC was completed, however, ROC and AUC were not. [Sections 3.3.4, 3.3.5 and 3.3.6](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.ydyv68x8q6e0) 3. Yes, however, not based on the complete range of expected analysis   Incomplete |

### 3.3.4 Thematic Analysis Testing

Quality Tests

| When Tested | Test Results | Improvements Made |
| --- | --- | --- |
| Tuesday Week Ten | UX Data   1. A spreadsheet was not used 2. Not all relevant data was extracted and some unnecessary data was extracted 3. N/A 4. N/A   Study Characteristics   1. Successful 2. Some demographic data was missing 3. N/A 4. N/A | 1. A spreadsheet was created for the UX data. 2. The results of the studies were reviewed to determine the relevant data. 3. More columns were added to the study characteristics table to store demographic data. |
| Week Eleven: final draft of report | UX Data   1. Successful 2. Successful 3. Successful 4. Successful   Study Characteristics   1. Successful 2. Successful 3. Successful 4. Successful |  |

Acceptance Tests

| # | User Story | Acceptance Tests | Status |
| --- | --- | --- | --- |
| 7 | As a radiologist, I want a product that will not interrupt my workflow so that I can maintain my focus on treating patients. | 1. Were studies of user experience and user acceptance identified? 2. Was thematic analysis conducted to identify factors that impact on the workflow? 3. By drawing on the thematic analysis and additional research, what was the impact of AI on the workflow? | 1. Yes, through a systematic review. 2. Yes, using a codebook and coding of data to finalise themes which were discussed. 3. This is found in section 4.2.   Successful |

### 3.3.5 Evaluation Testing

| When Tested | Test Results | Improvements Made |
| --- | --- | --- |
| Week Nine: Draft report that was emailed to client | 1. Successful 2. Some questions had brief answers with no analysis nor references. 3. Meta-analysis not complete 4. Thematic analysis not complete | 1. Expectations of the ECLAIR questions were communicated to the group, which were providing analysis and including research and references. 2. UX and technical teams had to continue working on their sections. |
| Week Eleven: final draft of report | 1. One question was incorrectly answered by discussing a source about quantitative x-rays (QXR), instead of our product qXR. 2. Some questions used non-scholarly sources. 3. 3.3 appeared unfinished because it lacked a conclusion and was missing elements ROC and AUROC. 4. Some results were missing from the results section and some discussion difficult to follow and lacking supporting evidence | 1. The incorrect question was re-written. 2. Questions that referenced non-scholarly sources were edited to remove the reference and insert a scholarly reference. 3. The technical team added a ‘conclusion and limitations’ section to finish their work and explain the missing analysis 4. The remaining results are added to the results section and the discussion and conclusion were re-structured to clearly and logically communicate the team’s discussion and conclusion |
| Week Eleven: final report | 1. Successful 2. Successful 3. Successful 4. Successful |  |

Acceptance Tests

| # | User Story | Acceptance Tests | Status |
| --- | --- | --- | --- |
| 3 | As a radiologist, I want to have an efficient AI product to detect TB so that I can identify issues and assist my patients in recovering as soon as possible. | 1. Was the efficiency of qXR researched? 2. Was qXR found to be efficient or not efficient and why? | 1. Yes, it was identified as part of the [benefits of qXR](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.mrypuxh9hq3). 2. It was found to be efficient with the reasoning in [section 2.2](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.mrypuxh9hq3).   Successful |
| 8 | As a radiologist, I want to know what problem qXR is solving in my worklife so that I can better understand the product. | 1. Is the purpose of qXR identified? 2. Are the intended users identified? 3. Is the problem that qXR is solving described? | 1. Yes, stated in [section 2.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj). 2. Yes, stated in [section 2.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj). 3. Yes, stated in [section 2.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj).   Successful |
| 9 | As a medical imaging provider, I want to have a knowledge of what qXR can bring us and what kind of patients it aims for. | 1. Are the benefits of qXR described? 2. Is it identified who the benefits are for? | 1. Yes, stated in [section 2.2](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.mrypuxh9hq3). 2. Yes, stated in [section 2.2](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.mrypuxh9hq3). 3. Yes, stated in [section 2.2](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.mrypuxh9hq3).   Successful |
| 10 | As a medical imaging provider, I want to know if there are any risks that may occur in using qXR so that I can be aware of and mitigate them | 1. Were potential risks of qXR identified? | 1. Despite extensive research, no sources on risks specific to qXR were found, however, this was satisfied by discussing general risks to all AI solutions in radiology, in [section 2.3](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.p1mm2hcxbggk).   Successful |
| 11 | As a medical imaging provider, I want the performance and validation of qXR to be wholly covered in the report so that I can evaluate its benefit to me | 1. Were all ECLAIR performance and validation questions answered with research? 2. Was the “how has performance been evaluated?” question answered by conducting a systematic review and meta-analysis? | 1. All questions answered with references. 2. Systematic review and meta-analysis processes were described and followed.   Successful |
| 12 | As a radiologist I want the specifications of the algorithm to be clear so that I can better understand how it functions | 1. Were qXR’s design specifications located? 2. Were the specifications clear? | 1. Yes, as described in [section 3.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.ydckl0jetufe). 2. Yes, as stated in [section 3.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.ydckl0jetufe).   Successful |
| 13 | As a radiologist, I want to know how qXR’s algorithm be trained so that I can better understand how it works | 1. Is qXR’s algorithm explained? | 1. Yes, in [section 3.2](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.xk5607o36x8v).   Successful |
| 14 | As a medical imaging provider, I require knowledge of how qXR’s performance has been evaluated so that I can ensure it is the best option | 1. Was a systematic review protocol created? 2. Was a systematic review undertaken following the protocol? 3. Was the data synthesised using meta-analysis? 4. Were the AI products benchmarked? | 1. [Yes](https://docs.google.com/document/d/1oJnamNJiI4bGTQpt8Bg1_v500bqte6NK5NP-_wd2I8k/edit?usp=sharing) 2. Yes, as described in [section 3.3.4](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.idtxe159mk3y) 3. Yes, as described in [section 3.3.6](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.ws0zvlz1q3v4) 4. Yes, as described in [section 3.3.7](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.khtruixjvyjn)   Successful |
| 15 | As a medical imaging provider, I want to ensure that the selected product is unbiased so that I can accept it as a valid product | 1. Have potential sources of bias been identified? 2. Have methods of mitigating this bias been suggested? | 1. Yes, [section 3.4](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.4h32iyqfn13k) 2. Yes, [section 3.4](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.4h32iyqfn13k)   Successful |
| 16 | As a medical imaging provider, I need to know whether the algorithm is adapted to new data so that I can long term use this algorithm | 1. Was it identified whether the algorithm was fixed or adapting?   Find whether the algorithm is fixed or adapting to new data | 1. Yes, [section 3.5](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.87gt5b8ojta0)   Successful |
| 17 | As a radiologist, I want to ensure that qXR can be easily integrated into my workflow so that I am not unnecessarily hindered by using it | 1. Were integration options identified and described? 2. Are the advantages and disadvantages of each option outlined? 3. Is a recommendation of which integration method made? | 1. Yes, [section 4.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.xng9gdqqtj69) 2. Yes, [section 4.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.xng9gdqqtj69) 3. Yes, [section 4.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.xng9gdqqtj69)   Successful |
| 18 | As an owner of a medical imaging practice, i want to know what are the IT requirements to implement qXR | 1. Were IT requirements identified? 2. Was a recommendation provided? | 1. Yes, [section 4.3](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.aazsam1t5atj) 2. Yes, [section 4.3](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.aazsam1t5atj)   Successful |
| 19 | As a medical imaging provider, I want to know what the data output by qXR be used for | 1. Was it explained how to export data from the system? | 1. Yes, [section 4.4](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.5woi8v41g6s8) |
| 20 | As a medical imaging provider, I want to know if the output of qXR will be given to other people instead of radiologists. | 1. Was it identified who can access the data? | 1. Yes, [section 4.5](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.yx25n24xr5mz) |
| 21 | As a radiologist I want to know if the result from qXR is interpretable or not. | 1. Was the output of the results identified? 2. Was the interpretability of the results described? | 1. Yes, [section 4.6](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.r3wg0yqvto4g) 2. Yes, [section 4.6](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.r3wg0yqvto4g) |
| 22 | As a medical imaging provider I want to be aware of the legal status of qXR in Australia. | 1. Was it found whether qXR complies with medical device laws? 2. Was it found whether qXR complies with data laws? | 1. Yes, [section 5.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.lj76shr1k7d4) 2. Yes, [section 5.2](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.59m15i60rkxc) |
| 23 | As a radiologist I want to know the licensing model for qXR. | 1. Was the licensing model for qXR described? | 1. Yes, [section 6.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.em0hu54n594n) |
| 24 | As a medical imaging team, I want to know what the training process involved to use qXR | 1. Was the user training process outlined? | 1. Yes, [section 6.2](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.syytgy6gd4gi) |
| 25 | As a radiologist I want to know how much maintenance is required once qXR has been implemented into the imaging system because I need to know how much resources it will require | 1. Was the potential maintenance required and how it is maintained identified? | 1. Yes, [section 6.3](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.5cv0fmxc13wa) |
| 26 | As a radiologist I want to know how the potential malfunctions and errors from qXR can be resolved because I need to know that any issue that arises will have instructions to solve the error | 1. Were the potential malfunctions and errors described? 2. Was it identified how these can be handled? | 1. Yes, [section 6.4](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.c1mn568fgtrt) 2. Yes, [section 6.4](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.c1mn568fgtrt) |

### 

### 3.3.6 Report Testing

Quality Tests

| When Tested | Test Results | Improvements Made |
| --- | --- | --- |
| Week Nine: Draft report that was emailed to client | 1. Successful 2. Referencing style was inconsistent, some sections lacked referencing entirely, report lacked cohesion, and informal language was used 3. N/A | 1. It was communicated to group members that the referencing style is APA7 and references need to be updated to match it. 2. It was communicated to group members that the expectation is that responses to ECLAIR questions are thoroughly researched which means that references are required. 3. Group members were advised to revise their draft responses to elevate their writing to academic writing. 4. It was decided that the managers would complete the final edit of the report, and a smaller group of people would help to create a cohesive report with a single and consistent voice. |
| Week Eleven: final draft of report | 1. Successful 2. Widespread incorrect referencing and some grammatical mistakes and informal language 3. N/A | 1. In-text references were fixed and the reference list was re-done to ensure all references were correct and in alphabetical order. 2. The managers edited the report to fix grammatical mistakes and elevate the language in some areas to be at an academic level. |
| Week Eleven: final report | 1. Successful 2. Successful 3. Report is complete so it will be emailed to the client before the presentation |  |

Acceptance Tests

| # | User Story | Acceptance Tests | Status |
| --- | --- | --- | --- |
| 1 | As a medical imaging provider, I want a comprehensive report so that I can justify my use of qXR. | 1. Does the report cover all ECLAIR evaluation questions? 2. Is there evidence of research and analysis? 3. Do the recommendations have a clear justification and explain whether qXR is worth investing in? | 1. Report contains an answer for every ECLAIR question. 2. All sections contain references and analysis. 3. The recommendations are backed by the evaluation using ECLAIR and reach a final recommendation of what to do next.   Successful |
| 4 | 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. | 1. Is there a clear summary of the evaluation conducted and recommend what to consider when investing in qXR? 2. Is there an overall recommendation of whether or not to invest in qXR? | 1. Yes, [section 7.1](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.mskrla5fquyt) 2. Yes, [section 7.2](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.lz7vgnsxcoln)   Successful |
| 5 | As a medical imaging provider, I want a reliable report so that my decision is well informed. | 1. Are the responses to all ECLAIR questions based on research? 2. Was a systematic review protocol created that includes the search queries, information sources, and eligibility criteria? 3. Was the systematic review protocol followed? | 1. All responses to ECLAIR questions contain references. 2. Systematic Review protocols were created to address those points. 3. Through testing, it was confirmed that the second systematic review correctly followed the protocol apart from minor errors that were simple to fix.   Successful |
| 6 | As a medical imaging provider, I want to have the report by my required date so that I can make decisions in a timely manner | 1. Is the report submitted to the client before the final presentation? | 1. The report will be emailed to the client the day before the presentation.   Expected to be successful. |

### 3.3.7 Presentation Testing

Quality Tests

| When Tested | Test Results | Improvements Made |
| --- | --- | --- |
| Week Twelve | 1. Successful 2. Slides not complete 3. Slides not complete 4. N/A 5. N/A 6. Presentation will be Wednesday (26/10) at 9am | 1. Slides have been created however, details and formatting has not been finalised and will be completed after this report is submitted. 2. Team members are aware of the testing criteria and some testing can only take place after the presentation has been completed. |

## 3.4 Quality Constraints

The project achieved its goal of evaluating the product qXR using the ÉCLAIR guidelines and using the evaluation to provide recommendations to the client. Overall, the project has been evaluated to have an acceptable quality, however, there are a range of areas for improvement.

### 3.4.1 Technical Knowledge

At the beginning of the project, the client provided articles for the group to read which were primarily example articles. The client also informed the group to complete the systematic reviews using PRISMA and to use the ÉCLAIR guidelines for evaluation. However, the concepts PRISMA, systematic review, meta-analysis, ÉCLAIR, and later thematic analysis were unfamiliar to all group members and therefore, had to be self-taught. These are complex research concepts which took the group a significant amount of time to learn by researching each one. Also, the group remained uncertain whether their understanding was correct due to a lack of guidance or mentorship. The group did contact the library for systematic review support, however, apart from sharing a general library guide, the library said that they only provide that guidance to post-graduate students.

After the client was concerned that the group’s first attempt of systematic review was not correct, the group undertook further research and developed a systematic review protocol to provide a clear description of the systematic review process being undertaken. This was emailed to the clients, however, the feedback received was, “the best way to check the protocol would be to try and see what you get, do some preliminary analyses, then adjust and improve. What you provided in the documents is rather high level, I can’t say if the protocol will fully work or not.” This did not clarify for us whether we were on the right track, however, we continued with the project, assuming that the protocol would be at least of an acceptable quality because it addressed all aspects of the PRISMA-Protocol which is a leading method for writing systematic review protocols.

The group started this project with very limited technical knowledge and had to self-learn all concepts and tools required to complete the project within the project time. Since all concepts were self-taught without guidance nor clarification, it is possible that some concepts were misunderstood and used incorrectly in this project, hence, a high quality cannot be completely guaranteed. However, the group conducted thorough research and worked to the best of their ability to correctly use all concepts in the project correctly.

### 3.4.2 Client Expectations

Our client, Simon Poon changed the scope for this project a few times early on in the project as described in [section 2.4](#_heading=h.9zlr6ux5effw). Additionally he shared new expectations late in the project and informed us that he would not be at the final presentation and our client is in fact Neysa Petrina. These circumstances likely impacted the quality of the project and has made it difficult to assess whether the final deliverables, presentation and report, meet the client’s expectations.

First, Simon’s role unexpectedly changed over the course of the project. Initially he said he would support and guide the group, and this included sharing articles for the first few weeks of the project. At our first client presentation, the majority of the feedback related to the group needing to provide additional explanation of what analysis we were doing. However, towards the end of the project, when the group began sharing the additional explanation and their findings with the client, he said he can’t provide feedback because he is the client, not a supervisor. He was also absent for three meetings, 21/9, 5/10, and 19/10. In the week 10 client meeting (12/10/2022) Simon said he thought that much of the ÉCLAIR guidelines responses would not be needed in providing the recommendations in the report. However, the group had spent a considerable amount of time on that section of the project because it had been included in the scope since 24/8 and was foundational to the project as it was the method of evaluation to identify the recommendations. Simon also said he had expected that the group would have recorded every hour of work done throughout the project in a logbook and provide it at the conclusion of the project. He said that this is how he would have worked out how much to pay us if we were being paid. However, the group did not know they were expected to do this because it was the first time he stated this expectation and with 2 weeks of the project remaining it was not possible to do.

In week 4 (24/8), Simon introduced us to Neysa Petrina who works in the project area for the purpose of providing a deeper context to the project. She recommended using the ECLAIR guidelines for evaluation. Over the course of the project she provided feedback to the group in most meetings, however, Simon remained the primary person we discussed the project scope and project with. On 19/10 the group received an email from Simon stating that he was unavailable for our final presentation and it would be more suitable for Neysa to evaluate us because she is the client. The group wasn’t aware until this email that Neysa was introduced to us to become the client. The group did speak with Neysa about her expectations for the final presentation and endeavoured to meet them. In the week 11 client meeting (19/10) Neysa shared expectations of being presented with how the diagnostic test accuracy of qXR compares to its competitors, the user experience of qXR and based on the ECLAIR guidelines, why she should choose qXR and important considerations when investing in it.

Therefore, the final project may not entirely meet the expectations of our client due to these changes, however, the group worked to the best of their ability to pivot each time they were faced with a change to the project.

### 3.4.3 Time

A major quality constraint for this project was time, particularly due to the two constraints explained above. Needing to self-learn all concepts for the project and the scope changing over the course of the project significantly limited the amount of time spent on the project itself. The group worked hard to complete the project; however, the quality would have been higher if more time was available. A particular situation where this made a difference was for the systematic review. Our client requested that the group undertake a systematic review, however, the systematic review that we undertook was closer to a rapid review because systematic reviews are expected to take at least 12 months. A systematic search was undertaken for the search queries used; however, it is highly unlikely that our search queries covered the entire breadth of research we were looking for. Whereas, if more time was available, more time could have been spent designing search queries and using more to identify all articles in the area of research.

# 4 Discipline knowledge and tools

## 4.1 Research discipline

The primary research discipline for this project was information systems in healthcare. The branch of healthcare this project investigated was the use of information systems, specifically artificial intelligence tools in diagnostic radiology. Team members who had completed ISYS3401 already had a basic understanding of this research discipline. Throughout the project team members learnt about how AI tools are used in radiology to assist in diagnosis and its potential benefits such as optimising the workflow and identifying diseases earlier.

Within this discipline, the team was required to evaluate a specific AI tool, qXR, using methods introduced by the team’s clients. The primary evaluation method was the ECLAIR guidelines and included systematic reviews, meta-analysis, and thematic analysis. The ECLAIR Guidelines are a practical set of questions to answer when deciding which commercial AI solution is most appropriate for radiologists to use (Omoumi et al., 2021). Systematic review is a method of literature review following a particular process and is widely accepted by the medical imaging community which is the team’s targeted audience. Systematic review combines the outcomes of various primary studies associated with each other by utilizing bias and error reduction methods (Ganeshkumar & Gopalakrishnan, 2013). Meta-analysis is the quantitative part of systematic review which is designed to combine and analyse the data from the systematic review results (Haidich, 2010). Thematic analysis is a flexible, qualitative research method for identifying and summarising themes within a dataset (Nowell, Norris, White, & Moules, 2017).

### 4.1.1 Academic background

According to the objectives of this project, the team members conducting this research all hold sufficient knowledge regarding the study of information systems. The team is constructed with students from the University of Sydney who are currently studying information systems as one of the majors. Four of them had already taken the unit, ISYS3401, as one of the degree cores. It introduced those team members to the research discipline, statistical analysis concepts, and evaluation models for user experience. These preliminary academic backgrounds provided a starting point for further research into this research discipline. The knowledge and ability could be presented as the following factors:

* Full understanding of the information system as an academic discipline.
* Practical experience regarding university level information system projects.
* Four had basic knowledge of the research discipline.
* Four had experience with models to evaluate user experience, these being the Technology Acceptance Model and the Unified Theory of Use and Acceptance of Technology model.
* Four had knowledge of statistical methods for quantitative analysis.
* Academic English composition skill.

### 4.1.2 PRISMA

The **PRISMA process** was used to conduct the systematic review, as requested by the client. To summarize what the PRISMA process is, referencing a guideline from the University of North Carolina (The University of North Carolina, Libguides: Creating a PRISMA flow diagram: Prisma 2020), it is a five-step procedure including collecting all the relevant articles can be found in multiple databases, removing duplicates, removing unrelated articles based on abstract, assessing the quality of the full text of the remaining researches and extracting the valuable data contained. To ensure this process was followed, systematic review protocols were developed following the PRISMA protocol process. To visually display the completed process, a PRISMA flow diagram was created for each systematic review. Once this process was complete, the team had relevant and high-quality data available for synthesis.

### 4.1.3 Thematic analysis

To synthesise the data for the user experience team, **thematic analysis** was chosen as the most appropriate method because the data extracted varied significantly and included qualitative data. Thematic analysis is specifically effective regarding evaluation of user experience and acceptance because it can often reveal deeper insights into the opinion and attitude of participants. There are six steps to be followed to complete thematic analysis, familiarisation of data, generation of initial codes, searching for themes, reviewing the themes, defining and naming themes, and producing reports (Kiger & Varpio, 2020).

For our project, we decided to conduct a primarily deductive thematic analysis by creating a codebook that used the UTAUT model as the conceptual framework prior to coding the data. However, to avoid restriction, team members also created new codes and themes that emerged from the data and added them to the codebook. The team individually coded the data and then met to discuss their views and reach a consensus for the codes. Finally, the teams wrote the results for each theme and discussed the impact of the themes on the workflow through the lens of radiologist’s user experience and acceptance.

.

### 4.1.4 The ECLAIR guidelines

The project uses the **ECLAIR guidelines** as the evaluation method which was published in 2021 and is used to determine which commercial AI solution in radiology is most appropriate for a radiologist’s or organisation’s needs. Using this guideline ensures that the team presents to the client a comprehensive evaluation and resulting recommendations that are based on relevant and important issues.

It consists of a total of twenty questions which are divided into five sections which are relevance, performance and validation, usability and integration, regulatory and legal aspects, and financial and support services considerations (Omoumi et al, 2021). All of these questions were answered through researching relevant information, analysing that information and providing a recommendation based on the answer.

### 4.1.5 Diagnostic test accuracy

To measure qXR’s performance against its competitors, their diagnostic test accuracies were benchmarked. Diagnostic test accuracy refers to the ability of a test to distinguish between a healthy person and a person with the disease or illness (Šimundić, 2009). The diagnostic test accuracy measures used were sensitivity and specificity.

Sensitivity is the ability of a test to identify patients with a disease correctly and specificity is the ability of a test to identify people without the disease correctly. However, the project team cannot directly compare these two values, and high-specificity experiments are likely to cause over-medicalization. Because highly specific experiments tend to capture some people who do not have the disease, it also creates additional anxiety for patients and wastes unnecessary resources (Swift et al., 2019). Referring to these two abilities simultaneously, the ability of AI products to identify patient status can be obtained and avoid false positives and negatives. At the same time, when comparing AI products, to determine the significance of the results, we usually compare another value under specific sensitivity/specificity conditions.

In the comparison process, in order to visualize our results and avoid the limitations of specificity and sensitivity, we used an SROC graph. An SROC graph is a summary of the performance of diagnostic testing for different studies on the same graph; SROC curves are often used to graphically show the association/trade-off between clinical sensitivity and specificity.

### 4.1.6 Meta-analysis

To synthesise the data for the technical team, meta-analysis was chosen because most of the data used the same diagnostic test accuracy measures. Meta-analysis is a method of statistically combining results from multiple studies to yield overall statistics with confidence intervals (Deeks, Higgins & Altman, 2022).

A Forest plot is an essential tool for meta-analysis. It provides each study's CI and the effect size (weight). This also means that when the project team conducts a systematic review without meta-analysis, the project team can also use a forest plot to determine which articles have a more significant impact on the project.

At the same time, forest plots can also reflect the degree of difference between different studies, and the difference in research results on the same topic is called heterogeneity. Heterogeneity is considered a prerequisite for whether or not a meta-analysis can be performed; when the directions of outcomes in the studies are similar, then heterogeneity is less likely, and a meta-analysis is feasible (A Guide to Understanding Meta-Analysis | Journal of Orthopaedic & Sports Physical Therapy, 2020).

## 4.2 Tools

In order to evaluate qXR and provide recommendations at the required level that the clients proposed, multiple tools were used by the team. The tools used were divided into four categories, team communication, file management, data analysis, and academic databases. Different tools and disciplines were used based on targeted obligatory functionalities.

### 4.2.1 Team communication

First of all, the team used the application ‘**Slack**’ to achieve effective team communication. This application was introduced by the team tutor, and is a real-time messenger for the team to communicate. It can be downloaded to mobile devices or accessed through browsers which demonstrates a high level of convenience. More importantly, different channels could be created in order to enhance the efficiency of communication. For the team, five channels were used in total, the general channel for general matters such as establishing deadlines and task distributions, the meeting channel for scheduling meetings and discussing the agenda prior to the meeting, the protocol reading channel for easy access to readings regarding protocol and sharing of new and relevant protocols finally the tech channel and the UX channel were used for each team’s inner communication and discussion of their specific report-related requirements. Additionally, team members were able to privately chat with other members of the Slack and this was used effectively in communicating individual needs.

The team was geographically distributed between two countries, therefore, **Zoom** was our main tool for holding group meetings. Not only did Zoom allow us to communicate remotely, but there were many other functions to assist the project team's efficient work. The "breakout room" feature allowed us to have breakout discussions and multi-threaded work as our team was split into the UX team and the tech team. At the same time, the "recording" function supported us in preparing deliverable demonstrations for our clients and tutors and allowing for feedback to be more precise.

### 4.2.2 File management

To effectively manage files, the team utilised Google Drive and Bitbucket. **Google Drive** was the primary tool used as it provided the team with cloud storage for many project documents, allowed the team to work on the same file simultaneously, and provided version control. The team separated files into designated folders to organise them and make them easy to track. **Bitbucket** was also used as a file storage according to the requirement from the project tutor. Even though it did not function as well as Google Drive because of the nature of this project as an evaluation instead of programming, it was still useful when it comes to distinguishing individual contributions when tracking the commit record. It was also used to store all meeting minutes that were recorded, allowing absent members to catch up as well as clearly displaying what occurred in the meeting and allowing us to understand what was further required.

### 4.2.3 Data analysis

All of the data extracted for analysis was manually entered into **Google Sheets** files. This ensured that the data was stored in an uniformly organised and structured way which makes it easier to be analysed.

The technical team was required to construct various graphs during the process of meta-analysis, **Tableau** was chosen as the tool for this purpose because it has a more negligible learning cost and lower mastery difficulty than **RStudio**. In addition to being a part of the process of meta-analysis, these graphs provide a visualisation of the results which can make understanding the results quicker and easier.

### 4.2.4 Article Collection

Since our project was research based, a foundational part was identifying high-quality articles for analysis. To assist in this process, **academic databases**, such as ProQuest, Pubmed, Embase, and ScienceDirect were searched. These databases were selected on the basis of being recommended by the University of Sydney Library and having a good reputation of high quality, unbiased, and peer reviewed articles. This meant that our searches did not include non-academic sources which can be found through Google Search or grey literature and articles with an ambiguous academic integrity through Google Scholar.

# 5 Group Processes

Our project was to create a technology evaluation research paper, therefore it was difficult for us to follow the extreme programming process as suggested by the unit coordinators. Due to the length of time it takes to research and write a high level report for this project, we were unable to consistently write up modified reports to garner feedback from our clients. Over the course of this semester we utilised a combination of techniques to complete the project, one of the methods used was the waterfall method, where we would all work on one task at a time before moving on to the next, this was effective as it allowed us to collaborate and more efficiently create a valid output. To begin the project itself, our group took time to research and develop an understanding of the research process, only once our entire group was familiar with the research process would we be capable of planning and conducting our research. We planned to complete all relevant research before we would begin the writing of the report for our client as this would simplify the process and prevent missteps. To ensure progress was being made throughout the semester, our group allocated extreme programming roles to the team members to help make sure the group was staying on track and we were outputting high quality work.

Based on the advice from our client, our team also built a unique structure and roles based on the objectives of the project, where we split the team in half based on the two main focus areas, UX and technical evaluation, which enabled us to work on the UX and technology sections at the same time. The two teams separately conducted their own research and once finished began writing the relevant report sections, this allowed the project to come together at a faster rate and dynamic teams also enabled us to quickly reconfigure in accordance with scope changes.

## 5.1 Application setup

Our team heavily used the tools Slack, Google Drive and Bitbucket. Despite a majority of our team having limited experience in several of these tools, we effectively used them throughout the semester for group collaboration, this section will detail the process of setting up these tools for continued use.

Our team found Slack straightforward to deploy and a good platform for instant messaging. Our team created several channels on slack based on different teams as well as more general information relevant to all members. We were able to join using a link and ensured the tutor was also in the channel in order to gauge our interactions and provide feedback.

Bitbucket was found to be more difficult because none of our group members had prior experience with it as a tool. Bitbucket operated through git which required our team to follow unique code instructions that took a while to understand. However, several team members have used git in the past which helped them understand how to use bitbucket quickly. Once the repository was created we found it easy to join and access the shared files, experienced team members were able to assist their less experienced counterparts with bitbucket to ensure they got up to speed with the software. The team reached a unanimous decision that bitbucket will be helpful for future projects, especially coding-related projects.

To help manage and allocate our team’s tasks we created a Gantt chart. However, due to the number of task reallocation and changes our team ended up using a spreadsheet on google drive where tasks were written, allocated and adjacent columns that included the task’s status and due date. This task tracker was created and managed by the group leader Cailin, who would ask the team to keep updating their task rows. Our team also would use the SCRUM standup method to ensure progress was being made in the weeks, which was facilitated by Cailin who would ask the team to write down the answers to the three questions weekly. Those questions include:

1. What work have you done since we last met?

2. What work are you planning on doing?

3. Roadblocks faced?

Throughout the semester our team used Google Drive the most. Our team found it made more sense to use this tool due to the fact our assignment is a collaborative research report. Google Drive allows the group to work on the document at the same time and see the real-time changes being made without needing to push and pull which can become tedious. As we had all used this tool in the past, setting up and understanding the tool was efficient and we were able to see results at a faster rate.

Evidence: [Google Drive](https://drive.google.com/drive/folders/1kfOahR374bFx8T5_THx_djK7jEDF5kdh), [Bitbucket](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/), [Slack](http://isys3888group-bmj6211.slack.com)

## 5.2 XP Roles

To ensure our group was making progress with the project throughout the semester as well as to make sure the content produced is high quality, XP roles were allocated. To make the most of our capstone unit, we decided to do weekly XP role rotations amongst the team members which gave everyone a chance to work on a variety of skills that are required by the different roles. However, reflecting on this choice to rotate roles, caused tremendous confusion amongst our group members, this issue is further explained in [section 6.4](#_heading=h.pc5viraq1r8f).

Role 1: Tracker

The tracker’s (intended) tasks included:

1. Calculate the projected velocity

2. Check-in with how the programmers are doing

3. Take action if the project is going off track e.g. meet with the client, ask another programmer to help etc.

Our team’s plan was for the tracker to be responsible for setting the tasks and due dates for each team member to complete in the next sprint according to the timeline and current workload. At the same time, when the phased tasks are completed, the tracker will check the quality of the project.

Role 2: Manager

The manager’s (intended) tasks included:

1. Records meeting minutes

2. Ensures proper processes for meetings are followed

Our team’s plan was for the manager to be responsible for assisting the leading project manager during the meetings and recording the meeting minutes.

Role 3: Customer Liaison and Lead Project Manager

The customer liaison and lead project manager’s (intended) tasks included:

1. Contacting the client and negotiating client meeting time
2. Look at the big picture of the project,
3. Communicate with the sub-team leaders
4. Create and allocate tasks
5. Schedule and lead group meetings

Due to the limited work required by the customer liaison, the role was combined with the lead project manager. Our team’s plan for the customer liaison and lead project manager was to ensure constant communication and project progress was being made, this was done through scheduling meetings, slack messages and emails.

Role 4 and 5: Programmer(s)

The programmer’s (intended) tasks included:

1. Researching and understanding the assignment requirements to then explain to their team members

2. Defining and allocating tasks from user stories

3. Estimates duration of user stories and tasks

4. Creating corresponding unit tests

This project does not contain coding, so our team redefined the role of the programmer to lead the two sections of the report, UX and technical. Our team’s plan was for our two team leads to guide their respective groups to ensure they finish their sections of the report on time and with high quality. The team leads planned on doing this by assisting their group members when their team members appeared heavily confused or off track.

Role 6: Tester

The tester’s (intended) tasks included:

1. Implements and runs tests (excluding unit tests)

2. Graphs test results

3. Notifies others when a test fails

Our team’s plan was for our tester to constantly review the new work made each week in the report, to ensure the content is accurate and original as well as including our conclusions. This role is to ensure the quality of the report remains high and to alert the group member to re-write their portion if they found it to not match the expectation required by our client.

Role 7: Doomsayer

The doomsayer (intended) tasks included:

1. Monitors project status

2. Looks ahead to identify any future issues

3. Notifies others when a problem or challenge arises

4. Notifies others of the risks involved

Our team’s plan was for our doomsayer to be responsible for project warnings, they would notice upcoming risks and inform the team of current threats to the project, we would also have this member propose solutions to mitigate risk.

Evidence: [Group contract and deed poll](https://drive.google.com/drive/folders/1C0bS96LmWLMyLzTi-RB788CfyD8niaN8)

## 5.3 Team Structure

As suggested by our client, our group project members split into two sub-teams, the sub-teams were divided up by the two main objectives of the report, the technical and the user experience evaluation of CAD for TB technologies. Each team was appointed a team manager to manage internal matters and completing tasks such as: understanding the assignment, task creation and allocation based on user stories and deadline management.

To make sure there was consistent communication between the clients and within the team, Cailin was chosen to be the project manager for the whole project, she communicated with clients through emails and organised and led meetings through zoom. Cailin was also in charge of task distribution and deadline management for the whole group and throughout the semester, would check on the status of each team member, and monitor the progress of the two sub-groups. For all the documentation tasks, including presentations, client reports, group reports, and progressive reports, the team was able to divide the task relatively evenly where each of the team members would self-allocate portions of the work to do from the templates provided from the assignment briefs. Our project manager made sure there existed no overlaps with work and would identify if there were missing parts of the assignment and would allocate that part to a member of the team who had already completed their given task. The exact task distribution is presented in the individual contribution part of this report.

Evidence: See [slack](http://isys3888group-bmj6211.slack.com) channel.

## 5.4 Tech Team

Kiran was the team manager of the technical team. His team members were Andersen, Zichen, and Zheyuan, the main task of the technical team was to evaluate the accuracy and sensitivity of the product qXR and compare it with other existing medical imaging products and then identify the best-performing product for our client.

To improve work efficiency, different CAD products were distributed to all the members of the technology team therefore all the gathering of articles could be done concurrently, while this did lead to further duplicates the process was much more streamlined.

1. CAD4TB - Kiran
2. RADIFY - Kiran
3. InferRead DR Chest - Zichen
4. JLD-02K (JVIEWER-X) - Zichen
5. INSIGHT CXR - Andersen
6. ChestEye - Zheyuan
7. AXIR - Zheyuan
8. VUNO - Andersen

Everyone in the tech team extracted data from their corresponding articles and were then equally assigned articles for further analysis, Andersen was responsible for most of the data cleansing work and Kiran introduced various analysis methods to the tech team. However many of the methods were not able to be implemented due to the data limitations as well as general difficulty applying them with limited discipline knowledge.

## 

## 5.5 UX Team

The manager of the user experience team was Bronte and the team members of the UX team were Cailin and Benjamin. Bronte guided the UX team throughout the entire semester as she spent an extensive amount of time learning and understanding the systematic review process and the ECLAIR method, this research outputted documents that outlined protocols for our team to follow and gave her team tremendous guidance. Bronte was able to give critiques on the work done by the user experience team members ensuring high quality work. The work of the User experience team involved conducting a systematic review, where the tasks were distributed evenly among the three members, this included article collection, data extraction, study characteristics extraction, creating the code book and coming up with conclusions to present to our client based on our findings.

## 5.6 Team Communication

Most of the communication was through our Slack channels and meetings conducted throughout the week. The language used by each team member was always appropriate, with no occurrence of harsh language ever being used. Due to this respectful and professional tone, arguments were avoided during the team collaboration process.

Throughout the semester our team would have one hybrid tutorial meeting on Tuesday, which would also include a tutor meeting and one group virtual meeting on Fridays, our team would also attend a weekly client meeting on Wednesday. These meetings were scheduled and led by our group project manager Cailin. The purpose of these meetings would be for the team to update the project managers on their progress and highlight issues, the meetings would often involve the lead project manager changing tasks and their level of priority based on upcoming assignment deadlines as well as client feedback.

Evidence: See [slack](http://isys3888group-bmj6211.slack.com) channel, [meeting minutes](https://drive.google.com/drive/folders/1B1a83AgfKcrpq5OR-qm5QMo6DdQMX8YS).

## 5.7 Client Communication

In spite of our original client being Simon Poon and his requirement to see us weekly for an hour and respond to any issues throughout the week, his commitment to this project deteriorated heavily by the end of the semester. Our original agreement was to meet on Wednesdays, in which he would provide feedback on the team’s previous work and instruct the team with further requirements.

At the start of our semester, Simon provided an extensive amount of research papers to help us get started on the project. However, this led to confusion as his help and scope of the project had the group see Simon more as a supervisor rather than a client. Several weeks into the project Simon realised how heavily reliant our team has been on him and Neysa in which he then made it clear in the last client meeting before the mid-semester break that he is a client, not a supervisor and any struggles with academic review could be helped by the librarians, which although understandable, it put our team off balance and lost.

A few weeks into the semester Simon included Neysa Petrina to be part of our capstone project to act as an external consultant as he believed her experience would benefit our team. Due to Simons’s lack of commitment, he heavily relied on Neysa’s attendance for the client meetings to make up for his absences, which included Simon missing three client meetings in which he would either last minute notify us or not at all. Simon also failed to inform the team that he will not be present for our final project deployment and allocated Neysa to be our client with one week’s notice.

Throughout the semester, our team’s meetings ended up not being as productive as planned. This was due to the high level of expectations Simon and Neysa had for our team. They had expected our team to start the research process right at the beginning of the semester, which was not possible because the team needed time to learn how to perform a systematic research review as the methods required were not learned in prior units of study. It was this failure of meeting expectations that caused friction between the client and students. We believe that our team would have benefited from a supervisor rather than a client as we required guidance and it was unrealistic to learn and apply the methods of research that were expected in this project completely independently. This issue will be expanded upon in the next paragraph.

Evidence: [Client communication](https://drive.google.com/drive/folders/1cL6X_sMAl_iENIAnzTq9V-1Hcl5d2x6i) emails, [meeting minutes](https://drive.google.com/drive/folders/1B1a83AgfKcrpq5OR-qm5QMo6DdQMX8YS).

## 5.8 Issues and Solutions

Several issues were faced by the team which had a significant effect on the quality of the report. Luckily the team’s managers were resilient and able to think on their feet and come up with solutions.

As previously mentioned in 5.7, one of the biggest issues the team faced was the client’s unreasonable task timeline expectation because they did not take into account how long it would take for our team to have a well-rounded understanding of what we were being asked to do. Their advice was to have the team seek librarians’ help to get started with the systematic review process, when three project leads went to the library to seek help, the librarians said they are unable to as we are not post-graduate or honours students. This lack of academic support resulted in our team needing to figure out how to conduct systematic reviews ourselves without any additional help. Our head of UX, Bronte took the initiative on researching and understanding the systematic review process, once Bronte had felt confident enough with her research, she then made protocol documentation to help guide the teams with their research. Cailin also made sure Bronte explained the review processes clearly to the team in the following meeting, ensuring that everyone in the group had a good understanding of the requirements and the steps to follow.

Another issue that our group faced was not tracking tasks consistently throughout the entire semester. Although the group made a Gantt chart at the start of the project, it was not kept up to date. A reason for this could be the constantly changing tasks prioritisation. At the halfway mark of the semester, our project lead built a task tracker spreadsheet that listed out all the assignments and associated tasks and had columns denoting the team member allocated to the task, status and task due date, This spreadsheet ensured the team was up to date with deadlines and allowed for team members to easily update the status of their allocated tasks.

Due to the nature of our report and having the majority of the semester with the team not having a clear timeline of work to do there was concern amongst the managers, who realised the lack of work done by the group members each week. The solution to address this concern, as Cailin suggested, was to take inspiration from SCRUM meetings and each Friday have our group answer the following questions:

1. What work have you done since we last met?

2. What work are you planning on doing?

3. Roadblocks faced?

However, reflecting on the team’s progress near the end of the semester, the managers realised they should have had the team members provide evidence for the weekly work done as many times the extent of work done was not evident in any shared documentation.

# 6 Reflections and Conclusions

## 6.1 Challenge and Risk Analysis

The biggest challenge our group faced was time, evidence of this is the majority of issues mentioned on the weekly SCRUM Slack update meetings, where ‘time’ was the most frequent issue amongst the entire team. As mentioned in several sub-sections of Section 5 of the report , a large portion of this semester involved our team having to figure out the requirements and processes on our own. Therefore the team had to rush research, analysis and report writing, user testing and editing and this rushed work led to a variety of risks. The rush impacted the level of quality of the paper as we were not able to take our time to edit, further, due to time limitations we were unable to present our client with a copy of the final report for feedback to improve upon before final submission, therefore, our team is at risk of not providing exactly what our client expects and thus not successfully delivering on our report.

Several challenges that the individual groups experienced were:

**Technical Side:** No articles cover the performance of all the certified AI products so it’s difficult to create valid comparisons across different articles. The measurement method or threshold for different articles to test sensitivity and specificity is frequently different in each article, therefore,the technical team was restricted in using results from different articles to do the comparisons. However, the technical team managed to do cross-article comparisons in the end.

**User Experience Side:** The main challenge for the user experience team is that there are few articles related to user experience of qXR. Therefore, the UX team needed to expand search scope to find any user experience articles related to medical imaging instead of just searching for qXR.

## 6.2 Limitations

Due to the nature of our project, there are some limitations for both the technical side and user experience side.

**Technical Side:** A major limitation for the technical team was a lack of articles and data of certified AI products. Some AI products such as VUNO only had the korean-language articles which can’t be used for evaluation as we had excluded these in our systematic review protocol. Some AI products such as JLD-02K had not had studies on detecting TB, therefore, the technical team couldn’t find any relevant articles. In addition, excluding many articles was also a limitation for the technical team, since the technical team needs to make sure the data isn’t biased, we were forced to exclude articles released by the product owner. Therefore, after filtering all the articles, the technical team only had 6 articles for evaluation. Another limitation is that the data provided in each article is not sufficient to do a complete meta-analysis, some articles only provided a pair of sensitivity and specificity values for each product which was not sufficient to draw a comprehensive SROC curve in the meta-analysis process. However, the technical team was able to address this limitation by drawing forest plots to compare sensitivity, specificity and odds ratios for each product in each article.

**User Experience Side:** The primary goal for the UX team is to evaluate user experience related to detecting TB, however, there were few articles related to this aspect, consequently, the UX team expanded their search scope to all medical imaging AIs in order to support the evaluation. As the team only had about 12 weeks to finish the evaluation, it was quite difficult to identify all the relevant articles in medical imaging, therefore we had limited time to identify the most appropriate articles for evaluation and may not have identified articles that would have been more appropriate.

## 6.3 Primary Strengths

The biggest strength our team had was our project leads going above and beyond their required tasks, as mentioned previously, for a long period of time we had no structure because we were unsure how best to approach the report. The UX lead, Bronte took the initiative and spent a large amount of time researching how to perform the systematic reviews, Bronte was one of our team’s biggest strengths as her learnings were shared with the rest of the team to inform the steps to follow in systematic review, she wrote a detailed document outlining the steps that was followed by all members.

Another strength our team had was the suggested frameworks to follow when writing our report, the ECLAIR and PRISMA guidelines. This helpful suggestion from our client meant that our group did not need to spend time finding a way to evaluate the product. These guidelines have a large amount of research proving the high level of quality a research paper would hold through use of the ECLAIR and PRISMA guidelines.

An additional strength our team had was that a majority of our team members completed the unit ISYS3401. Therefore a large portion of the group already had a large amount of experience with the UTAUT models and researching the various CAD for TB certified products.

Our final strength was knowledge of the significance our report would have. Through our research we see that TB is extremely serious if left diagnosed too late hence why there is an incredible need for these CAD for TB products to be implemented in all radiological practices. This importance kept our team motivated throughout the semester as we knew that our research will save stakeholders time comparing the various CAD products on the market by suggesting the highest performing technology as well as noting the most significant factors that impact the clinical workflow when integrating this technology.

## 6.4 Programming Practices

An issue our group faced is that the team members who were meant to be rotating the manager, tracker, tester and doomsayer did not do all the required tasks for that role. These essential roles were required to ensure high quality reports and ensure deadlines would be met. As the managers knew the importance of these roles, these roles were taken over by Bronte and Cailin. These leads consistently checked the current tasks being worked on, monitoring the progress and quality of the tasks.

Furthermore, due to inability to fulfil their roles, the managers did not have the time capacity to create a system to note down issues and consistently update it, instead our solution was for our project lead to ask the group at the end of the meetings if they had any issues and if issues were raised, then the team would then discuss the action steps needed to take place. Our project lead also would heavily encourage the group members to also raise issues on slack throughout the week and on Fridays. The third SCRUM question was dedicated to issues faced where threads could be made to help solve the issues through group collaboration.

# 7 Individual Contributions

Table of Technical Contributions

| Name of Task | **Bronte Lee** | **Zichen Li** | **Cailin Feldman** | **Jingbang Men** | **Kiran Papalkar** | **Baoxi Liu** | **Zheyuan Zhang** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Executive Summary | [UX Findings](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.c2xkcg2dat9d) | [Overview](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.c2xkcg2dat9d) |  |  | [Problem Articulation](https://docs.google.com/document/u/0/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit) | [Technical Findings](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.stcjof8qfezi) | [Recommendation](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#) |
| What problem is the application intended to solve, and who is the application designed for? |  |  |  |  | [**✓**](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) |  |  |
| What are the potential benefits, and for whom? |  | [**✓**](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) | [2.2 What are the potential benefits, and for whom?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.mrypuxh9hq3) |  |  |  |  |
| What are the risks associated with the use of the AI system? |  |  |  |  |  |  | [**✓**](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) |
| Are the algorithm’s design specifications clear? |  |  |  |  |  | [**✓**](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) |  |
| How was the algorithm trained? |  |  |  |  |  | [**✓**](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) |  |
| How has performance been evaluated? |  | [wrote 3.3.2 and the first 2 articles of 3.3.5](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) |  |  | [Wrote 3.3.1, 3.3.3, 3.3.4, 3rd and 4th article of 3.3.5, SROC graph + write up of 3.3.5, 3.3.6](https://docs.google.com/document/u/0/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit) |  |  |
| Technical Systematic Review Protocol |  | [Objectives, data management, outcomes and prioritisation](https://docs.google.com/document/d/1oJnamNJiI4bGTQpt8Bg1_v500bqte6NK5NP-_wd2I8k/edit) |  |  | [Wrote sections: Rationale, Information Sources, Search Strategy, Data Collection Process, Data Items, Data Synthesis, Meta-biases, Confidence in cumulative evidence](https://docs.google.com/document/u/0/d/1oJnamNJiI4bGTQpt8Bg1_v500bqte6NK5NP-_wd2I8k/edit) | [Search strategy](https://docs.google.com/document/d/1oJnamNJiI4bGTQpt8Bg1_v500bqte6NK5NP-_wd2I8k/edit#heading=h.4qbnyg5edg9a)  [Data management](https://docs.google.com/document/d/1oJnamNJiI4bGTQpt8Bg1_v500bqte6NK5NP-_wd2I8k/edit#heading=h.4qbnyg5edg9a) |  |
| Technical Systematic Review |  | [evaluate AI products InferReadDR and JLD-02K](https://docs.google.com/document/d/1W8iuZQUIGnEWvmU8oMOaIRMLAHrjquiFD5JuZYmemcU/edit) |  |  | [Identified 12 articles related to DTA for CAD4TB and RADIFY](https://docs.google.com/document/d/1W8iuZQUIGnEWvmU8oMOaIRMLAHrjquiFD5JuZYmemcU/edit) | [evaluate AI products LUNIT and Vuno](https://docs.google.com/document/d/1W8iuZQUIGnEWvmU8oMOaIRMLAHrjquiFD5JuZYmemcU/edit) | [evaluate AI products ChestEyeand AXIR](https://docs.google.com/document/d/1W8iuZQUIGnEWvmU8oMOaIRMLAHrjquiFD5JuZYmemcU/edit) |
| Technical meta-analysis |  | [Draw the forest plots for 2 selected articles](https://docs.google.com/document/d/1Ur6pdGAVgvLBD3Fs5BcM9zVMRXc8qpzNshj0M-SUofo/edit) |  |  | [Drew Qin and Kagujje forest plots](https://docs.google.com/document/u/0/d/1Ur6pdGAVgvLBD3Fs5BcM9zVMRXc8qpzNshj0M-SUofo/edit)  [Drew SROC graph (3.3.5)](https://docs.google.com/document/u/0/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit) | [Draw the forest plots for article](https://docs.google.com/document/d/1Ur6pdGAVgvLBD3Fs5BcM9zVMRXc8qpzNshj0M-SUofo/edit) 7 | [Draw the forest plots for article 5 and 6](https://docs.google.com/document/d/1Ur6pdGAVgvLBD3Fs5BcM9zVMRXc8qpzNshj0M-SUofo/edit) |
| Have the developers identified and accounted for potential sources of bias in their algorithm? |  |  |  |  | [**✓**](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) |  |  |
| Is the algorithm fixed or adapting as new data comes in? |  |  |  |  | [Major rewrites](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) |  | [**✓**](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.vqn6m0hq7aj) |
| How can the product be integrated into your clinical workflow? |  |  | [4.1 How can the product be integrated into your clinical workflow?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.xng9gdqqtj69) |  |  |  |  |
| How exactly does the product impact the workflow? | [Abstract, introduction, methods, systematic review of qXR, and systematic review of UX](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.d6ue6fdugvfe) |  |  |  |  |  |  |
| UX Systematic Review protocol | [Data synthesis, meta-biases, and confidence in cumulative evidence](https://docs.google.com/document/d/1fvetkO7wzLRR7s-9ukJXNDwvOyJ5Wf3UZVyRBuy1Jz0/edit#heading=h.jo6c9edby6pe) |  |  | [Introduction + Methods](https://docs.google.com/document/d/1fvetkO7wzLRR7s-9ukJXNDwvOyJ5Wf3UZVyRBuy1Jz0/edit#heading=h.jo6c9edby6pe) |  |  |  |
| qXR Systematic Review protocol | I[ntroduction section and made minor edits to adapt all other sections from the UX Systematic Review protocol to be specific to qXR](https://docs.google.com/document/d/10bP45WN_cmBoHZwW8HTuMwZA_uikEy_RQw8eqWT-X9c/edit?usp=sharing) |  |  |  |  |  |  |
| UX Systematic Review | [qXR tab and articles 137-156 in AI in radiology tab](https://docs.google.com/spreadsheets/d/1RhzX0JjUH6D0ZRoSSnJewTDqgtng03jA3h9OQ1VyBVI/edit?usp=sharing) |  | <https://docs.google.com/spreadsheets/d/1RhzX0JjUH6D0ZRoSSnJewTDqgtng03jA3h9OQ1VyBVI/edit?usp=sharing> - science direct and the lancet | [All the articles excluding Bronte and Cailin’s](https://docs.google.com/spreadsheets/d/1RhzX0JjUH6D0ZRoSSnJewTDqgtng03jA3h9OQ1VyBVI/edit?usp=sharing) |  |  |  |
| UX Codebook | [Introduction, methods, and performance expectancy theme](https://docs.google.com/document/d/17DE7Ovc5DS25iIhcqy0-nlGFoVUyjGQpqp0LRJdit_s/edit?usp=sharing) |  | <https://docs.google.com/document/d/17DE7Ovc5DS25iIhcqy0-nlGFoVUyjGQpqp0LRJdit_s/edit> effort expectancy, social influence, facilitating conditions | [personal attitude](https://docs.google.com/document/d/17DE7Ovc5DS25iIhcqy0-nlGFoVUyjGQpqp0LRJdit_s/edit?usp=sharing) |  |  |  |
| UX data extraction | 1 [Study characteristics except demographics](https://docs.google.com/spreadsheets/d/1MW8KNZpBNYQtpLGp4bm2e5EJi1HzDjVB8iNpfZuWj_Y/edit?usp=sharing)  2 [Data for study ID 1](https://docs.google.com/spreadsheets/d/16Dn5pyBYoqCTGyaVS5_1dafeEufuFVUn7lz_4Q3MsFI/edit?usp=sharing) |  |  | 1. [Study characteristics demographics part](https://docs.google.com/spreadsheets/d/1MW8KNZpBNYQtpLGp4bm2e5EJi1HzDjVB8iNpfZuWj_Y/edit?usp=sharing)  2. [Data for study 2-9](https://docs.google.com/spreadsheets/d/16Dn5pyBYoqCTGyaVS5_1dafeEufuFVUn7lz_4Q3MsFI/edit?usp=sharing) |  |  |  |
| UX thematic -analysis | 1 [Initial coding](https://docs.google.com/spreadsheets/d/1JmPz_Qg_n8Q5x-Ze6AmBM927AczSYmx2/edit?usp=sharing&ouid=103908336510759123757&rtpof=true&sd=true)  2 [Results and discussion for performance expectancy theme](https://docs.google.com/document/d/1GCc8bKJnMmTPZT7ax6uYwynt9DtzkOsS6xIfqJLTZ_w/edit#heading=h.el48eb22wpro) |  | <https://docs.google.com/document/d/1GCc8bKJnMmTPZT7ax6uYwynt9DtzkOsS6xIfqJLTZ_w/edit#heading=h.el48eb22wpro> effort expectancy, facilitating conditions | [Social influence + personal attitude](https://docs.google.com/document/d/1GCc8bKJnMmTPZT7ax6uYwynt9DtzkOsS6xIfqJLTZ_w/edit#heading=h.el48eb22wpro) |  |  |  |
| What are the requirements in terms of IT infrastructure? |  |  | [4.3 What are the requirements in terms of IT infrastructure?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.aazsam1t5atj) |  |  |  |  |
| Interoperability - How can the data be exported for research and other purposes? |  |  |  | [4.4 How can the data be exported for research and other purposes?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.aazsam1t5atj) |  |  |  |
| Will the data be accessible to non-radiologists? |  |  | [4.5 Will the data be accessible to non-radiologists?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.yx25n24xr5mz) |  |  |  |  |
| Are the AI model’s results interpretable? |  |  | [4.6 Are the AI model’s results interpretable?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.r3wg0yqvto4g) |  |  |  |  |
| Does the AI application comply with the local medical device regulations? | [✔](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.lj76shr1k7d4) |  |  |  |  |  |  |
| Does the AI application comply with the data protection regulations? | [✔](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.59m15i60rkxc) |  |  |  |  |  |  |
| What is the licensing model? |  |  | [6.1 What is the licensing model?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.em0hu54n594n)  last paragraph | [6.1 What is the licensing model?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.em0hu54n594n)  Everything except the last paragraph |  |  |  |
| How are user training and follow-up handled? |  |  |  | [6.2 How are user training and follow-up handled?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.em0hu54n594n) |  |  |  |
| How is the maintenance of the product ensured? |  |  | [6.3 How is the maintenance of the product ensured?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.5cv0fmxc13wa) |  |  |  |  |
| How will potential malfunctions or erroneous results be handled? |  |  | [6.4 How will potential malfunctions or erroneous results be handled?](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.c1mn568fgtrt) |  |  |  |  |
| Summary of areas of evaluation with individual recommendations | [7.1.4 Regulatory and legal aspects](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.shjxmcn7xabg) |  | [7.1.3 Usability and Integration](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.l15cgxi210tq) |  | [7.1.2 and recommendations of 7.1.1](https://docs.google.com/document/u/0/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit) |  | [7.1.1 Relevance](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.d2anqiyw52pp) |
| Overall recommendation |  | [7.2 recommendation](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#) |  |  |  |  |  |

Table of Non-Technical Contributions

| **Name of Task** | **Bronte Lee** | **Zichen Li** | **Cailin Feldman** | **Jingbang Men** | **Kiran Papalkar** | **Baoxi Liu** | **Zheyuan Zhang** |
| --- | --- | --- | --- | --- | --- | --- | --- |
| XP Summary | [Completed XP roles section](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/wiki/XP%20Summary) | [Wrote 12 core practices](https://docs.google.com/document/d/1Qz_7_luCl_EE9qLvzWbDASXyJEwoCXc0_OuyLAuSYHA/edit) |  |  | [Iteration Summary](https://docs.google.com/document/u/0/d/1Qz_7_luCl_EE9qLvzWbDASXyJEwoCXc0_OuyLAuSYHA/edit) |  |  |
| Scope | [For version 1, I edited and elaborated on most sections, particularly deliverables and scope. For the revised version I re-wrote the background and aim sections](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Scope%20Statement.pdf) | [objectives and success criteria](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Scope%20Statement.pdf) |  | [First draft](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Scope%20Statement.pdf) | [Contributed to most sections, wrote human resources and other resources. Helped edit whole doc for version 2](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Scope%20Statement.pdf) |  |  |
| 1st Presentation | [Expected outcomes and deliverables slide](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Week%203%20Group%202%20Slide%20Deck.pdf) | [Member responsibilities slide](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Week%203%20Group%202%20Slide%20Deck.pdf) | [Created slide deck](https://docs.google.com/presentation/u/0/d/1mK0g73wGf2sa7chJEEwdXjU4YzEnP2NzN3lbsPK59XM/edit) | [Constructed the project timeline.](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Week%203%20Group%202%20Slide%20Deck.pdf) | [Intro Slide](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Week%203%20Group%202%20Slide%20Deck.pdf) | [project scope](https://docs.google.com/presentation/d/1mK0g73wGf2sa7chJEEwdXjU4YzEnP2NzN3lbsPK59XM/edit#slide=id.g1456fc6c867_0_210) | [project scope](https://docs.google.com/presentation/d/1mK0g73wGf2sa7chJEEwdXjU4YzEnP2NzN3lbsPK59XM/edit#slide=id.g1456fc6c867_0_210) |
| Initial Plan | [✔](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Initial-Plan.md) |  |  |  |  |  |  |
| README File | [✔](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/README.md) |  |  |  |  |  |  |
| Progressive Reports | [Description and scope sections for all reports](https://drive.google.com/drive/folders/1wmhRfFS-GCS4xLa0cFBfLqllJcOt0VUA) | [Achievements and major deliverables for all 3 progressive reports](https://drive.google.com/drive/u/0/folders/1wmhRfFS-GCS4xLa0cFBfLqllJcOt0VUA) | Quality sections | [Planned activities and major issues.](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Progressive%20Report.pdf) | [Risks and external dependencies in week 5](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/Progressive%20Report.pdf)  [External dependencies and overall status in week 9 and 11](https://drive.google.com/drive/u/0/folders/1wmhRfFS-GCS4xLa0cFBfLqllJcOt0VUA) | [Effort for the week 5](https://docs.google.com/document/d/1Qf7r3OnWmtWTlbrXaqJbOWH4dlOWt3X1/edit)  [RIsk and effort for week 9 and 11](https://drive.google.com/drive/u/0/folders/1wmhRfFS-GCS4xLa0cFBfLqllJcOt0VUA) | [Planned activities for week 5, Time for week 9 and 11](https://drive.google.com/drive/u/0/folders/1wmhRfFS-GCS4xLa0cFBfLqllJcOt0VUA) |
| First Group Report | [3 Evaluation](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.pd60i6w4xjr) | [Tools to build the system](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit) | [Executive Summary](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.bq44qy9r3z7e)  [1.1](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.smb0kwj703qx)  [1.2](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.bsfkyvpsiabt) [1.3](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.e2zt0957hued) [1.4](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.rnod3gr3wsix) [1.5](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.oar4gf350c1a)  [1.6](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.qav5g5od0exv) | [6.3](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.qcbhin75he9k)  [6.4](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.u8u2evoqissw)  [6.5](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.tzh9zapxybbx)  [7.1](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.77obe56hea4b)  [7.2](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.hh15t66n0bby)  [7.3](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.zc101frcr9dz) | [2 Effort section Overview of system from user view](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.xix9dn1c6l6n) | [6.1](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.77obe56hea4b)  [6.2](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.77obe56hea4b)  [7.2](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.77obe56hea4b)  [7.4](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.77obe56hea4b) | [4 System structure overview](https://docs.google.com/document/d/1q0sRp2ZpBMlMjl53FUaUgykhQ0brlNBl/edit#heading=h.6rcsiyk1922t) |
| User Stories |  | [✔](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report-Files/User%20Stories.md) |  |  | [Wrote first seven user stories and joint contribution to remaining stories](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Report%20Documents/User%20Stories.pdf) [Compiled user stories into formatted table](#_heading=h.leghe36c926s) | [Write some of the user stories](https://docs.google.com/document/d/15CzlLgdh9CZgOBHP4IaoLiQ9_B5y5WA4Cxy-E9gqyQA/edit) | [Write some of the user stories](https://docs.google.com/document/d/15CzlLgdh9CZgOBHP4IaoLiQ9_B5y5WA4Cxy-E9gqyQA/edit) |
| Meeting Minutes | [Week 2](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-2/), [week 4 client meeting](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-4/client-meeting-week-4.md), [week 5 tutor meeting](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-5/tutor-meeting-5.md), [week 5 tutorial meeting ,](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-5/tutorial-meeting-5.md) [week 6 client meeting](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-6/client-meeting-week-6.md), [week 10 managers meeting minutes](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-10/week10managersmeetingFri.md) | [Week3 meeting minutes](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-3/)  [Week7 meeting minutes](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-7/)  [Week11 meeting minutes](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-11/) |  | [Week4](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-4/week4groupmeetingFriday.txt)  [group meeting](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-4/week4groupmeetingFriday.txt)  [Week4 tutorial meeting](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-4/week4tutorialmeeting.txt)  [Week5 group meeting](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/Week-5/Fridaygroupmeeting.txt)  [All the Week 10 meetings](https://drive.google.com/drive/folders/1p_sQTsjJSHyghTlvJBU4FHKMGr4URE20), 4 meetings |  |  | [Week9 tutorial meeting](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/commits/52818f54f42e919655a49d400f58f62c889db97a)  [Week9 group meeting](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/commits/299863033a89f0026bca328763a520e62bd5c034) |
| First group project demo | [Slides 11, 12, 15, 17, 18](https://docs.google.com/presentation/u/0/d/1rpqTaLHact1AOirPYDEqwEb0E-HSNnX8uiuYecP-lHE/edit) | [Demonstration of qXR vs InferRead DR](https://docs.google.com/presentation/d/1rpqTaLHact1AOirPYDEqwEb0E-HSNnX8uiuYecP-lHE/edit) | <https://docs.google.com/presentation/d/1rpqTaLHact1AOirPYDEqwEb0E-HSNnX8uiuYecP-lHE/edit?usp=sharing> slide deck design, intro and requirements slide | [Project pipeline + Evaluating UX for Other Medical Imaging AI Products](https://docs.google.com/presentation/d/1rpqTaLHact1AOirPYDEqwEb0E-HSNnX8uiuYecP-lHE/edit?usp=sharing) | [User stories, technical article gathering, technical article summary, contributed to excluded products](https://docs.google.com/presentation/d/1rpqTaLHact1AOirPYDEqwEb0E-HSNnX8uiuYecP-lHE/edit?usp=sharing) | [Slides 8(qXR vs Lunit vs VUNO), 10(Excluded Products)](https://docs.google.com/presentation/d/1rpqTaLHact1AOirPYDEqwEb0E-HSNnX8uiuYecP-lHE/edit#slide=id.g11ac1abd191_0_0) | [Demonstration of qXR vs ChestEye](https://docs.google.com/presentation/d/1rpqTaLHact1AOirPYDEqwEb0E-HSNnX8uiuYecP-lHE/edit#slide=id.g15351e7af24_2_0)  [Next Steps](https://docs.google.com/presentation/d/1rpqTaLHact1AOirPYDEqwEb0E-HSNnX8uiuYecP-lHE/edit#slide=id.g15351e7af24_2_0) |
| Second presentation | [Quality of work slide](https://docs.google.com/presentation/d/1QKzF1JPhqQd9GyI5PBYnrRoD06RLYM57Rc0aRnZK8KM/edit?usp=sharing) | [Demonstration](https://docs.google.com/presentation/d/1QKzF1JPhqQd9GyI5PBYnrRoD06RLYM57Rc0aRnZK8KM/edit#slide=id.g15024f5e6b8_1_0) | [Slide deck design](https://docs.google.com/presentation/d/1QKzF1JPhqQd9GyI5PBYnrRoD06RLYM57Rc0aRnZK8KM/edit?usp=sharing)  [intro](https://docs.google.com/presentation/d/1QKzF1JPhqQd9GyI5PBYnrRoD06RLYM57Rc0aRnZK8KM/edit?usp=sharing) | [Group process](https://docs.google.com/presentation/d/1CaUGvRq_3NjSAG6SgAUOFc-M1hGC-uDNIAhPYtBrBGw/edit#slide=id.g14b5028479a_2_0) | [System Specification Slide](https://docs.google.com/presentation/d/1CaUGvRq_3NjSAG6SgAUOFc-M1hGC-uDNIAhPYtBrBGw/edit#slide=id.g14b5028479a_2_0) | [Discipline](https://docs.google.com/presentation/d/1QKzF1JPhqQd9GyI5PBYnrRoD06RLYM57Rc0aRnZK8KM/edit#slide=id.g14b5138ff0d_0_52) | [Plan for Remaining Requirements/User Stories Potential Project Risks](https://docs.google.com/presentation/d/1CaUGvRq_3NjSAG6SgAUOFc-M1hGC-uDNIAhPYtBrBGw/edit#slide=id.g14b2df86c4a_0_20) |
| Final Tutor presentation | [Quality and testing slides and introduction and ECLAIR summary for demonstration](https://docs.google.com/presentation/d/1TAHQtbPaf5O9o3q0RclxCmzbLjNPe1CFLkST8WLIL_g/edit?usp=sharing) | [Response to key changes](https://docs.google.com/presentation/d/1TAHQtbPaf5O9o3q0RclxCmzbLjNPe1CFLkST8WLIL_g/edit#slide=id.g11ac1abd191_0_0) | <https://docs.google.com/presentation/d/1TAHQtbPaf5O9o3q0RclxCmzbLjNPe1CFLkST8WLIL_g/edit?usp=sharing> - overview and slide deck design | [UX demonstration model part + quality of group process](https://docs.google.com/presentation/d/1TAHQtbPaf5O9o3q0RclxCmzbLjNPe1CFLkST8WLIL_g/edit?usp=sharing) | [Tech demonstration + System Specification](https://docs.google.com/presentation/d/1TAHQtbPaf5O9o3q0RclxCmzbLjNPe1CFLkST8WLIL_g/edit?usp=sharing) | [discipline and tools](https://docs.google.com/presentation/d/1TAHQtbPaf5O9o3q0RclxCmzbLjNPe1CFLkST8WLIL_g/edit#slide=id.g11ac1abd191_0_50) | [System Architecture and Design](https://docs.google.com/presentation/d/1TAHQtbPaf5O9o3q0RclxCmzbLjNPe1CFLkST8WLIL_g/edit#slide=id.g122cf293abc_0_54)  [Final recommendation](https://docs.google.com/presentation/d/1TAHQtbPaf5O9o3q0RclxCmzbLjNPe1CFLkST8WLIL_g/edit#slide=id.g122cf293abc_0_54) |
| Final client presentation | [Testing Slides, 3.2.1, 3.2.2, 3.2.3, regulatory and legal aspects slides](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit?usp=sharing) | [1.2 + 2.3.1 + 6.1](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_7878) | [made the slide deck design and did all my assigned eclair questions](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_2565) | [Introduction + 4.2.5 + 4.2.6 + 4.2.7 + 4.2.8 + 4.4 + 6.1 + 6.2](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_2565) | [Assigned eclair questions and 3.3.4 Heterogeneity](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_2565) | [2.1 Algorithm’s Design Specifications](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_4339)  [2.2 Algorithm Training](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_4339)  [2.3.4 Results and Discussion](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_4339)  [part of 3.3.5 + diagrams of 3.3.4](https://docs.google.com/document/d/1wMEKHrzlRuY-XHFqm4lb6Agv6A-yVkUJL-1puGev4xY/edit#heading=h.yiz4tlm02v3t) | [2.3 What are the risks associated with the use of the AI system?](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_2565)  [3.3.3 Systematic Review](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_2565)  [3.5 Is the algorithm fixed or adapting as new data comes in?](https://docs.google.com/presentation/d/1zn4O-M9rGN7O4YT96KSX70ptx5lbZ2TZrxuFuR4EVvk/edit#slide=id.g16fffef4640_0_2565) |
| Leading group meetings |  |  |  |  |  |  |  |
| Drafting up task and allocation and tracking progress |  |  | <https://docs.google.com/spreadsheets/d/1IjFLQp5BcPfdfJADUKgXEYQhPAJDeKwPJGPjLh763aY/edit?usp=sharing> |  |  |  |  |
| Final Group Report | [Sections 3, 2.4, 4.2.4, re-write of sections 2.5.3 and 2.5.4, major edit to 4.1 to clarify the research discipline,moderate edits to, 4.1.1 (changed some dot points to increase relevance), 4.1.4 (second paragraph), and general edits to sections 1, 4.2, Appendix, and References](https://docs.google.com/document/d/1JpVNF6F7MQNBZ7sxIElc_3WW_QTkc8X0/edit?usp=sharing&ouid=103908336510759123757&rtpof=true&sd=true) | [Section 6 reflections and conclusions](https://docs.google.com/document/d/1JpVNF6F7MQNBZ7sxIElc_3WW_QTkc8X0/edit#) | [1 Introduction](#_heading=h.njgp2uyxxd67) re-write of [5 Group Processes](#_heading=h.jtg1kyb1wxtq) and [6 Reflections and Conclusions](#_heading=h.sd705oembn5k) | [4.1.1+4.1.2+4.1.3+4.1.4+4.2.1+4.2.2+5+5.1+5.3+5.6+5.7+5.8](https://docs.google.com/document/d/1JpVNF6F7MQNBZ7sxIElc_3WW_QTkc8X0/edit#) | [Wrote Sections 2.1, 2.2 and 2.3, major rewrite of sections 2.5.1 and 2.5.2, major editing of sections 5 and 6](https://docs.google.com/document/d/1JpVNF6F7MQNBZ7sxIElc_3WW_QTkc8X0/edit#) | [4.1(Intro)+4.1.5+4.2(Intro)+4.2.1(Zoom)+4.2.3+5.2+5.5](https://docs.google.com/document/d/1JpVNF6F7MQNBZ7sxIElc_3WW_QTkc8X0/edit#) | [System Architecture and Design](https://docs.google.com/document/d/1JpVNF6F7MQNBZ7sxIElc_3WW_QTkc8X0/edit#) |

# Appendix

## A1. Location of Meeting Minutes and Client Communication

The meeting minutes are available in the group’s [Bitbucket](https://bitbucket.org/andersen_liu/isys3888_tu15_02_p56/src/master/).

## A2. Involved articles

[Articles](https://docs.google.com/document/d/1W8iuZQUIGnEWvmU8oMOaIRMLAHrjquiFD5JuZYmemcU/edit) identified by tech team.

[Articles](https://docs.google.com/spreadsheets/d/1RhzX0JjUH6D0ZRoSSnJewTDqgtng03jA3h9OQ1VyBVI/edit#gid=0) identified by UX team.

## A3. Systematic Review Protocols

[UX Systematic Review Protocol](https://docs.google.com/document/d/1fvetkO7wzLRR7s-9ukJXNDwvOyJ5Wf3UZVyRBuy1Jz0/edit#)

[Technical Systematic Review](https://docs.google.com/document/d/1oJnamNJiI4bGTQpt8Bg1_v500bqte6NK5NP-_wd2I8k/edit#)

[qXR Systematic Review Protocol](https://docs.google.com/document/d/10bP45WN_cmBoHZwW8HTuMwZA_uikEy_RQw8eqWT-X9c/edit?usp=drive_web&ouid=110701808833773385280)

## A4. Academic Websites

<https://pubmed.ncbi.nlm.nih.gov/>

<https://www.proquest.com/?accountid=14757>

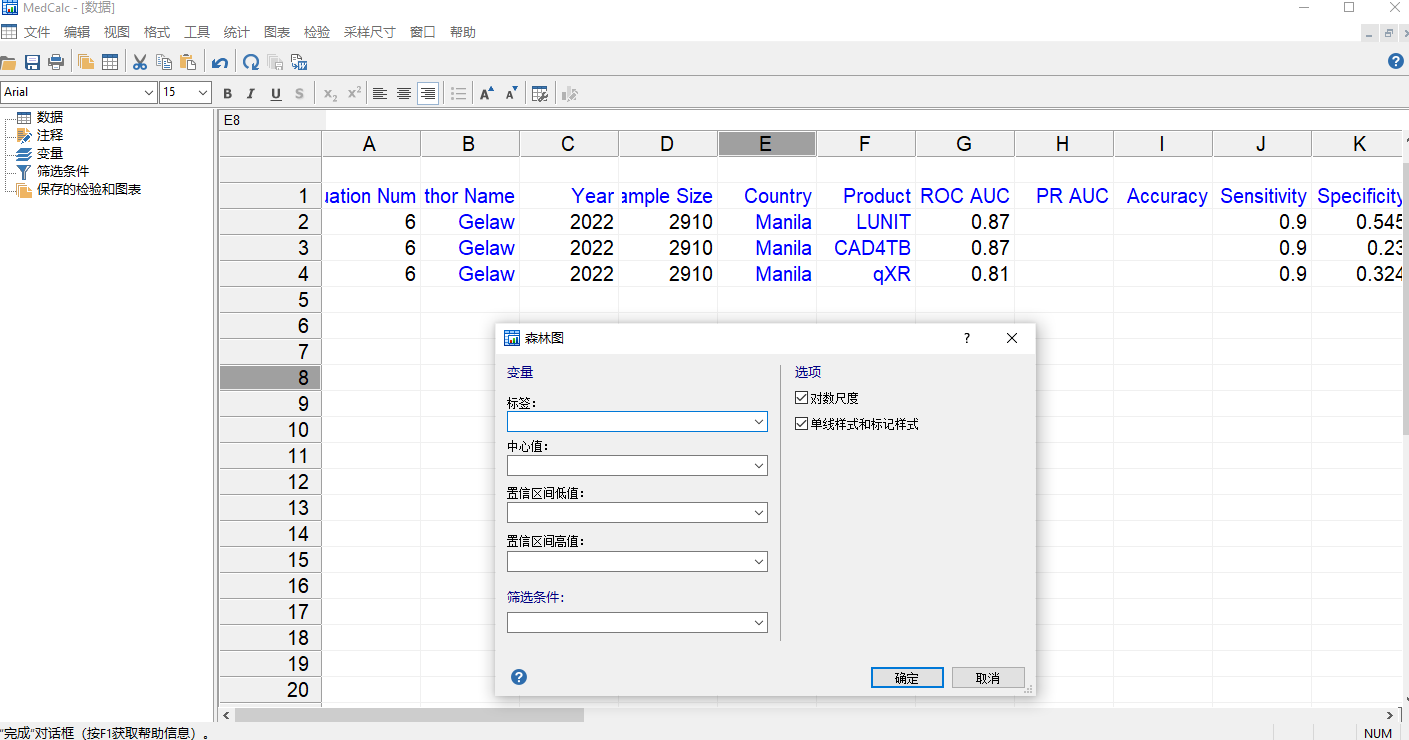
<https://www.thelancet.com/>

<https://www.sciencedirect.com/>

## A5. Meta Analysis Tools

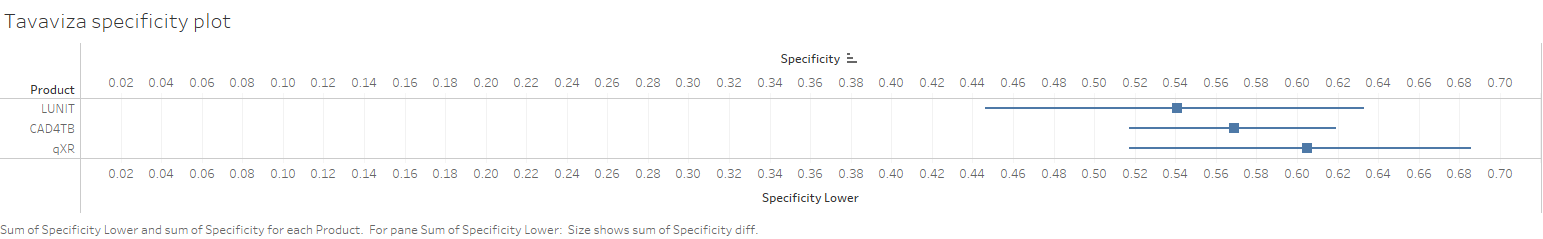
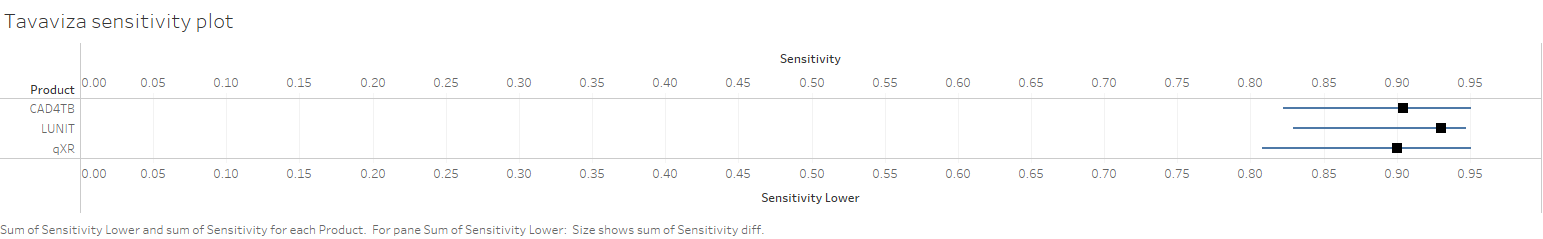
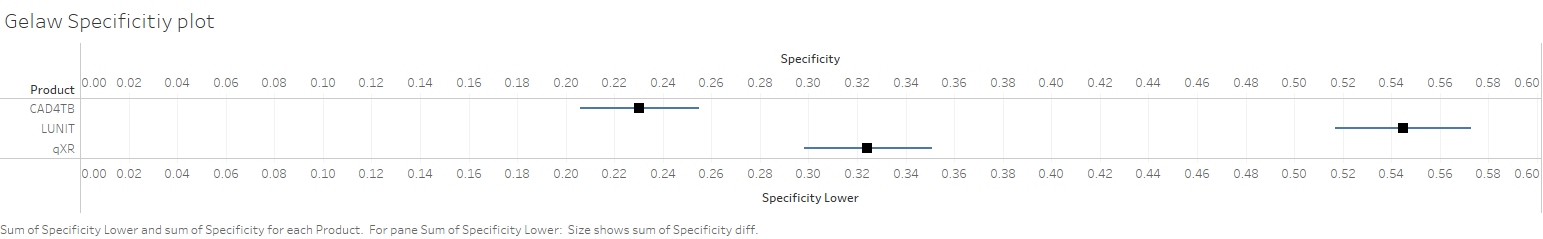
# 

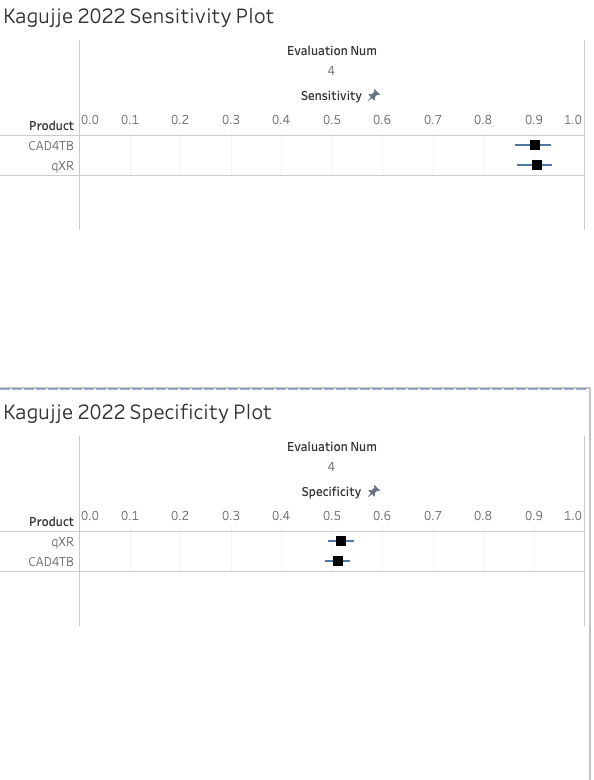
(Tableau)

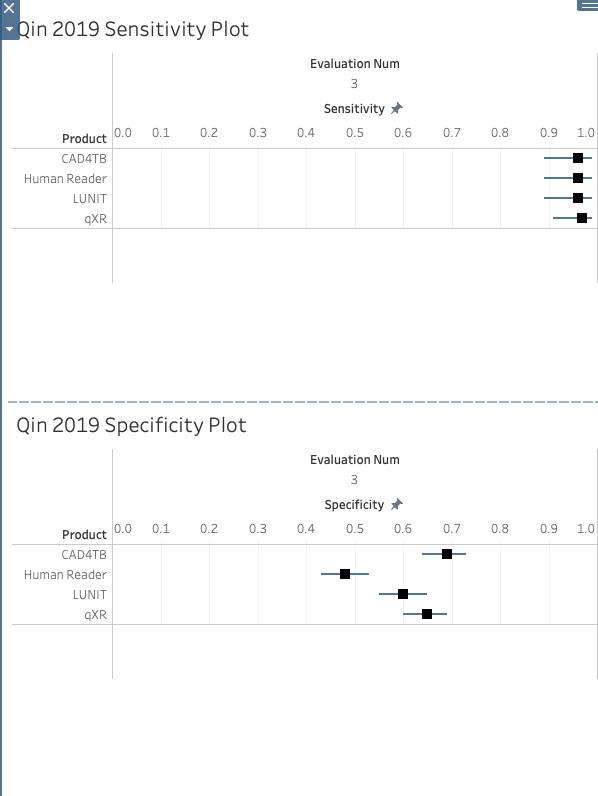


(MedCalc)

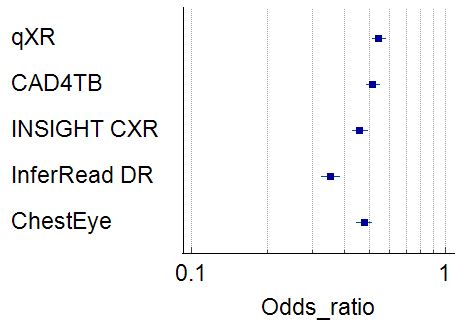
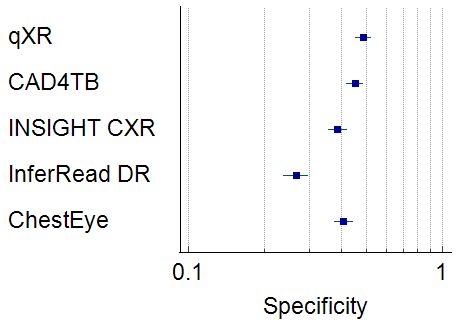
## A6. Forest Plots



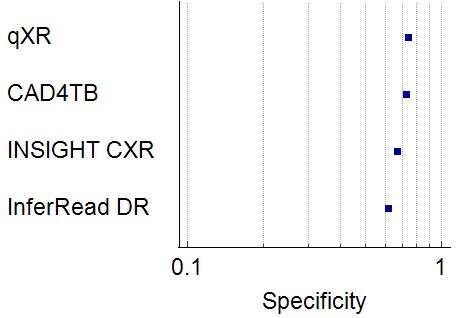




Zhi Zhen Qin 2021

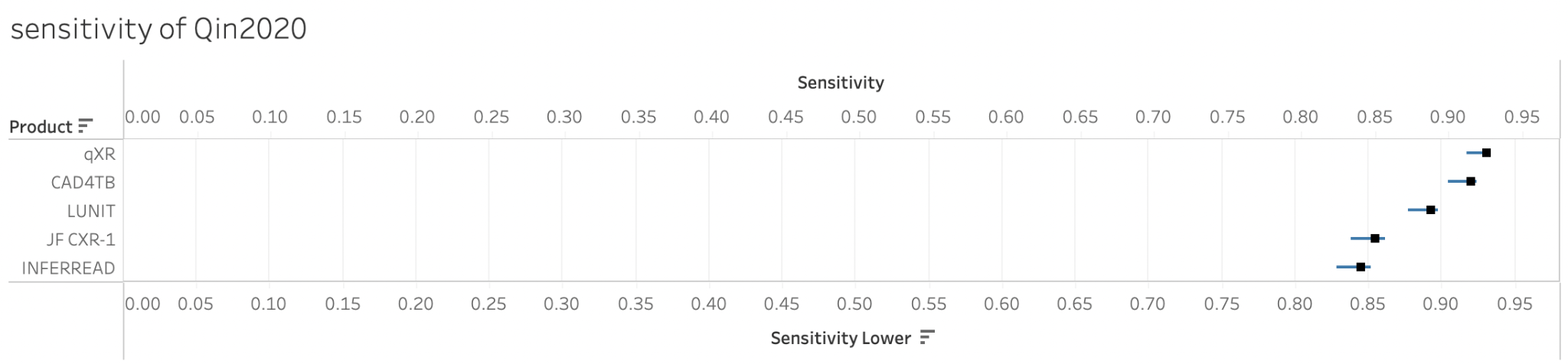


Andrew J

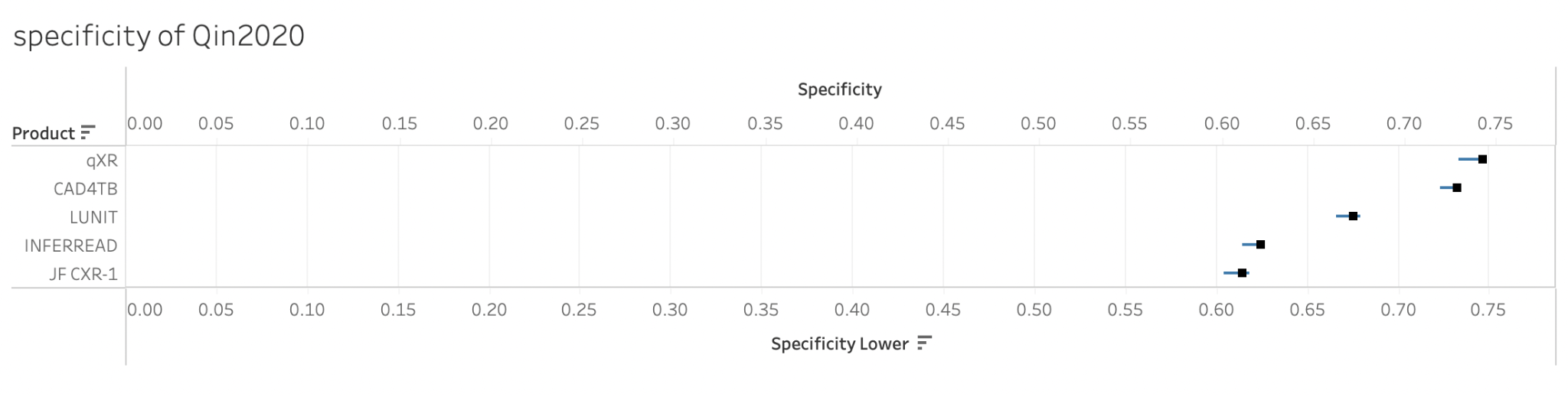


**Qin, 2020**

**Sensitivity with 70% specificity**

****

**Specificity with 90% sensitivity**

****

# 

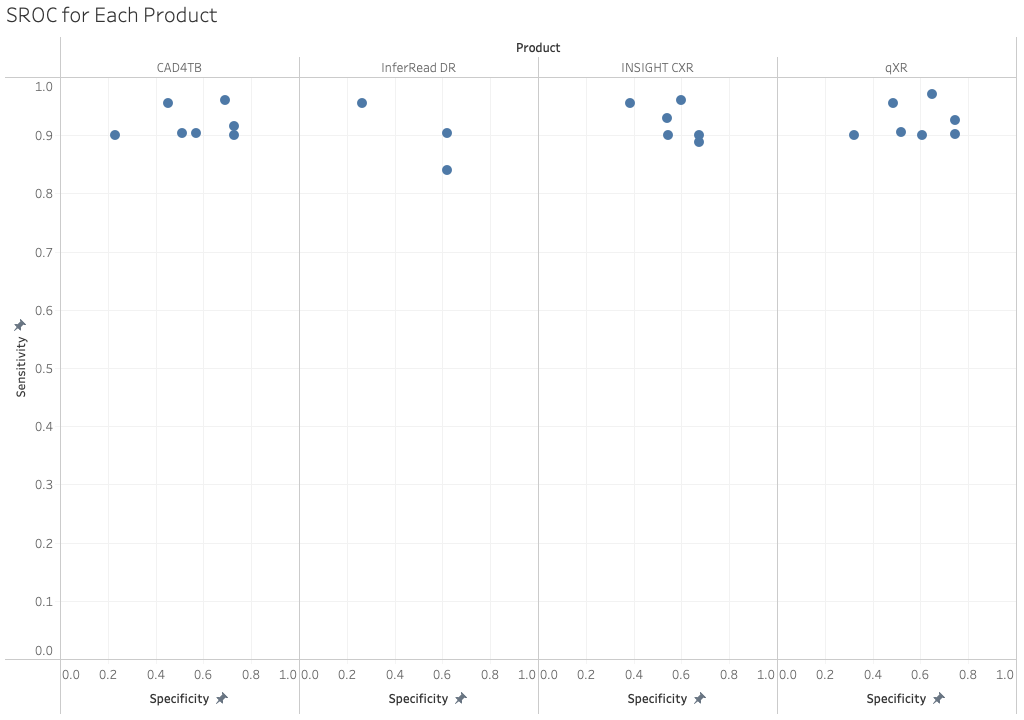
# 

# 

# 

# 

## A7. SROC Graph



# 

# 

# 

# 

# 

# 

# 

# References

Deeks JJ, Higgins JPT, Altman DG (editors). (2022). Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). Cochrane. Available from <https://training.cochrane.org/handbook/current/chapter-10>.

Ekelund, S. (2012). ROC Curves—What are They and How are They Used? Point of Care: The Journal of Near-Patient Testing & Technology, 11(1), 16–21. <https://doi.org/10.1097/poc.0b013e318246a642>

Ganeshkumar, P., & Gopalakrishnan, S. (2013). Systematic reviews and meta-analysis: Understanding the best evidence in Primary Healthcare. *Journal of Family Medicine and Primary Care*, *2*(1), 9. <https://doi.org/10.4103/2249-4863.109934>

Haidich, A. B. (2010, December). Meta-analysis in Medical Research. Hippokratia. Retrieved October 15, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3049418/>

Israel, H., & Richter, R. R. (2011). A guide to understanding meta-analysis. *Journal of Orthopaedic & Sports Physical Therapy*, *41*(7), 496–504. <https://doi.org/10.2519/jospt.2011.3333>

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: Amee Guide no. 131. *Medical Teacher*, *42*(8), 846–854. <https://doi.org/10.1080/0142159x.2020.1755030>

*Libguides: Creating a prisma flow diagram: Prisma 2020*. PRISMA 2020 - Creating a PRISMA flow diagram - LibGuides at University of North Carolina at Chapel Hill. (n.d.). Retrieved September 4, 2022, from <https://guides.lib.unc.edu/prisma>

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis. *International Journal of Qualitative Methods,* *16*(1), 160940691773384. <https://doi.org/10.1177/1609406917733847>

Omoumi, P., Ducarouge, A., Tournier, A., Harvey, H., Kahn, C. E., Louvet-de Verchère, F., . . . Richiardi, J. (2021). To buy or not to buy—evaluating commercial AI solutions in Radiology (the Eclair guidelines). *European Radiology,* *31*(6), 3786-3796. <https://doi.org/10.1007/s00330-020-07684-x>

Šimundić A. M. (2009). Measures of Diagnostic Accuracy: Basic Definitions. *EJIFCC*, *19*(4), 203–211.

Swift, A., Heale, R., & Twycross, A. (2019). What are sensitivity and specificity? Evidence Based Nursing, 23(1), 2–4. <https://doi.org/10.1136/ebnurs-2019-103225>