TRITECH
Sefydliad | Institute

Hywel Dda University Health Board

Evaluation report
Multimodal AI analysis of Prostate Cancer indicators to reduce patient backlogs and improve patient care.

Report Produced on 15th May 2023

Report prepared by:

Professor Chris Hopkins, Consultant Clinical Scientist & Head of Innovation & the Tritech Institute, Hywel Dda University Health Board

With support from:

Dr Rachel Gemine, Deputy Head, Mr Billy Woods, Clinical Scientist, Professor Keir Lewis, Clinical Lead, Tritech, Hywel Dda University Health Board; Dr Manish Patel, Dr Andrew Thompson, Dr Richard Hammersley & Dr Esin Karahan from Jiva.ai

Clinical lead(s):

Mr Sohail Moosa, Consultant Urologist, Hywel Dda University Health Board

Evaluation Lead:

Professor Chris Hopkins, Consultant Clinical Scientist & Head of Innovation & the Tritech Institute, Hywel Dda University Health Board

Funded by: Moondance Cancer Initiative, Innovation Time Award 2021

Who We Are

In 2021 the Tritech Institute was launched. We are a team based in a bespoke facility within the Hywel Dda University Health Board comprising of industry-leading engineers, scientists and clinicians.

Our Institute

Here at TriTech Institute, we support thedevelopment of healthcare solutions on a local, national, and global level offering designers and manufacturers a single point of access to the NHS through a collaborative and agile approach.

What We Offer

The team's advanced skills in clinical and research design are combined with technical engineering expertise to manage the whole innovative pathway from early unmet need, through to concept design, prototyping, clinical testing, and real-world service evaluations.

Our Services

We provide specific services and solutions for clinical engineering, research and innovation and Value-Based healthcare and can also support with grant writing and submission.

Executive summary

Prostate cancer affects 1 in 8 men throughout their lifetime. Due to increasing longevity and increased awareness the incidence and prevalence of prostate cancer is increasing. Prostate cancer is predominantly diagnosed using Magnetic Resonance Imaging (MRI) scans; however these scans require specialist interpretation and timely reporting. A lack of radiologists and particularly urology specialist radiologists can be a limiting factor, especially as demand grows lead to delays in the diagnostic pathway.

An artificial intelligence (AI)/machine learning (ML) based MRI diagnostic aid for prostate cancer may support clinical decision making and reduce time to interpret MRI. JivaRDX (a class IIa medical device, pending MHRA approval) is a radiology-facing application that predicts the presence of cancerous tissue from prostate MRI scans, and is intended for use as a diagnostic aid. Operationally, JivaRDX can integrate into the radiology workflow non-disruptively by automatically annotating imaging files and therefore requires minimal intervention and training. Jiva have previously demonstrated a proof-of-concept achieving detection and localisation of prostate cancer from MRI scans (87% sensitivity, 67% specificity); bone, tissue and organs differentiated with 96.8% specificity. It has been found to perform within reported MRI diagnostic accuracy in the clinic (58-96% sensitivity, 23-87% specificity).

The JivaRDX AI/ML-based MRI platform was evaluated over eighteen months across all 4 acute hospital sites within Hywel Dda University Health Board (HDUHB).

As part of the current evaluation, we retrospectively analysed scans taken of 121 patients with suspected prostate cancer. These anonymised scans, combined with levels of blood prostate specific antigen (PSA) (a known biomarker for prostate cancer) and patient age, were used to create a feasibility demonstration of the JivaRDX platform as a multimodal predictor of the presence of the disease. The initial evaluation could prepare for clinical pilot readiness and quantify early Value-Based healthcare impact as well as diagnostic accuracy.

For this evaluation, we used several sources of qualitative and quantitative data. We aimed to evaluate three main areas:

- 1.**Technology evaluation**, to test the accuracy of the multimodal model and any early positive or adverse impact of AI technology in situ.
- 2.**Value analysis** with Health Technology Wales to determine the potential value of this technology through completion of the NICE META gap analysis tool
- 3.**Clinical and patient perspectives** through focus groups to determine acceptability of using JivaRDX in diagnostic pathways

High Level Outcome

Technology evaluation

In conclusion the JivaRDX AI/ML based MRI model was found to provide 77% sensitivity, 65% specificity & 69% accuracy in detecting prostate cancer. Multiple studies have shown Radiologist specificity at 57%. Whilst these results are extremely promising, further analysis is required before JivaRDX is moved into routine clinical care within Hywel Dda.

Clinical and patient perspectives

Eleven out of fifteen MDT staff responded and all eleven MDT members had a positive opinion of the Jiva RDX MRI diagnostic aid. All highlighted the positive impact the AI/ML model could have on patient safety, outcomes, teamwork, communication and efficiency.

Only three out of twenty patients responded to the questionnaire but all three patients had a positive opinion of the AI/ML. In general, our patients reported enthusiasm on the ability of AI/ML to be a positive influence in medicine. They felt healthcare AI/ML was a positive step forward and those patients who understood the concept of AI were supportive of developing AI tools for a variety of different healthcare applications.

Value analysis

The value case results undertaken by Health Technology Wales suggest that JivaRDX is less costly than standard care.

This report presents the findings of the evaluation, which covers the period 4th November 2021 to 15th May 2023. Based on this evaluation, several key recommendations are made:

Recommendations

Recommendation 1: [Improve sensitivity & specificity of AI MRI outcomes]

The future of clinically successful healthcare AI relies on robust accuracy. Increasing specificity from 90% to 95% amounts to cutting false positives (and false alerts) by two-fold. Jiva.ai algorithms must feature both high sensitivity AND high specificity in a real-world clinical and radiological setting.

Recommendation 2: [Regulatory approval]

Clinical investigation to be submitted to MHRA for the study 'Jiva.ai MRI validation of JivaRDX for Prostate Cancer' for the company to seek regulatory approval across the UK.

Recommendation 3: [Integration of multimodal AI]

Examine the integration opportunity of using JivaRDX with the Fuji REiLI (artificial intelligence (AI) enablement) and Synapse (Image analysis) platforms to streamline data integration and flow.

Recommendation 4: [Improve patient understanding of AI]

The patients' view on the implementation of AI in radiology is still mainly unexplored territory. Successful implementation of AI in radiology requires the assessment of our patient views towards the technology.

Recommendation 5: [Account for Data variability in a real-world clinical environment]

The 121 cases from Hywel Dda were sourced from different clinical sites which have used a variety of MRI imaging devices and protocols. Data variability is therefore a real-world issue, and any deployed AI tool needs to take this into account. Jiva should undertake a retrospective multicenter study.

The aim of the next larger study should also compare performance of the software against an independent radiology expert, e.g., showing that JivaRDX is not worse or better than standard care in terms accuracy of detecting clinically relevant lesions in MRI scans performed for prostate cancer.

Table of Contents

Index of tables and Figures

Abbreviations

3T MRI	Tesla (T) is the unit of measurement indicating the strength of a magnetic field. A 3T (Telsa) MRI is twice as strong as a standard 1.5T MRI
AS	Active Surveillance
AI	Artificial Intelligence
DICOM	Digital Imaging and Communications in Medicine
DPIA	Data Protection Impact Assessment
EPR	Electronic Patient Record
HDUHB	Hywel Dda University Health Board
ML	Machine Learning
MRI	Magnetic Resonance Imaging
mpMRI	Multiparametric MRI
META	Medtech Early Technical Assessment
NICE	National Institute for Clinical Excellence
NHS	National Health Service
PCa	Prostate Cancer
PSA	Prostate-specific antigen
PACS	Picture Archive and Communication System
R&I	Research and Innovation
SCP	Single Cancer Pathway

Acknowledgements

A note of sincere thanks to all of the Urology multi-disciplinary team members in Hywel Dda and Swansea Bay University Health Boards, Mr Sohail Moosa, Consultant Urologist & Clinical lead for the study, Mrs Nina Ralph (Radiology) and Mr Dan Owen (Cyber Security) for their support with image capture and Dr Manish Patel, Dr Andrew Thompson, Dr Richard Hammersley, Dr Esin Karahan and Dr Samir Khan from Jiva.ai for their time and contributions to this evaluation. We would also like to thank all our patients for providing information and feedback through interviews & surveys. We would like to acknowledge Health Technology Wales who contributed to the NICE META data collection and our thanks also to the Moondance Cancer Initiative for funding the evaluation and to Dr Megan Mathias MBE & Wendy Evans for their unwavering support.

Introduction

Prostate cancer background and context

In the UK over 11,500 men die of prostate cancer (PCa) every year and the disease continues to be a top three cause of male cancer-related death [Prostate Cancer UK, 2022] . It is the most common cancer in males with incidence forecast to increase from 56,780 cases in 2020 to 66,639 cases (+17%) by 2030. Moreover, prostate cancer mortality in the UK is forecast to increase from 13,168 to 17,116 deaths (+30%) over the same 10-year period [Global cancer observatory figures].

Earlier detection/diagnosis of cancer can improve clinical outcomes [NHS England National Cancer Strategy 2015-2020; CRUK]. The typical treatment pathway involves a GP referral followed by a biopsy - a procedure that is both intrusive and uncomfortable. It is also known to exhibit a high incidence of infection that can spiral into serious complications. In May 2019, the UK National Institute for Health and Care Excellence recommended multiparametric MRI (mpMRI) as the standard first-line investigation for suspected clinically localised prostate cancer, shifting the emphasis on improving diagnosis towards radiology. NICE guidelines stipulate that at least 92.5% of all suspicious cases must have an MRI scan.

The clinical unmet need in the PCa diagnostic pathway is exemplified in the landmark [PROMIS,2017] and [ProtecT,2020] studies. Current NICE guidelines [NICE guidance NG131, 2019] stipulate the preferred clinical pathway which is a blood test for PSA followed by MRI scan before determination of biopsy.

Radiology issues

Human subjectivity in interpreting scans has been problematic. Multiple studies have shown low specificity (57%) in diagnosing prostate cancer by mpMRI, which can lead to unnecessary biopsies. Furthermore, difficulties in clinical care pathways are increasingly affected by an increasing shortage of qualified personnel to analyse scans effectively. The 5% increase in UK consultant radiologist headcount compares to a 26% increase in MRI scans alone over 6 years ago [Royal College of Radiologists (RCR) Annual report 2015-2016]. Additional demand for mpMRI will place further strain on the workforce and exacerbate human error – the current misdiagnosis rate is ~15-30% [Brady 2017 Insights into Imaging]. In the UK over 306,000 scans wait more than six weeks to be processed [NHS Constitution for England, 2021]. To plug the gap more than £165M is pumped into external referrals. Importantly, the vast majority of litigations in prostate cancer misdiagnosis related to the timeliness of diagnosis delivery. The COVID-19 pandemic has also led to backlogs in all scanning and cancer diagnosis [Health and Social Care Committee, 2021]. It is estimated that there are up to 2300 undiagnosed cases of prostate cancer per week in 2020 [Cancer research UK, 2020]. Therefore speeding but the accurate reporting of MRI scans should help achieve a timelier and true diagnosis for many people and hopefully lead to better clinical outcomes.

1. Risk of complications:
   The reported rate of over-diagnosis [Over diagnosis and overtreatment of prostate cancer, 2014] is 67% (approx. 4000 patients per year in Wales). This leads to unnecessary biopsies, from which 80% of patients suffer at least one complication (e.g., rectal bleeding, erectile dysfunction), with 1.25% of these complications becoming life threatening (e.g., bacterial infection leading to sepsis).

2. Cost effective models for targeting severe PCa:
   There is currently a lack of cost-effective methods available to reduce the uncertainty in the active surveillance (AS) of PCa patients, confidently predicting disease progression and drug performance, or enabling treatment switching for metastatic cancer, based on a robust understanding of patient demographics and diagnostic history [NICE guideline NG131, 2019]. Better prediction of these factors during AS or at subsequent decision stages could prevent the cost and human burden of radical interventions in PCa [Prediction models in cancer care, 2019]. Multimodal analysis using AI is a promising approach to enhance AS to the benefit of

both patients and healthcare providers.

On diagnosis itself three fundamental clinical issues remain:

- Accuracy: Misdiagnosis due to poor intra-radiologist specificity (ranging from 36% to 66%) and high subjectivity [A systematic review on multiparametric MR imaging in prostate cancer detection, 2017] lead to unnecessary biopsies, which cause downstream complications that have direct human and financial implications.

- Speed to decision: The current delay in diagnosis can stretch over 56 days [prostatecanceruk.org, 2017] in the UK, compared to 14 days for breast cancer. This situation will only worsen post-COVID; so early and fast diagnosis is crucial.

- Lack of multimodal analysis: Unification of modalities, such as PSA together with mpMRI, is widely accepted as a better indicator of disease [Artificial intelligence at the intersection of pathology and radiology in prostate cancer, 2019]. However, a clinically proven tool still does not exist for PCa diagnosis. JivaRDX hope to provide such a tool.

Jiva RDX background and context

JivaRDX is an artificial intelligence (AI)/machine learning (ML) based MRI diagnostic aid for PCa. Its intended application is as a novel diagnostic aid (class IIa medical device pending MHRA approval) to detect tumours. JivaRDX is a radiology-facing application that identifies cancerous tissue presence from prostate MRI scans. Operationally, JivaRDX can integrate into the radiology workflow non-disruptively by automatically annotating imaging files and therefore requires minimal intervention and training.

It is anticipated that JivaRDX will reduce over diagnosis and minimise the human and financial burden on healthcare systems. Addressing the key market requirements, JivaRDX is designed to increase diagnostic accuracy, increase speed to decision and empower clinicians to gain more insightful diagnoses. It will mean that patients receive faster diagnoses and commencement of treatment.

It has been developed to:

- Improve the accuracy of prostate cancer (PCa) detection and localisation.
- Improve patient experience and outcomes.
- Reduce pressures on radiologists.
- Improve clinical care pathway efficiency and reduce delays.

Jiva has a demonstrated a proof-of-concept achieving detection and localisation of PCa from 3T MRI scans (87% sensitivity, 67% specificity); Bone, tissue and organs differentiated with 96.8% specificity. It has been found to perform within reported mpMRI diagnostic accuracy in the clinic (58-96% sensitivity, 23-87% specificity), [Detection of clinically relevant prostate cancer from multi-parametric 3Tesla MRI scans using artificial intelligence, 2019].

Evaluation Introduction

In November 2021, the Tritech Institute within Hywel Dda University Health Board (HDUHB) was commissioned by Moondance Cancer to undertake a real-world clinical evaluation of JivaRDX, across all four Radiology acute sites within Hywel Dda University Health Board. The project was funded by Moondance Cancer.

Aim

The aim of this evaluation was to test for non-inferiority, in terms of sensitivity and specificity, of the software against independent clinical & radiology experts and quantify early diagnostic and Value-Based healthcare benefits. The evaluation seeks to address the three priority areas:

1. Diagnosis: The effectiveness of the Jiva.ai solution to diagnose PCa early more accurately in the clinical pathway (MRI scan).

2. Evaluate whether the Jiva.ai solution is scalable and specifically designed to accelerate radiological assessment and thereby speed treatment decisions.

3. Value-Based healthcare:
   to quantify the benefit of the Jiva.ai solution to tangible patient outcomes.

Methods

We built a data pipeline and acquired end-to-end data transmission in order to validate the machine learning model. The data collection and systems pathway is outlined below:

1. Patient data was anonymised at source from the Radiology system to provide anonymised patient studies (an example of the anonymisation can be seen in figure 1).

2. A local record was kept of each patient study to allow analysis of the images at the end of the project (see point 8 below).

3. Inclusion and exclusion criteria were applied in order to determine study participants.

4. Anonymised patient studies were shared with a Consultant Urologist in order to validate Likert scoring data (Table 1) and biopsy results (where available).

5. Anonymised patient studies were transferred by our cyber team to Jiva.ai via an encrypted file sharing platform.

6. Anonymised patient studies were passed through the JivaRDX ML platform.

7. The outcomes of the anonymized patients studies was passed back to the Health Board via the cyber team and encrypted file sharing platform.

8. Anonymised patient studies were de-anonymised.

9. Patient studies were shared with Mr Moosa, Consultant Urologist in order to clinically validate the outcomes of each patient study.

10. Sensitivity, Specificity and Accuracy results were shared with Jiva.ai after each iteration.

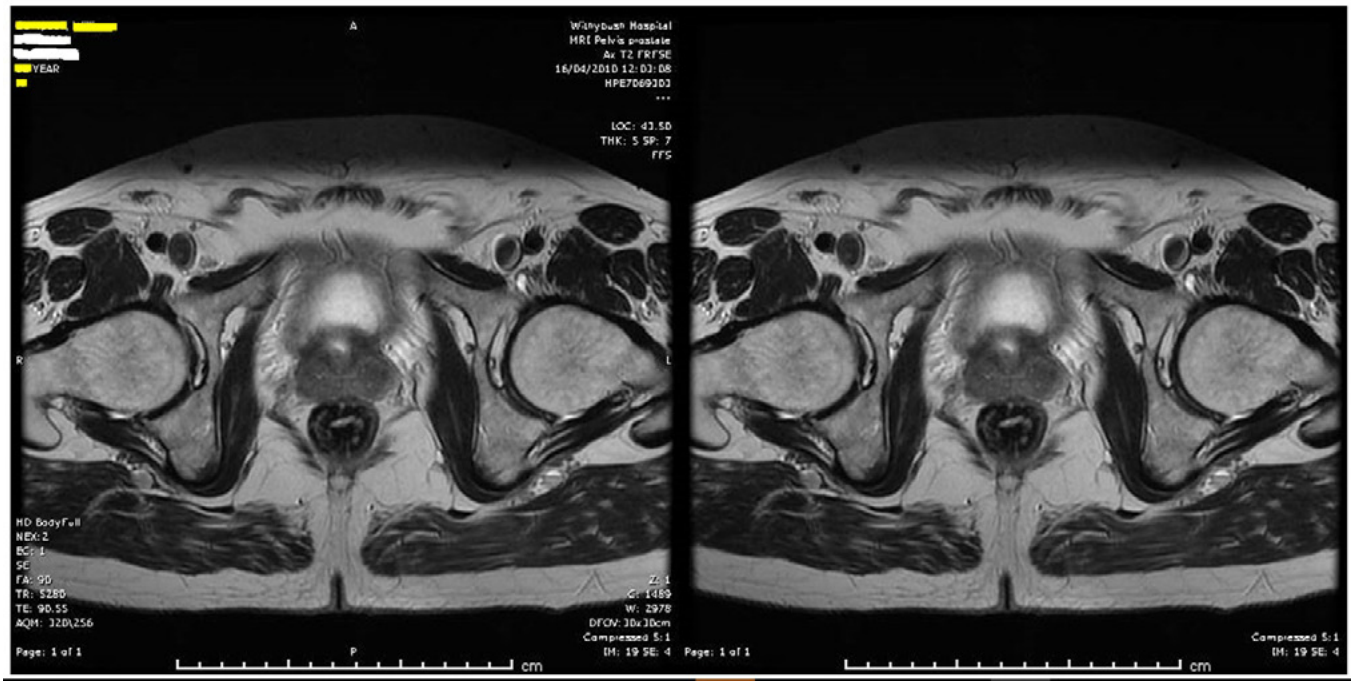


Figure 1: Example of an anonymized MRI scan image.

Procedure

As part of the current evaluation we retrospectively analysed scans taken from patients with suspected prostate cancer.

Sample size
The target number of patients from the study was 121.

Inclusion criteria
Participants were:

- Between 49 and 75 years of age.
- Have Likert scoring indexes available between 1-5 (table 1).
- Have been MRI scans during 2020 & 2021.

Exclusion criteria

Participants that:

1. Are outside of the age ranges of 49 to 75 years of age. (Most studies previously undertaken to assess the diagnostics used for prostate cancer, excluded patients outside of the age ranges of 50-70 years of age).

2. Do not have Likert scores available.

3. Have had MRI images taken before 2020. (Previous to 2020, Hywel Dda had different MRI protocols on each acute site).

Linkert score	Clinical significance
1	Cancer is highly unlikely to be present
2	Cancer is unlikely to be present
3	Caner is equivocal
4	Cancer is likely to be present
5	Cancer is highly likely to be present

Table 1 – Likert score for prostate cancer.

Risks and mitigation

Risks around data security and protection were discussed with all teams and were mitigated against (see table 2).

Risk	Mitigation
Meeting required accuracy standards (technical): For the solution to be of value to imaging providers it must at minimum, match or exceed the existing accuracy level of human interpretation (current standard of care).	Jiva has already undertaken testing that demonstrated a sensitivity and specificity above that of existing practice. In addition, we have a highly skilled and experienced team, and a thorough evaluation framework in place.
There is no indicative signal that improves diagnostic accuracy by adding other factors such as PSA and age. It is possible that the lack of data does not yield a significant signal.	A negative signal is still a positive, publishable result that bears clinical value. The clinical and Jiva teams will be actively sourcing exogenous data sources in parallel that could be used to augment the data in the case that lack of data is the cause of lack of signal.
Cybersecurity standards (technical): it is imperative that patient confidentiality is maintained and the solution is kept secure.	Jiva have built certified, secure and tested solutions to transfer data as required for processing and AI analysis. A Full DPIA and Cyber security assessment has been undertaken by the health board.

Risk	Mitigation
PACS integration (clinical): in order for the solution to be successfully adopted it must be seamlessly integrated into the existing workflow, however there are numerous PACS providers with which the solution will need to integrate.	As with cybersecurity, Jiva.ai leverages the capability of its partners for PACS integration, ensuring scalability across NHS organisations, as well as for this project.
Data access (clinical, ethical): data quality and appropriate annotation is imperative for creation of clinical grade AI diagnostics. Clinical sites must have access to required data sets and have the expertise to label them (if required). Consents and approvals must also be in place.	Much of the due diligence around data access and requirements has already been undertaken between the partners through an existing collaboration.

Table 2 – Risks and mitigations.

Privacy and confidentiality statement

All patient data was anonymised prior to leaving the Health Board. A Data Protection Impact Assessment (DPIA) was undertaken to help assess privacy risks to individuals in the collection, use and disclosure of personal information. A failure to properly embed appropriate data and privacy protection measures may result in a breach of data protection law, a declaration of incompatibility with the Human Rights Act, or prohibitive costs in retrofitting a system to ensure legal compliance or address community concerns about privacy. A cybersecurity assessment was also undertaken prior to commencement.

Objectives

Each key milestone for the project was prepended by a set of tasks encased in a work package to clearly delineate the goals of the project. Milestones 1-4 were undertaken by the Tritech institute and Milestone 5 was undertaken by Health Technology Wales.

- **Tritech Milestone 1:** Build a technical data pipeline and acquire end-to-end data transmission
  -This milestone is critical for a smooth transmission of data between source and predictor. Primarily, this will involve development of APIs between peers as well as completion of administrative tasks for permissioning of data. It will also cover creation of data sharing agreements and discovery of data types at source (some of which has already been covered in previous work between the partners).
- **Tritech Milestone 2:** Build initial validated model and tuning of AI/ML
  -Jiva already has a working model which will need to be calibrated for 1.5T images. The goal will be to validate the model with a blind data set from source. This will require DICOM file anonymisation. This will form a full deployment of an AI prostate diagnostic tool at the site, which can be replicated across other sites.
- **Tritech Milestone 3:** Build a multimodal model to include PSA analysis
  -Create a feasibility fusion model by combining different data verticals - imaging, PSA levels and patient age (at least) to test whether these pieces of data inform a better than practice or better than original diagnostic indicator.
- **Tritech Milestone 4:** Clinical and patient pathway acceptability
  -Establish acceptability and alignment with patient needs and experiences.
- **Health Technology Wales Milestone 5:** Value analysis/health economics assessment
  -Establish the key clinical and resource benefits and risks of using an AI product in clinical practice.

Results

Milestone 1: Build a technical data pipeline and acquire end-to-end data transmission.

Project setup and approvals
Initial discussions with our Cyber and Information Governance teams confirmed a full DPIA would not be required due to the anonymisation of all data being released to Jiva.

Pre-process data for release to Jiva
All MRI scans were anonymised locally by Nina Ralph, radiology system manager, and securely transferred to the third party for analysis by our cyber security specialist Dan Owen.

Establish data transfer and test Extraction, Transfer and Load
Secure transfer of the anonymised data to Jiva servers followed by AI/ML analysis by JivaRDX and return of the results to HDUHB clinicians was achieved over the duration of the project. This established a replicable data compliant pipeline for the continued execution of this project.

The technical data pipeline and end-to-end data transmission was achieved but moving forwards we would wish to examine the integration opportunity of using JivaRDX with the Fuji REiLI (artificial intelligence (AI) enablement) and Synapse (Image analysis) platforms to streamline data integration and flow.

Milestone 2: Build initial validated model and tuning of AI/ML

Validate JivaRDX analysis of 1.5T MRI images within Hywel Dda environment on patient case
An iterative approach to JivaRDX evaluation and refinement was taken for this evaluation, with AI/ML model development followed by clinical evaluation of the outputs performed three times in consecutive series. The reason for this design was to enable a responsive assessment of the AI/ML model to better understand the impact of data characteristics and their influence on the model outputs and overall performance levels.

Iteration 1: Assessment of a ProstateX, JivaRDX model (March, 2022 results)

The JivaRDX algorithm was initially trained on the ProstateX dataset [PROSTATEx Challenge, 2017] and the model applied to the HDUHB patient cohort data to predict the presence of clinically significant prostate cancer. The individual results on a patient-by-patient basis and the overall sensitivity, specificity and accuracy of the predictions are shown in the tables below; note two patients were excluded from the results.

These initial results (appendix 1) showed a sensitivity of 65% in detecting prostate cancer but with a high false positive rate and specificity of only 22%, with overall accuracy below 50%.

Sensitivity (March 22)	**65%**
TRUE POSITIVE	17
FALSE NEGATIVE	9
Specificity	**22%**
TRUE NEGATIVE	4
FALSE POSITIVE	14
Accuracy	**48%**

Table 3 - Iteration 1: Assessment of a ProstateX JivaRDX model (March, 2022 results)

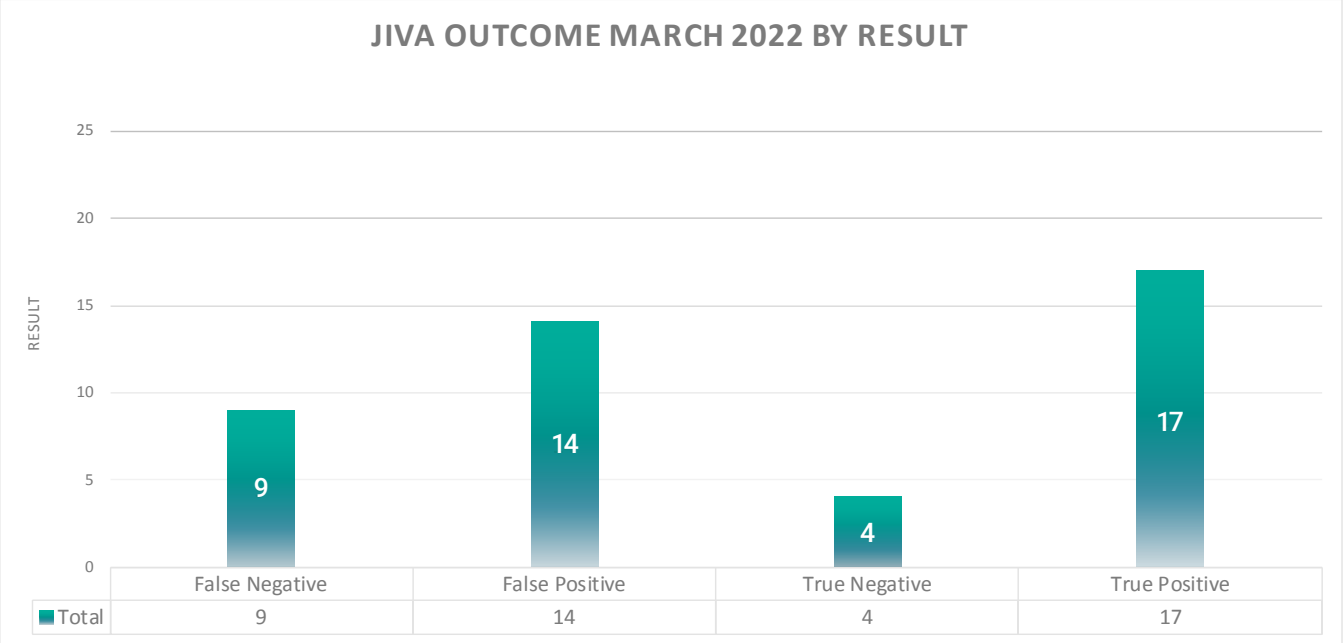


Figure 2 - Iteration 1: Assessment of a ProstateX, JivaRDX model (March 2022 results).

Iteration 2: Assessment of a PI-CAI JivaRDX model (June, 2022 results)

The quality of the training data has a large impact on AI/ML model development and predictive capability. To address issues with training data quality the project team devoted considerable time and resource to sourcing higher quality training data. Of more than ten data sources investigated, the PI-CAI (AI & Radiologists at Prostate Cancer Detection in MRI) Grand Challenge data (https://pi-cai.grand-challenge.org/) was selected as the most appropriate for this project; this is an extensive real-world dataset containing 1500 annotated multi-centre, multi-vendor bpMRI prostate examinations, and including a mix of MRI flux strength (T).

The JivaRDX algorithm AI/ML model was trained against a PI-CAI dataset of 1294 patient cases comprising 1074 benign cases and 220 prostate cancer cases, based on ISUP > 2 (https://pubmed.ncbi.nlm.nih.gov/27150257/). A randomised split of the dataset was performed to establish a training dataset (85% of the data) and a blind subset for technical validation (15% of the data). Numerous model architectures and variants were developed involving hundreds of different hyper parameter configurations over approximately 2000 computational hours. Six models were selected as high performers based on performance criteria of sensitivity >87% and specificity >83%; these models comprised three variants based on ResNet architecture and three variants based in CBRTall architecture (see table).

Model	Version	Accuracy	Precision	Sensitivity	Specificity
ResNet	A	92.70%	80.90%	97.40%	90.80%
ResNet	B	91.40%	76.00%	100.00%	88.00%
ResNet	C	92.70%	87.00%	87.00%	95.00%
CBRTall	A	88.20%	66.70%	89.70%	87.90%
CBRTall	B	86.80%	65.20%	93.80%	84.60%
CBRTall	C	86.80%	63.80%	96.80%	83.80%

Table 4 - Performance of JivaRDX models for blind technical validation against PI-CAI data.

ResNet models generally outperformed the CBRT in terms of accuracy: ResNet variant A (the base model) was selected as the candidate model to take forward based on highest performance characteristics of 97.4% sensitivity, 90.8% specificity and 92.7% accuracy.

The JivaRDX ResNet A model was applied to the HDUHB patient cohort data to predict the presence of clinically significant prostate cancer. The individual results on a patient-by-patient basis and the overall sensitivity, specificity and accuracy of the predictions are shown in the tables below.

Compared to the prior ProstateX JivaRDX assessment, the results (appendix 2) indicated an improved sensitivity (96%) and accuracy (66%) in detecting prostate cancer. However, specificity remained low at 22% suggesting many people would need further testing including many unnecessary biopsies.

Sensitivity (June 22)	**96%**
TRUE POSITIVE	25
FALSE NEGATIVE	1
Specificity	**22%**
TRUE NEGATIVE	4
FALSE POSITIVE	14
Accuracy	**66%**

Table 5 - Iteration 2: Assessment of a PI-CAI JivaRDX model (June results)

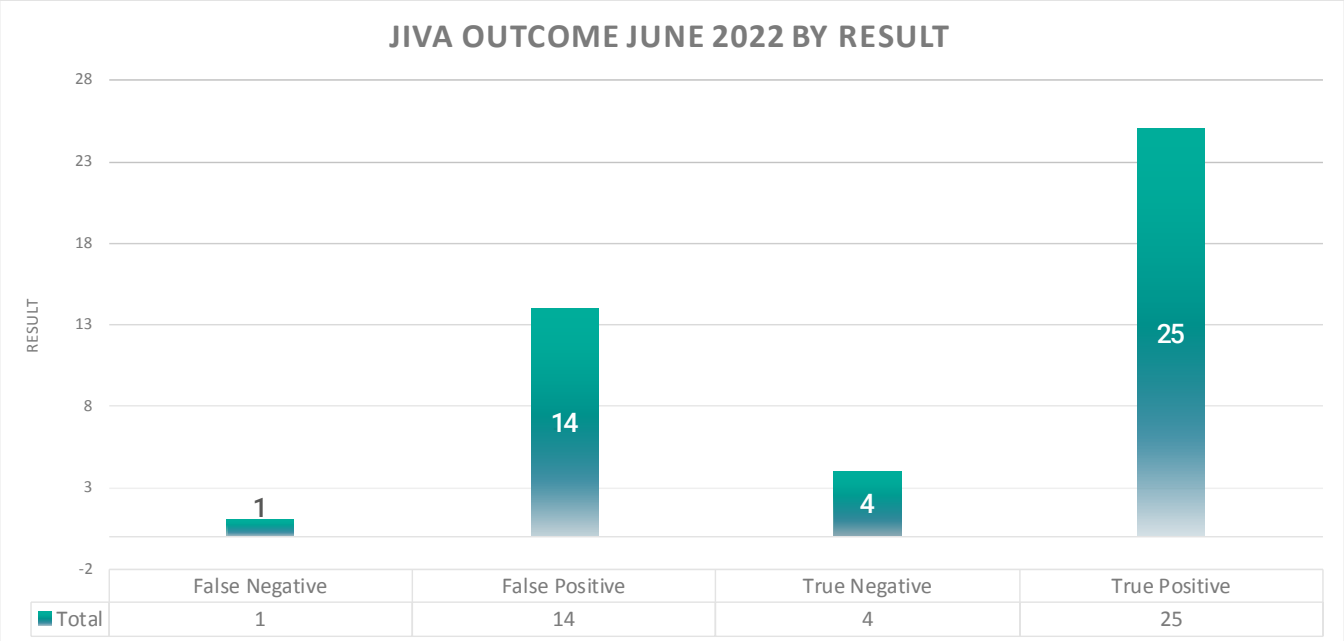


Figure 3 - Iteration 2: Assessment of a PI-CAI JivaRDX model (June results)

During the evaluation we noted variability in the MRI scanning data pack elements from the 44 HDUHB patient cases. In particular, the MRI DWI sequence compositions in 18 cases differed from the remaining 26 cases. The DWI issue arises from some DICOM files indicating the DWI is derivative data, and does not appear to be full DWI. However, close inspection of the DWI data/images indicates no significant difference is observed compared to data that is confirmed to contain full DWI data. Based on this, JivaRDX was applied equally to the full 46 patient datasets although the team exercised vigilance for any detectable bias. This is likely because the 46 cases from Hywel Dda are sourced from different clinical sites which have used a variety of instruments and protocols. Data variability is therefore likely a real-world issue and any deployed tool needs to take this into account.

Iteration 3: Optimisation of a PI-CAI JivaRDX model v2 (July, 2022 results)

On feedback from clinical evaluation of the first PI-CAI model results, further model optimisation was explored to improve specificity of detecting clinically significant cancer. Underperformance in specificity can arise from AI models overfitting during the training process. To address this, the PI-CAI training dataset was split into 5 non-overlapping subsets and the AI model retrained against these independent subsets.

Validation of the retrained model against the HDUHB patient cohort resulted in an inversion of sensitivity and specificity performance. Close inspection of the probability scales output by JivaRDX pointed to thresholds, in this case 0.77, above which all reported predictions were correct. Collectively, the results indicate the algorithm can be calibrated with continued clinician feedback to afford a maximal balance between sensitivity and specificity.

Sensitivity (July 22)	**27%**
TRUE POSITIVE	7
FALSE NEGATIVE	19
Specificity	**94%**
TRUE NEGATIVE	17
FALSE POSITIVE	1
Accuracy	**55%**

Table 6 - Iteration 3: Optimisation of a PI-CAI JivaRDX model v2 (July 22 results)

Iteration 4: Optimisation of a PI-CAI JivaRDX model v3 (March, 2023 results)

On further feedback from clinical evaluation of the first PI-CAI model results, further model optimisation was explored to improve specificity of detecting clinically significant cancer.

Validation of the retrained model against an additional HDUHB patient cohort of 75 patients resulted in an improved sensitivity and specificity performance. Collectively, the results indicate the algorithm can be calibrated with continued clinician feedback to afford a maximal balance between sensitivity and specificity.

Final (revised) results, March, 2023.

Sensitivity (MARCH 23- 3DC)	**76.92%**
TRUE POSITIVE	20
FALSE NEGATIVE	6
Specificity	**65.22%**
TRUE NEGATIVE	30
FALSE POSITIVE	16
Accuracy	**69.44%**

Table 7 - Iteration 4: Optimisation of a PI-CAI JivaRDX model v3 (March, 2023 results).

Milestone 3: Develop a multimodal model that includes MRI and PSA data

The feasibility of developing a multimodal AI/ML approach to JivaRDX for prostate cancer detection was tested using the PI-CAI dataset, which includes PSA, PSA density and age data along with the MRI scans.

	Sensitivity	**Specificity**
MRI only	93%	77%
Multimodal MRI + age + PSA	78%	81%
Multimodal MRI + age + PSAd	81%	93%

These results demonstrate it is indeed feasible to develop a working multimodal model that ingests these data types, and that analysis of multimodal data allows iterative learning and improvements in short time that will improve the diagnostic accuracy of the algorithms.

Feasibility assessments of site compatible automated data transfer integrations

The company Fuji, our Radiology system provider, has a strong vision for integrating AI with their

imaging platform and a mechanism has since been developed. REiLI (AI enablement) and Synapse (Image analysis). Explorations between Jiva.ai and Fujifilm have identified the REiLi platform as a route to enable a clinical deployment of JivaRDX at HDUHB clinical sites, including compliant automated data transfer integrations. The findings support the case to test actual integration of JivaRDX and Fuji platform processes.

Clinician testing of JivaRDX for clinical practice

As the learnings from WP2 and WP3 indicated a need for further calibration of the JivaRDX algorithm we limited clinician testing to obtaining staff feedback on the potential for multimodal AI/ML without progressing to hands on testing.

Milestone 4: Clinical and patient pathway acceptability

It was a key part of this evaluation to look at clinical and patient pathway assessments. We have evaluated professional and patient acceptability in order to establish alignment with staff/patient's needs and experiences.

Patient and Staff acceptability was undertaken using an online patient and staff questionnaire within Microsoft Forms (appendices 5&6) to ascertain the Multi-Disciplinary Team's (MDT) & patients understanding and acceptability on the use of AI/ML in the diagnosis of prostate cancer.

Summary of the Clinical Engagement:

We asked the MDT for their opinions on the impact of AI/ML in terms of :

- Patient Safety
- Patient outcomes
- Efficiency
- Teamwork & Communication

Eleven out of fifteen MDT staff responded and all eleven MDT members had a positive opinion of the Jiva RDX MRI diagnostic aid. All highlighted the positive impact the AI/ML model could have on patient safety, outcomes, teamwork, communication and efficiency. All staff thought the model would have an extremely positive impact upon cancer diagnosis and safety.

"Reading MRI is subjective and needs years of experience. AI can certainly shorten the learning curve. I do not see it as a replacement of a good radiologist but an aid to a clinician". MDT01

"Potential for improved reliability of reporting and streamlined workflow. Also, additional source of reporting from a urologist point of view". MDT02

"AI is a useful tool for diagnosis and planning of radiotherapy treatment - how this is integrated into current workflow patterns are important. It will be important to evaluate and validate any AI systems into routine clinical care to ensure robust safety for patients". MDT03

"Will reduce time in target delineation". MDT04

"Will enable more objective and accurate analysis of prostate MRI (as well as in many other clinical situations)" MDT05

Summary of the Patient Engagement:

Patients reported a lack of knowledge on AI, citing that they needed more information around its role and implementation. Findings cannot be generalised due to poor levels of patient engagement. Only three out of twenty patients responded to the questionnaire but all three patients had a positive opinion of the AI/ML.

In general, our patients reported enthusiasm on the ability of AI/ML to be a positive influence in medicine. They felt healthcare AI/ML was a positive step forward and those patients who understood the concept of AI were supportive of developing AI tools for a variety of different healthcare applications.

"Quickly identifies patterns and correlations based on vast data resources." ,P01

"I am sure using AI can only help decision making",P02

"100% positive on safety, treatment outcomes, diagnosis and monitoring",P03

Discussion

The findings from this evaluation indicate that there is a clinical need for new diagnostic processes in prostate cancer. JivaRDX looks feasible in the real world and is popular at least with clinical staff. However, more work needs to be undertaken to improve JivaRDX accuracy with more training. More work also needs to be undertaken looking at clinical outcomes including the impact on diagnosis times, biopsy rates and survival, to assess its true value:

1. The nature and extent of variability on the real-world MRI data collected from the various locations in Hywel Dda needs to be better understood.

2. A secure, robust and automatable solution for sharing patient sensitive data needs to be found e.g. a solution for automated anonymisation.

3. As clinicians expect to see the JivaRDX prediction within their workflow, more information needs to be gathered on how this might happen and how Jiva can return the necessary values to the clinical systems.

4. Fujifilm (MRI supplier in NHS Wales) provides a platform for integrating AI solutions with their systems, e.g. REiLI (AI enablement and Synapse (Image analysis)

Integration with the Fujifilm platform needs to be tested to determine the practicalities of commercial deployment of JivaRDX in the HDUHB clinical setting.

5. There are trade-offs between sensitivity and specificity that the JivaRDX algorithm can be tuned for e.g. by improving specificity this will reduce the number of false positives and thereby reduce the number of biopsies. Close work with clinicians is needed to understand where and how the algorithm can best be calibrated to maximise clinical impact.

6. The feasibility assessment of multimodal AI/ML demonstrated a marked improvement in the specificity of JivaRDX predictions

Collectively, the project results justify progression to further clinical development of JivaRDX through real-world testing. Exposure to larger and more varied datasets enables better calibration of the AI/ML output that also accounts for wider variations in clinical settings. Further research and development is justified also in multimodal application, which was found to improve the specificity of predictions.

Health Technology Wales NICE META summary:

Milestone 5: Value analysis/health economics assessment

Health Technology Wales have assessed the value of this technology through completion of the NICE's META gap analysis (Appendix 3). A deeper, Value-Based assessment and health economics analysis was also undertaken with Health Enterprise East to construct a decision analytical model to evaluate the relative cost-effectiveness of JivaRDX.

This META report considers the use of JivaRDX, which is a software application designed to assist radiologists in analysing multiparametric magnetic resonance imaging (mpMRI) scans of people with suspected prostate cancer. JivaRDX uses computer vision and machine learning techniques to read, interpret, analyse, and generate findings from mpMRI data. It is intended to be a diagnostic aid rather than a standalone diagnostic tool. It could support radiologists by drawing attention to areas of the mpMRI scan that may have been missed or misinterpreted. It could also provide reassurance to the radiologist in confirming their suspicion about a potential prostate cancer tumour, especially in more marginal cases.

The introduction of JivaRDX into the healthcare system could improve diagnostic accuracy and thereby lead to earlier detection of prostate cancer and fewer false positive results leading to 'unnecessary' biopsies. There has been limited evidence collected on the diagnostic accuracy and clinical effectiveness of JivaRDX to date. However, the company plan to collect evidence in an upcoming study, which will estimate the diagnostic accuracy of JivaRDX in comparison to an independent consultant radiologist. The study will involve a retrospective multicentre data analysis of 265 cases of suspected prostate cancer. Preliminary findings show that JivaRDX has improved sensitivity and specificity in comparison to the current standard of care.

There is limited evidence available on the cost-effectiveness of JivaRDX this stage. However, the company have developed a preliminary health economic analysis using a decision analytic model. The results showed JivaRDX to be more effective (0.125 QALYs) and less costly (£154) than standard care and can therefore be considered dominant [Health Technology Wales (HTW), July 2022].

When developing future clinical and economic evidence collection plans, consideration should be given to how well the study reflects the likely use of the technology if it were to be adopted in clinical practice. In particular, it is important to ensure that the evidence collected reflects the key comparison of interest, which is JivaRDX alongside radiologist interpretation in comparison to radiologist interpretation alone. This will provide the strongest support for the technology's value proposition and thereby improve the case for adoption.

Jiva.ai enlisted the support of Health Enterprise East to construct a decision analytical model to evaluate the relative cost-effectiveness of JivaRDX as part of the META tool work.

Jiva.ai engaged with Health Enterprise East to construct a decision analytical model to evaluate the relative cost-effectiveness of JivaRDX, an artificial intelligence (AI) system developed by Jiva.ai Ltd. JivaRDX is a diagnostic adjunct to be used with multi-parametric magnetic resonance imaging (mpMRI) to improve both sensitivity and specificity of MRI in the diagnostic pathway for the detection of prostate cancer (PCa), as recommended by the NICE diagnostic pathway NG131 [Ref: WP4.2-1].

The perspective of the analysis was the NHS. The patients modelled were based on the characteristics of the sample used by the Prostate MRI Imaging Study (PROMIS) [Ref: WP4.2-2], with a mean age of 63 years old. The comparator reflects the diagnostic care pathway for PCa as recommended by NICE (NG131) [Ref: WP4.2-1]. Health outcomes were modelled in Quality-Adjusted Life Years (QALYS), applying the weightings from the study by Faria et al [Ref: WP4.2-3].

The decision analytical model combined a decision tree model for the diagnostic outcomes of repeated testing undergone by a cohort of patients, with a Markov model to simulate the results over a timeframe of 40 years.

The results from the decision analytical model suggest that the JivaRDX system is cost-effective from an NHS perspective, resulting in mean cost savings of £154 per patient over his lifetime (95% CI: -£ 206; -£ 98) based on 0.91 Sensitivity and 0.88 specificity performance parameters.

The incremental cost-effectiveness ratio (ICER) of the intervention stands at -£ 1,683 per QALY, and a probabilistic sensitivity analysis of the results indicates that the innovation is cost-effective in 97.5% of iterations at the NICE threshold of £20,000/QALY and £30,000/QALY. While cost savings are statistically significant at the 5% significance level, uncertainty remains over the potential impact on health outcomes, where the confidence interval indicates that there is insufficient evidence to reject the null hypothesis (H0 = JivaRDX has no impact on health outcomes).

Summary

There was uncertainty around the potential impact on health outcomes with insufficient evidence to reject the null hypothesis that JivaRDX has no impact on health outcomes.

Evaluation Conclusions

The evaluation focused upon identifying the impact of JivaRDX an artificial intelligence (AI)/ machine learning (ML)-based MRI diagnostic aid for Prostate Cancer over an eighteen month period across all acute sites within Hywel Dda University Health Board (HDUHB).

Does JivaRDX work?

In conclusion the JivaRDX AI/ML based MRI model was found to provide 77% sensitivity, 65% specificity & 69% accuracy in detecting prostate cancer. Multiple studies have shown Radiologist specificity at 57%. Whilst these results are extremely promising, further analysis is required before JivaRDX is moved into routine clinical care within Hywel Dda.

Does it show to be acceptable to staff?
Yes, all staff interviewed had a positive opinion on the Jiva RDX MRI diagnostic aid. The staff who responded to the survey all highlighted the positive impact the AI model would have on patient safety, outcomes, teamwork, communication and efficiency. All staff thought the model would have an extremely positive impact upon cancer diagnosis and safety.

We considered the barriers and enablers to implementation and training, validation and accuracy featured prominently throughout the survey.

There was interest in using the technology for further research to determine the effectiveness of other clinical modalities to reduce treatment lengths and to test it on other conditions or symptoms. More information is needed to convince all staff of the long-term effectiveness of the MRI diagnostic aid, but the technology was well received by all.

Does it show to be acceptable to patients?
Since patients are the intended recipients of many AI innovations, more carefully distinguishing their understanding, values, and priorities is important for ensuring that these advances are not just well-received but are developed and implemented in a joint ethical way that improves patient care. In situations where patients interface directly with AI technologies, patients bear the largest risk should implementation be done incorrectly or unethically, to the extent that patients will be asked to accept the potential risks associated with innovative applications of AI in healthcare, there is an ethical obligation to ensure that patient values and needs are incorporated into our thinking and plans.

Our 3 responders reported enthusiasm about the ability of AI to be a positive influence in medicine. They felt healthcare AI was a positive step forward: to heal as many patients as possible. Those patients who understood the concept of AI were supportive of developing AI tools for a variety of different healthcare applications but there were others who didn't understand its application to healthcare and refused to respond to the survey.

Can JivaRDX be implemented as a service within NHS Wales?
From all the information collected during this service evaluation, the current answer is no. Jiva have previously demonstrated a proof-of-concept achieving detection and localisation of PC a from 3T MRI scans (87% sensitivity, 67% specificity) and our results demonstrate that in a real-world environment we achieved 77% sensitivity, 65% specificity & 69% accuracy.

The JivaRDX artificial intelligence (AI)/machine learning (ML)-based MRI platform was found to be sensitive and specific and to be acceptable to staff and some cancer patients. However, the longer-term clinical benefits and patient understanding of AI need further analysis before JivaRDX is moved into routine clinical care within Hywel Dda.

Does JivaRDX have Value?
The results from the Health Enterprise East decision analytical model suggest that the JivaRDX system resulted in average cost savings of £154 (95% CI:-£ 206; -£ 98) per patient over the modelled time horizon. There is uncertainty around the potential impact on health outcomes as the confidence interval indicates that there is insufficient evidence to reject the null hypothesis that JivaRDX has no impact on health outcomes. However, the base case result suggests that JivaRDX resulted in a gain of 0.125 QALYs per patient over the modelled time horizon.

In summary
The JivaRDX system has proven to be sensitive and specific in terms of its diagnosis potential and further studies are required to refine the model. We found that the system has the capability to integrate into our current clinical systems and pathways. Furthermore, our engagements with clinical teams and patients identifies a general positive reaction to the use of AI as long as there are safeguards in place. On the next page we list our recommendations for future work regarding the implementation of this innovation in the PCa diagnostic pathway.

Recommendations & the way forward

Recommendation 1: [Improve sensitivity & specificity of AI MRI outcomes]

The JivaRDX system has proven to be sensitive and specific. The future of clinically successful healthcare AI relies on robust accuracy. Increasing specificity from 90% to 95% amounts to cutting false positives (and false alerts) by two-fold. Jiva.ai algorithms must feature both high sensitivity AND high specificity in a real-world clinical and radiological setting.

Recommendation 2: [Regulatory approval]

Clinical investigation to be submitted to MHRA for the study 'Jiva.ai MRI validation of JivaRDX for Prostate Cancer' for the company to seek regulatory approval across the UK.

Recommendation 3: [Integration of multimodal AI]

Examine the integration opportunity of using JivaRDX with the Fuji REiLI (artificial intelligence (AI) enablement) and Synapse (Image analysis) platforms to streamline data integration and flow.

Recommendation 4: [Improve patient understanding of AI]

The patients' view on the implementation of AI in radiology is still mainly unexplored territory. Successful implementation of AI in radiology requires the assessment of our patient views towards the technology.

Recommendation 5: [Account for Data variability in a real-world clinical environment]

The 121 cases from Hywel Dda were sourced from different clinical sites which have used a variety of MRI imaging devices and protocols. Data variability is therefore a real-world issue, and any deployed AI tool needs to take this into account. Jiva should undertake a retrospective multicenter study.

The aim of the next larger study should also compare performance of the software against an independent radiology expert, e.g., showing that JivaRDX is not worse or better than standard care in terms accuracy of detecting clinically relevant lesions in MRI scans performed for prostate cancer.

Appendices

Appendix 1 – Iteration 1: Assessment of a ProstateX, JivaRDX model March, 2022 results.

JIVA Outcome March 2022	1 signifies clinically significant cancer is detected and a 0 signifies no clinically significant cancer detected. 0 is Likert score 1-2 and 1 is Likert score 3,4&5 March 2022	Actual equivalent	Matches original Likert Score Yes / No	Result
	JIVA AI generated Test result	Reality		
Patient1	0	1	False	False Negative
Patient2	1	1	True	True Positive
Patient3	1	0	False	False Positive
Patient4	0	1	False	False Negative
Patient5	1	1	True	True Positive
Patient6	1	0	False	False Positive
Patient7	0	1	False	False Negative
Patient8	1	1	True	True Positive
Patient9	1	1	True	True Positive
Patient10	1	0	False	False Positive
Patient11	1	0	False	False Positive
Patient12	1	1	True	True Positive
Patient13	0	1	False	False Negative
Patient14	1	0	False	False Positive
~~Patient15~~	~~1~~	~~0~~		
~~Patient16~~	~~1~~	~~0~~		
Patient17	0	0	True	True Negative
Patient18	1	1	True	True Positive
Patient19	0	0	True	True Negative
Patient20	0	1	False	False Negative
Patient21	1	1	True	True Positive
Patient22	0	0	True	True Negative
Patient23	1	0	False	False Positive
Patient24	1	0	False	False Positive
Patient25	1	1	True	True Positive
Patient26	1	1	True	True Positive
Patient27	0	1	False	False Negative
Patient28	1	0	False	False Positive
Patient29	1	1	True	True Positive
Patient30	1	1	True	True Positive
Patient31	1	1	True	True Positive
Patient32	1	1	True	True Positive
Patient33	1	0	False	False Positive
Patient34	0	1	False	False Negative
Patient35	1	1	True	True Positive
Patient36	0	1	False	False Negative
Patient37	1	0	False	False Positive
Patient38	1	0	False	False Positive
Patient39	1	0	False	False Positive
Patient40	0	1	False	False Negative
Patient41	1	0	False	False Positive
Patient42	1	1	True	True Positive
Patient43	1	1	True	True Positive
Patient44	1	1	True	True Positive
Patient45	0	0	True	True Negative
Patient46	1	0	False	False Positive

Appendix 2 – Iteration 2: Assessment of a PI-CAI JivaRDX model June, 2022 results.

JIVA Outcome June 2022	1 signifies clinically significant cancer is detected and a 0 signifies no clinically significant cancer detected. 0 is Likert score 1-2 and 1 is Likert score 3,4&5 June 2022	Actual equivalent	Matches original Likert Score True / False	Result
	JIVA AI generated Test result	Reality		
Patient1	1	1	True	True Positive
Patient2	1	1	True	True Positive
Patient3	1	0	False	False Positive
Patient4	1	1	True	True Positive
Patient5	1	1	True	True Positive
Patient6	0	0	True	True Negative
Patient7	1	1	True	True Positive
Patient8	1	1	True	True Positive
Patient9	1	1	True	True Positive
Patient10	1	0	False	False Positive
Patient11	1	0	False	False Positive
Patient12	1	1	True	True Positive
Patient13	1	1	True	True Positive
Patient14	1	0	False	False Positive
~~Patient15~~		~~0~~		
~~Patient16~~		~~0~~		
Patient17	0	0	True	True Negative
Patient18	1	1	True	True Positive
Patient19	0	0	True	True Negative
Patient20	1	1	True	True Positive
Patient21	1	1	True	True Positive
Patient22	0	0	True	True Negative
Patient23	1	0	False	False Positive
Patient24	1	0	False	False Positive
Patient25	1	1	True	True Positive
Patient26	1	1	True	True Positive
Patient27	1	1	True	True Positive
Patient28	1	0	False	False Positive
Patient29	1	1	True	True Positive
Patient30	1	1	True	True Positive
Patient31	1	1	True	True Positive
Patient32	1	1	True	True Positive
Patient33	1	0	False	False Positive
Patient34	1	1	True	True Positive
Patient35	1	1	True	True Positive
Patient36	1	1	True	True Positive
Patient37	1	0	False	False Positive
Patient38	1	0	False	False Positive
Patient39	1	0	False	False Positive
Patient40	1	1	True	True Positive
Patient41	1	0	False	False Positive
Patient42	1	1	True	True Positive
Patient43	0	1	False	False Negative
Patient44	1	1	True	True Positive
Patient45	1	0	False	False Positive
Patient46	1	0	False	False Positive

Appendix 3 – Medtech Early Technical Assessment
Summary compiled by: Health Technology Wales (HTW), July 2022

Key assessment points

Summary

This META report considers the use of JivaRDX, which is a software application designed to assist radiologists in analysing multiparametric magnetic resonance imaging (mpMRI) scans of people with suspected prostate cancer. JivaRDX uses computer vision and machine learning techniques to read, interpret, analyse, and generate findings from mpMRI data. It is intended to be a diagnostic aid rather than a standalone diagnostic tool. It could support radiologists by drawing attention to areas of the mpMRI scan that may have been missed or misinterpreted. It could also provide reassurance to the radiologist in confirming their suspicion about a potential prostate cancer tumour, especially in more marginal cases.

The introduction of JivaRDX into the healthcare system could improve diagnostic accuracy and thereby lead to earlier detection of prostate cancer and fewer false positive results leading to 'unnecessary' biopsies. There has been limited evidence collected on the diagnostic accuracy and clinical effectiveness of JivaRDX to date. However, the company plan to collect evidence in an upcoming study, which will estimate the diagnostic accuracy of JivaRDX in comparison to an independent consultant radiologist. The study will involve a retrospective multicentre data analysis of 265 cases of suspected prostate cancer. Preliminary findings show that JivaRDX has improved sensitivity and specificity in comparison to the current standard of care.

There is limited evidence available on the cost-effectiveness of Jiva RDX this stage. However, the company have developed a preliminary health economic analysis using a decision analytic model. The results showed JivaRDX to be more effective (0.125 QALYs) and less costly (£154) than standard care and can therefore be considered dominant. However, there was uncertainty around the potential impact on health outcomes with insufficient evidence to reject the null hypothesis that JivaRDX has no impact on health outcomes.

When developing future clinical and economic evidence collection plans, consideration should be given to how well the study reflects the likely use of the technology if it were to be adopted in clinical practice. In particular, it is important to ensure that the evidence collected reflects the key comparison of interest, which is JivaRDX alongside radiologist interpretation in comparison to radiologist interpretation alone. This will provide the strongest support for the technology's value proposition and thereby improve the case for adoption.

Regulatory and HTA requirements

The company do not have regulatory approval at present but work is ongoing to achieve the appropriate certification. The company are planning to undertake an initial assessment of the analytical validity of JivaRDX in order to file for regulatory approval with the Food and Drug Administration (FDA). This will involve a retrospective analysis focusing on diagnostic accuracy (sensitivity and specificity) and technical safety.

The company plan to obtain a CE mark via the conformity assessment body TÜV SÜD. The company is also aware they will need a UK Conformity Assessed (UKCA) mark to be sold in the UK. This is a new post-Brexit requirement for goods placed on the market in England, Wales and Scotland. The UKCA marking came into effect on 1 January 2021 but, in most cases, it will be possible to use the CE marking until 1 January 2023 to allow businesses time to adjust to the new requirements.

JivaRDX is a digital health technology and as such, it is useful to refer to the 'NICE evidence standards framework for digital technologies'. This document provides guidance on evidence requirements for digital health technologies by categorising them into different 'evidence tiers' according to the functions that the technologies perform. Based on our discussions, JivaRDX is expected to fall within 'Tier C: Interventions'. This tier level seems appropriate based on the following functional components listed in the NICE evidence standards framework, which seem pertinent to JivaRDX:

1. "Diagnose" – "Uses data to diagnose a condition in a patient, or to guide a diagnostic decision made by a healthcare professional."
2. "Calculate" – "Tools that perform clinical calculations that are likely to affect clinical care decisions."

Key considerations for evidence requirements within this tier are as follows:
- Demonstrating effectiveness for treat, active monitoring, calculate or diagnose functions
- Demonstrating the accuracy and reliability of measurements and information provided, recorded or transmitted by the technology,
- Using ongoing data collection to demonstrate usage and value of the technology
- Ensuring that appropriate safeguarding measures are in place
- User satisfaction and acceptability
- Showing relevance to current care pathways through a successful pilot in the UK health and social care system

Note that this list is not exhaustive and the manufacturer is advised to consult the NICE evidence standards framework for full details. Furthermore, note that evidence requirements are higher for technologies that are considered to be higher-risk. JivaRDX may fall within this higher risk category because of the seriousness of the consequences if the technology fails to perform as described. Furthermore, the technology uses machine learning algorithms which are classified as a risk factor within the evidence standards. However, it should be noted that 'adaptive algorithms' which continuously learn and update automatically are considered to be of higher risk than 'fixed algorithms'. Therefore, the company's plan to implement JivaRDX as a fixed algorithm should make this risk factor less of a concern. The risk level may be further mitigated by the involvement of the clinician in the decision making process. The intention is not for JivaRDX to replace the expert judgement of the clinician but rather to supplement it and this should lessen the perceived risk of the technology.

Note also that there are specific considerations for the development and validation of machine learning algorithms to ensure that potential biases are minimised when generating the training and validation data sets for the algorithm. The key consideration is to ensure that there is separation between the dataset used for training and validating the algorithm. To address this concern, the manufacturer has split their data into a training (60%) and validation dataset (40%). The source data is only sent to Jiva.ai if it is deemed to be training data while validation data is sent directly to ARU, who are independently conducting their evaluation. Therefore, there is minimal chance of the algorithm training set being contaminated.

Value proposition

The value proposition defines the benefits that could be attained by patients and the health and social care system if a technology was adopted into practice. Key considerations are whether the technology would improve patient quality of life, survival and resource use or costs in the health and social care system. Limited evidence has been collected to date to support JivaRDX's value proposition. However, there is potential for patient and system benefits if JivaRDX leads to improvements in diagnostic accuracy compared to standard care.

Improvements in specificity would lead to fewer patients with false positive results undergoing an 'unnecessary' biopsy. Clearly, this would be a benefit to the healthcare system as there is a cost for each biopsy and, at a time of high demand, freeing up the time and resources for use in other patient cases would release some pressure and reduce the backlog of patients on the waiting list. There could also be an impact of patient outcomes as having a suspicious case that needs further investigation can have a detrimental psychological impact (especially if there is a long wait for the subsequent biopsy). Furthermore, prostate biopsies are invasive procedures and are associated with known complications, such as pain, discomfort, bleeding, urine retention and erectile dysfunction. Most of these complications would have a short-term impact on quality of life but some may prove to be ongoing issues and so have an ongoing impact on quality of life. There may also be additional costs associated with managing these complications. Reducing the number of biopsies that prove to be false positive could therefore lead to improvements in quality of life and a reduction in costs.

There is further potential for benefits if JivaRDX leads to an improvement in sensitivity meaning that it is more likely for prostate cancer to be accurately detected. Improvements in sensitivity typically translate into faster diagnosis time and therefore faster time to treatment. This should mean that there would be cases where cancer is detected and treated at an earlier stage, thereby leading to improvements in disease control and ultimately improving patient's chances of survival. There could also be improvements in quality of life as the chances of recurrence and disease progression should be lower by catching and treating cancer at an earlier stage. Therefore, some patients may be able to avoid treatments for recurrent or progressed disease that they would otherwise have endured. There would also be system benefits associated with earlier detection and treatment as it may mean that costly treatments for recurrent and advanced disease can be avoided (such as multiple cycles of systemic therapy).

Clinical treatment pathway

JivaRDX is intended to be used alongside mpMRI scans that are currently used for prostate cancer diagnosis. The current use of mpMRI scans is largely informed by the diagnostic pathway specified in the NICE guideline on prostate cancer (NG131). The guideline recommended that mpMRI is offered as the first-line investigation for people with suspected clinically localised prostate cancer with results reported using a 5-point Likert scale. However, there was an exception as it also recommends that mpMRI should not be routinely offered to people with prostate cancer who are not going to be able to have radical treatment.

Following the mpMRI scan, NICE recommends that a biopsy is offered to people with a Likert score of 3 or more. In people with a Likert score of 1-2, NICE recommends omitting a prostate but only after discussing the risks and benefits with the person and reaching a shared decision.

The diagnostic pathway would be very similar if JivaRDX was to be implemented into practice but the key difference would be the decision-making process when interpreting the mpMRI scans. Radiologists would analyse the mpMRI scan as per their normal clinical protocol but they would also have the output from JivaRDX. JivaRDX analyses the scans to identify tumour lesions of clinical relevance and outputs a 'positive' or 'negative' result based on Prostate Imaging Reporting and Data System (PI-RADS) scores. A negative result indicates a PI-RAD score of 1-2 (considered to be clinically not relevant) while a positive result indicates a PI-RAD score of 3-5 (clinically relevant).

Thus, JivaRDX would act as a diagnostic aid to the radiologist as it provides a 'second read' of the mpMRI scan. Note, however, that JivaRDX is based on the mpMRI analysis only and it would not have the surrounding context available to the radiologist such as patient history, results of PSA tests or digital rectal examinations (DRE). Therefore, it is anticipated that the radiologist would still have to exercise their own judgement when making their decision.

Note that there is currently a discrepancy in the scoring system proposed by NICE and the scoring system used by JivaRDX. NICE recommends the Likert system whereas JivaRDX uses the PI-RAD system. The Likert system is more commonly used in the UK NHS whereas PI-RAD is more commonly used in the US healthcare system. The manufacturer is aware of this discrepancy and is investigating how it could impact the interpretation of mpMRI

The PICO statement

It is useful to consider the potential PICO (population, intervention, comparator and outcomes) statement for the technology as this framework is typically used in technology assessments to structure evidence reviews. Based on our discussions, the potential PICO statement for the technology is expected to be as follows:

Population: people with suspected prostate cancer undergoing an mpMRI scan

Note that people are typically referred for investigation for suspected prostate cancer based on their PSA level, DRE findings and other risk factors (such as age).

Intervention: JivaRDX assessment alongside radiologist interpretation of mpMRI findings

Comparator: Standard radiologist assessment of mpMRI findings

Outcomes: Diagnostic accuracy (sensitivity, specificity, negative predictive value, positive predictive value), time to diagnosis, time to treatment, adverse events, mortality, quality of life, patient satisfaction, healthcare resource use (including number of biopsies)

The outcomes listed above are relatively typical for health technology assessment of technologies involved in diagnosis. The core outcomes relate to diagnostic performance but, ideally, this should be supplemented with outcomes that demonstrate the impact of changes in diagnostic accuracy. As outlined above in the value proposition section, there are numerous patient and system benefits that could be attained through improvements in accuracy. Ideally, these benefits should be captured within the evidence base by measuring improvements in clinical outcomes, quality of life as well as reductions in resource use and costs. However, in practice, it is not always possible to capture all these outcomes within clinical studies. In particular, it can be difficult to capture the long-term benefits of changes in the diagnostic pathway, such as improvements in overall survival.

The PICO above reflects the use of JivaRDX as part of initial prostate cancer diagnosis. The company are aware that, in current practice, mpMRI is also undertaken as part of active surveillance (i.e. the close monitoring of known prostate cancer). The use of JivaRDX in this setting could be a future area of development for the company. Note that the PICO would be different for this population as it is a separate indication and it is therefore likely that most Health Technology Assessment bodies would assess it as a separate use case.

Measuring clinical effectiveness

At present, there is limited evidence on the diagnostic accuracy and clinical effectiveness of JivaRDX. However, the company plan to collect evidence in an upcoming study, which will estimate the diagnostic accuracy of JivaRDX in comparison to an independent consultant radiologist (with two radiologists used in the study) for the detection of clinically significant prostate cancer. The study will involve a retrospective multicentre data analysis of 265 cases of suspected prostate cancer. The primary aim of the study is to demonstrate the non-inferiority of JivaRDX in comparison to interpretation by a consultant radiologist. Diagnostic accuracy will be reported in terms of sensitivity, specificity, positive predictive value and negative predictive value.

The manufacturer has preliminary results on the diagnostic accuracy of JivaRDX, which show that it has a sensitivity of 91% and specificity of 88% for detecting clinically significant cancer cases. This compares favourably with the company's estimate of diagnostic accuracy with the current standard of care within the NHS (i.e. interpretation by a radiologist), which is estimated to have a sensitivity of 88% and specificity of 45% for detecting clinically significant cancer cases.

The company note that the evidence collection plans to date have largely been intended as a pilot study demonstrating proof-of-concept and allowing the company to file for regulatory approval. Further studies may be undertaken to demonstrate clinical validity and utility and this could include additional outcomes of interest, such as biopsies avoided, speed of diagnosis and quality of life. If it is feasible then the study may also aim to demonstrate superiority rather than non-inferiority.

When undertaking further research, consideration should be given to how well the study reflects the likely use of the technology if it were to be adopted in clinical practice. In particular, it has been proposed that JivaRDX would be used in conjunction with radiologist interpretation rather than instead of it. Therefore, the appropriate comparison would be JivaRDX alongside radiologist interpretation in comparison to radiologist interpretation alone (as outlined in the PICO above). The study described above appears to be a comparison of JivaRDX versus radiologist interpretation alone and therefore doesn't quite match the comparison of interest. Such a comparison is still useful as it demonstrates the accuracy of JivaRDX but further work and some assumptions would be required to come up with combined accuracy for JivaRDX alongside radiologist interpretation.

Consideration should be given to how well the study population reflects the likely population that would use the technology in clinical practice. A key component within this consideration is the prevalence of cancer in the study population and how closely it matches the likely prevalence of cancer in clinical practice. There may also be generalisability issues that require consideration if using data from a different country with a different healthcare system.

Economic data collection

At present, there is limited evidence on the cost effectiveness of JivaRDX. However, the company have developed a preliminary health economic analysis using a decision analytic model. The analysis has not been published but the manufacturer shared an executive summary with HTW for the purpose of writing this META tool report and we discussed some aspects of the model during our facilitation meeting. Note, however, that the following summary and commentary is based only on partial information and is not intended to be a full examination and critique of the economic analysis.

The model consists of an upfront decision tree for diagnosis and a Markov model to estimate longer-term impact (up to 40 years). The analysis estimates the cost effectiveness of JivaRDX in comparison to standard care for the diagnosis of prostate cancer from the perspective of the UK NHS. Standard care in the economic model was based upon the diagnostic pathway recommended in the NICE clinical guideline on prostate cancer (NG131). A hypothetical cohort of patients were modelled based on the characteristics of the Prostate MRI Imaging Study (PROMIS).

The diagnostic accuracy of mpMRI and JivaRDX in addition to mpMRI was based on the sensitivity and specificity values described in the above section on 'measuring clinical effectiveness'. Diagnostic accuracy values for the PSA test and the prostate biopsy were based on previous studies. It was assumed that people would be re-tested following a negative result every two years and that the accuracy of first and subsequent tests would be equivalent. The potential complications of biopsy were captured in the model with a possibility of patients experiencing erectile dysfunction, urinary tract infection (UTI), haematuria and sepsis. If patients have sepsis, then there was also assumed to be a 20% probability of mortality.

Transition probabilities in the Markov component of the model were based on a regression analysis by Faria et al. using data from the PROMIS study. The Markov model consists of three health states; 'localised/progression free', 'metastatic disease' and 'death' and the transition probabilities between health states varies depending on the treatment or management strategy adopted. The modelled probability of progressing to metastatic disease and death was higher in people receiving non-surgical management such as, radiotherapy or watchful waiting) compared to surgical management (radical prostatectomy). People correctly diagnosed with prostate cancer (i.e. true positives) were assumed to receive surgical or non-surgical management based on treatment proportions from a previous study (24% received prostatectomy). People with undetected prostate cancer (i.e. false negatives) were assumed to have the same probability of progression and death as non-surgical management.

Health outcomes were expressed in terms of quality adjusted life years (QALYs), which are estimated by combining life year estimates with quality of life values associated with being in a particular health state. Quality of life weightings were sourced from previous studies by Faria et al. and Korfage et al. Men with localised prostate cancer were assumed to have the same quality of life as similarly aged men with no cancer. The treatment and management of localised cancer was assumed to have no impact on quality of life but progression to metastatic disease resulted in a reduction in quality of life.

The cost associated with mpMRI scans appear to be based on NHS reference costs with different costs depending on whether contrast is used or not. The addition of JivaRDX was assumed to cost £25 per scan, which reflects the company's planned funding model (see section below on funding and commissioning). The majority of other costs within the model appear to have been sourced from NHS reference costs. The exception is the cost associated with management strategies, such as radical prostatectomy and radiotherapy. These costs appear to be based on a previous study and have been implemented as annual costs.

The results showed that JivaRDX resulted in average cost savings of £154 (95% CI: -£ 206; -£ 98) per patient over the modelled time horizon. There is uncertainty around the potential impact on health outcomes as the confidence interval indicates that there is insufficient evidence to reject the null hypothesis that JivaRDX has no impact on health outcomes. However, the base case result suggests that JivaRDX resulted in a gain of 0.125 QALYs per patient over the modelled time horizon. As such, the base case results suggest that JivaRDX is more effective and less costly than standard care and is therefore dominant.

This early stage economic analysis is useful for showing the potential value of adding JivaRDX into the NHS system and it could be developed further as more evidence becomes available. Future developments should consider how well the study reflects the likely use of the technology if it were to be adopted in clinical practice. As was the case with the clinical evidence to date, the analysis appears to be a comparison of JivaRDX versus radiologist interpretation alone. However, the key comparison of interest is JivaRDX alongside radiologist interpretation in

radiologist interpretation. Ideally, the combined accuracy would be sourced directly from clinical studies. However, if this is not possible then consideration may have to be given to how to combine the individual accuracy of JivaRDX and radiologist interpretation. This would involve making a simplifying assumption around what happens when there is a discrepancy between the radiologist's interpretation and JivaRDX. For example, it could be assumed that a 'positive' result with either approach would need to be investigated under biopsy. Such an approach would increase detection of cancer (i.e. high sensitivity) but it would be at the expense of false positives (i.e. low specificity). Alternatively, it could be assumed that a 'positive' result is needed with both approaches in order for a biopsy to be undertaken. This would reduce the likelihood of false positives but it would be at the expense of sensitivity.

A further area of development for the economic model could be the introduction of more health states in the model. As things stand there are only two alive states; 'localised/progression free' or 'metastatic disease' and this means that different stages of disease are combined within the non-metastatic health state. Therefore, this approach may miss some of the benefit of earlier diagnosis as there is a possibility of disease progression in undiagnosed patients resulting in no change to their health state (i.e. if it progresses but not to the extent of metastatic disease). The introduction of an intermediate health state between localised and metastatic disease would help to capture this stage of disease (and its associated costs and quality of life). Previous diagnostic models in prostate cancer have often used a 'locally advanced disease' health state in addition to localised and metastatic disease (such as Mowatt et al. 2013).

A further area of development for the economic model could be in the estimation of quality of life values. The company note that they are planning to collect quality of life data using the EuroQol five dimensions survey (EQ-5D). This data may be useful as an additional source of quality of life inputs in the analysis. In particular, they may be useful for earlier time points in the diagnostic pathway. For example, they may be able to estimate the quality of life impact of experiencing biopsy complications (which are currently assumed to have no quality of life impact). There could also be value in exploring whether quality of life would be different in other stages of the pathway. As things stand, there is no decrement associated with having localised cancer and its associated treatment. However, there are known complications with prostate cancer treatment (such as erectile dysfunction and urinary incontinence) and it is likely that they would have an impact on quality of life.

There could also be value in exploring alternative cost inputs. In particular, whether an alternative approach could be used to capture management costs associated with radiotherapy, radical prostatectomy, watchful waiting and metastatic cancer. As things stand, these are implemented as average costs per patient per year. However, we know that active treatment options would involve a 'one-off' upfront cost, such as the radical prostatectomy procedure, followed by ongoing costs, such as follow-up costs. Therefore, the average cost is likely to have involved spreading the cost of an upfront procedure over the patient's expected lifetime. This can cause problems in an economic analysis as it is likely to mean that total costs are underestimated in people that die earlier than expected and overestimated in people that die later than expected. This issue is compounded by the use of discounting, which is undertaken as standard in economic analysis (NICE recommends a discount rate of 3.5% per year). This means that future costs and benefits are valued lower than current costs and benefits and therefore the time point that they are incurred can make a difference to results. For these reasons, it may be advisable to implement treatment costs as an upfront cost whether a separate cost applied for any ongoing costs (such as follow-up).

Funding and commissioning

The company have considered two options for charging for the JivaRDX service; the first option involves charging a cost per patient while the second option involves charging an annual subscription fee. A software license fee of £30,000 was considered for the annual subscription model while a cost of £25 per scan was considered for the per patient model. Having considered both options, the company preference is to offer the service on a cost per patient basis. As well as including access to the application, the cost would also cover implementation and setup as well as basic training to use the service.

The relevant healthcare resource group (HRG) codes for the technology will be those relating to the use of MRI. Note that HRG codes do not distinguish between mpMRI and conventional MRI scans but they do vary depending on the number of areas scanned and whether contrast is used. Based on costs listed in the National Tariff Payment System 2022/23, the cost of an MRI scan varies from £116 for a scan of one area without contrast to £201 for scans of more than three areas. In addition, there is a further increased MRI cost of £236 covering instances where extensive patient repositioning is required. The tariff costs include the cost of reporting as well as the cost of the scan. However, the cost of reporting is presented separately and in most cases a cost of £24 per scan has been estimated. A higher reporting cost of £31 is assumed where two or more areas are scanned or extensive repositioning is required.

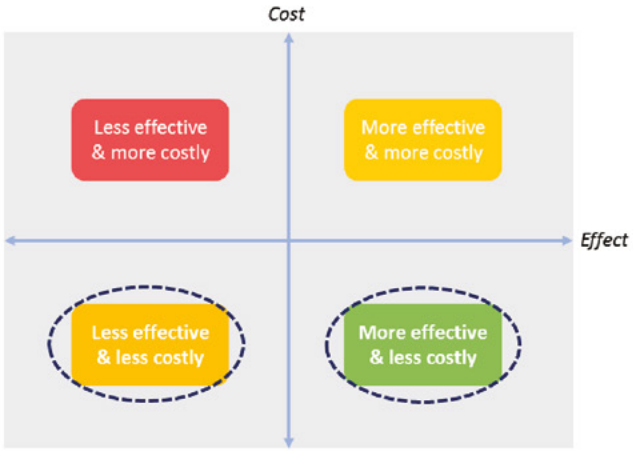
Adoption and impact

No specific adoption issues were identified through the META tool process beyond areas covered in other sections of this report (see specifically section on evidence requirements for digital health technologies).

The company do not think that specialist training is required to use the technology. The technology is designed to be deployed as a background process i.e. it will not require a user to trigger the process. However, there is likely to be some lower level training required to ensure that users are capable and comfortable with using the technology and that they understand the outputs.

The company do not think that the technology will require specific data storage arrangements. However, note that consideration needs to be given to how patient data is stored and used by the application to ensure that it complies with data governance and protection laws.

Value proposition graph



At present, there is limited evidence on the clinical and economic impact of JivaRDX. Therefore, there is uncertainty around the net impact that JivaRDX may have on costs and QALYs. However, based on our discussions and the evidence collected to date, the two most likely value propositions for the technology are as follows:

1. JivaRDX in combination with radiologist assessment is less costly and more effective than radiologist assessment alone
2. JivaRDX in combination with radiologist assessment is less costly and less effective than radiologist assessment alone

In both cases, JivaRDX in combination with radiologist assessment is predicted to be less costly than radiologist assessment alone. This judgement is based on the high specificity of JivaRDX shown in the preliminary evidence collected to date, which would translate to fewer 'unnecessary' biopsies of false positive mpMRI results. The scale of the improvement in specificity compared to clinical practice is so high that it seems likely that the cost savings through reduced biopsies would outweigh the additional upfront cost of JivaRDX.

The impact of JivaRDX on health outcomes is less certain. As is typical for diagnostic interventions, the likely impact on health outcomes is likely to be primarily driven by improvements in sensitivity as the key health gains are seen by detecting cancer at an earlier stage. Based on the preliminary evidence, JivaRDX appears to be better in both accuracy domains (sensitivity and specificity) and therefore it is plausible that it could be more effective as well as less costly (i.e. dominant). Indeed, the preliminary economic analysis concluded exactly this in its base case analysis. However, the analysis also concluded that there was uncertainty around the impact of JivaRDX on health outcomes. This uncertainty is unsurprising in some ways as the difference in sensitivity estimates is more marginal than the large gain in specificity and therefore when considering variation based on confidence intervals, it is likely that there could be instances where sensitivity is better with radiologist interpretation than JivaRDX.

While the above value propositions seem the most likely based on the evidence to date, it will be important to reconsider this position when there is an estimation of the combined accuracy of JivaRDX and radiologist assessment rather than JivaRDX alone. As outlined above, sensitivity and specificity are crucial in determining both

Conclusion

The technology does not currently have a CE mark and this will need to be obtained to enable payers, NICE, HTW, or other Health Technology Assessment bodies to appropriately consider the evidence for its clinical and cost effectiveness. There is a credible value proposition for JivaRDX but there is limited evidence, at present, to support a full health technology assessment. If pursuing a technology appraisal of JivaRDX, then Jiva.ai are encouraged to develop plans for gathering additional evidence.

Next steps

The company should consider adding details for the technology to HealthTech Connect in order to explore further options for development support for the technology or submission to NICE.

The technology does not currently have a UKCA or CE mark and this/these will need to be obtained to enable payers or NICE to appropriately consider the evidence for its clinical and cost effectiveness.

The company should approach one of the UK notified bodies permitted to issue UKCA certification - https://www.gov.uk/government/publications/medical-devices-uk-approved-bodies/uk-approved-bodies-for-medical-devices

The company is recommended to consider the gaps as outlined in this report and seek ways to generate evidence to address them.

The Academic Health Sciences Networks, approached directly or through HealthTech Connect, may provide further information and support for evidence development for the technology.

UK regulators

MHRA Innovation Office - http://bit.do/MHRA-innovation-office

BSI Overview of European Directives, notified bodies and CE mark process - http://bit.do/BSI-md-notified-body-guide

Medical Devices: UK Approved Bodies (for UK conformity assessment) - https://www.gov.uk/government/publications/medical-devices-uk-approved-bodies

MHRA guidance on medical device stand-alone software including apps - http://bit.do/MHRA-stand-alone-software

US regulators

Clinical investigations

MHRA clinical investigations of medical devices – guidance for manufacturers - http://bit.do/MHRA-manufacturers-guidance

MHRA clinical investigations of medical devices – statistical considerations - http://bit.do/MHRA-statistical-considerations

Oxford PROM Group - http://bit.do/Ox-PROM-Grp

NIHR toolkit - http://www.ct-toolkit.ac.uk/

Economic evaluations/HTA – checklists and models

Drummond appraisal checklist for economic evaluations - https://www.nlm.nih.gov/nichsr/edu/healthecon/drummond_list.html

INHAHTA checklist - http://bit.do/INHAHTA-checklist

Resource use and cost

NHS Costing Manual (services and tariffs) - https://www.gov.uk/government/publications/nhs-costing-manual

National Tariff Payment System - https://improvement.nhs.uk/resources/national-tariff/

Disclaimer

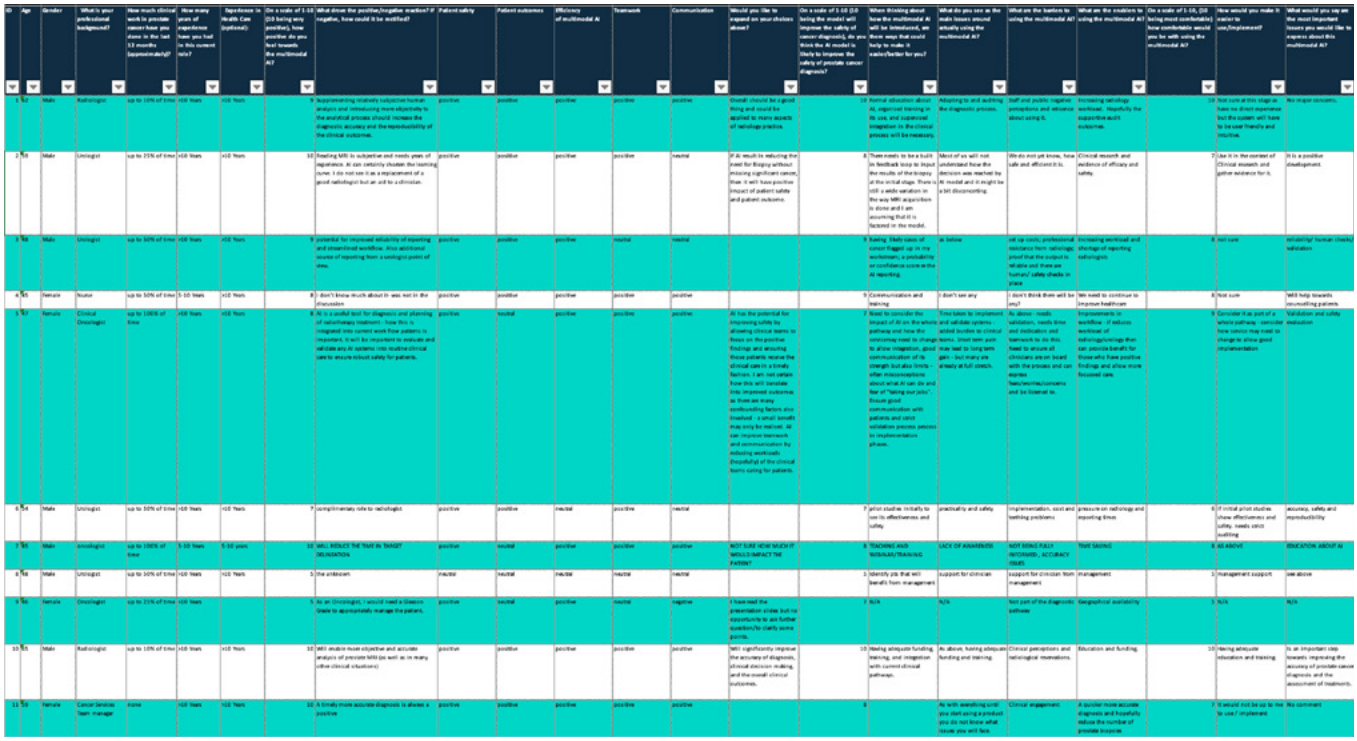
Companies should be aware that the opinions provided in the facilitated assessment and in the resulting report cannot be taken as expert advice or as indicative or suggestive of any future position, and will not be regarded as relevant to any future decision that may be taken by NICE in its role of evaluating products for use in the NHS or wider health arena.

The opinions provided are based on scientific knowledge publicly available at the time of the consultation, and will not account for future changes and developments in scientific knowledge or regulatory requirements. Companies should be aware that the opinions provided are expressed without any form of warrantee or guarantee. All opinions are presented solely upon the information provided and available at the time of consideration and are solely for the Company's use.

Appendix 4 – Patient data template, June 2022

	Blinded	PSA score prior to MRI	PSA Density	JIVA Outcome March 2022	1 signifies clinically significant cancer is detected and a 0 signifies no clinically significant cancer detected. 0 is Likert score 1-2 and 1 is Likert score 3,4&5 March 2022	Actual equivalent	Matches original Likert Score Yes / No	Result		JIVA Outcome June 2022	1 signifies clinically significant cancer is detected and a 0 signifies no clinically significant cancer detected. 0 is Likert score 1-2 and 1 is Likert score 3,4&5 June 2022	Actual equivalent	Matches original Likert Score True / False	Result		
	Likert score				JIVA AI generated Test result	Reality					JIVA AI generated Test result	Reality				
Patient1	5	19.2	0.188	Patient1	0	1	False	False Negative		Patient1	1	1	True	True Positive		
Patient2	3	4.3	0.108	Patient2	1	1	True	True Positive		Patient2	1	1	True	True Positive		
Patient3	2	8.4	0.056	Patient3	1	0	False	False Positive		Patient3	1	0	False	False Positive		
Patient4	3	10.4	0.260	Patient4	0	1	False	False Negative		Patient4	1	1	True	True Positive		
Patient5	3	18.4	0.409	Patient5	1	1	True	True Positive		Patient5	1	1	True	True Positive		
Patient6	2	16	0.160	Patient6	1	0	False	False Positive		Patient6	0	0	True	True Negative		
Patient7	3	6.1	0.161	Patient7	0	1	False	False Negative		Patient7	1	1	True	True Positive		
Patient8	4	6.1	0.153	Patient8	1	1	True	True Positive		Patient8	1	1	True	True Positive		
Patient9	5	70.8	1.416	Patient9	1	1	True	True Positive		Patient9	1	1	True	True Positive		
Patient10	1	6.5	0.130	Patient10	1	0	False	False Positive		Patient10	1	0	False	False Positive		
Patient11	2	3.5	0.059	Patient11	1	0	False	False Positive		Patient11	1	0	False	False Positive		
Patient12	4	11.3	0.182	Patient12	1	1	True	True Positive		Patient12	1	1	True	True Positive		
Patient13	4	54.8	1.566	Patient13	0	1	False	False Negative		Patient13	1	1	True	True Positive		
Patient14	2	6.1	0.111	Patient14	1	0	False	False Positive		Patient14	1	0	False	False Positive		
~~Patient15~~	~~3~~	~~7.1~~	~~0.093~~	~~Patient15~~	~~1~~	~~0~~				~~Patient15~~		~~0~~				
~~Patient16~~	~~1~~	~~11.1~~	~~0.128~~	~~Patient16~~	~~1~~	~~0~~				~~Patient16~~		~~0~~				
Patient17	1	3.1	0.044	Patient17	0	0	True	True Negative		Patient17	0	0	True	True Negative		
Patient18	5	48.7	0.974	Patient18	1	1	True	True Positive		Patient18	1	1	True	True Positive		
Patient19	2	7.5	0.150	Patient19	0	0	True	True Negative		Patient19	0	0	True	True Negative		
Patient20	3	10.8	0.193	Patient20	0	1	False	False Negative		Patient20	1	1	True	True Positive		
Patient21	4	9	0.180	Patient21	1	1	True	True Positive		Patient21	1	1	True	True Positive		
Patient22	2	5.4	0.245	Patient22	0	0	True	True Negative		Patient22	0	0	True	True Negative		
Patient23	2	8.3	0.184	Patient23	1	0	False	False Positive		Patient23	1	0	False	False Positive		
Patient24	2	3.5	0.074	Patient24	1	0	False	False Positive		Patient24	1	0	False	False Positive		
Patient25	3	7.7	0.233	Patient25	1	1	True	True Positive		Patient25	1	1	True	True Positive		
Patient26	5	3.7	0.074	Patient26	1	1	True	True Positive		Patient26	1	1	True	True Positive		
Patient27	3	8.5	0.057	Patient27	0	1	False	False Negative		Patient27	1	1	True	True Positive		
Patient28	2	9.7	0.069	Patient28	1	0	False	False Positive		Patient28	1	0	False	False Positive		
Patient29	4	5	0.143	Patient29	1	1	True	True Positive		Patient29	1	1	True	True Positive		
Patient30	4	5	0.132	Patient30	1	1	True	True Positive		Patient30	1	1	True	True Positive		
Patient31	4	3.1	0.103	Patient31	1	1	True	True Positive		Patient31	1	1	True	True Positive		
Patient32	5	14.7	0.420	Patient32	1	1	True	True Positive		Patient32	1	1	True	True Positive		
Patient33	2	5.4	0.083	Patient33	1	0	False	False Positive		Patient33	1	0	False	False Positive		
Patient34	3	5.5	0.092	Patient34	0	1	False	False Negative		Patient34	1	1	True	True Positive		
Patient35	4	7	0.175	Patient35	1	1	True	True Positive		Patient35	1	1	True	True Positive		
Patient36	3	3.7	0.028	Patient36	0	1	False	False Negative		Patient36	1	1	True	True Positive		
Patient37	2	7.5	0.067	Patient37	1	0	False	False Positive		Patient37	1	0	False	False Positive		
Patient38	2	6.1	0.102	Patient38	1	0	False	False Positive		Patient38	1	0	False	False Positive		
Patient39	2	4.9	0.089	Patient39	1	0	False	False Positive		Patient39	1	0	False	False Positive		
Patient40	3	4.3	0.096	Patient40	0	1	False	False Negative		Patient40	1	1	True	True Positive		
Patient41	2	7.2	0.131	Patient41	1	0	False	False Positive		Patient41	1	0	False	False Positive		
Patient42	4	8.4	0.240	Patient42	1	1	True	True Positive		Patient42	1	1	True	True Positive		
Patient43	3	5.6	0.072	Patient43	1	1	True	True Positive		Patient43	0	1	False	False Negative		
Patient44	3	4.4	0.088	Patient44	1	1	True	True Positive		Patient44	1	1	True	True Positive		
Patient45	2	7.2	0.103	Patient45	0	0	True	True Negative		Patient45	1	0	False	False Positive		
Patient46	2	5	0.067	Patient46	1	0	False	False Positive		Patient46	1	0	False	False Positive		
								SENSITIVITY (MARCH)		**65%**				**SENSITIVITY (JUNE)**		**96%**
								TRUE POSITIVE		17				TRUE POSITIVE		25
								FALSE NEGATIVE		9				FALSE NEGATIVE		1
								SPECIFICITY		**22%**				**SPECIFICITY**		**22%**
								TRUE NEGATIVE		4				TRUE NEGATIVE		4
								FALSE POSITIVE		14				FALSE POSITIVE		14

Appendix 5 - Multimodal AI analysis of Prostate Cancer Staff Questionnaire



Multimodal AI analysis of Prostate Cancer

Introduction: this survey intends to assess the reaction of healthcare professionals to the multimodal AI analysis content, format, and use. It may be used to explore:

- usefulness of the multimodal AI
- perceptions of efficacy of the multimodal AI
- format/presentation of the multimodal AI
- content of the multimodal AI
- barriers/solutions to adhering to particular process measures
- barriers/solutions to implementation

* Required

* This form will record your name, please fill your name.

1. Name *

2. Age *

3. Gender *

- ◯ Female
- ◯ Male
- ◯ Non-binary
- ◯ Prefer not to say

4.

What is your professional background? *

- ◯ Nurse
- ◯ Healthcare Assistant
- ◯ Radiologist
- ◯ Urologist
- ◯ Other

5. How much clinical work in prostate cancer have you done in the last 12 months (approximately)? *

- ◯ none
- ◯ up to 10% of time
- ◯ up to 25% of time
- ◯ up to 50% of time
- ◯ up to 100% of time

6.

How many years of experience have you had in this current role? *

- ◯ <1 Year
- ◯ 1-2 Years
- ◯ 2-5 years
- ◯ 5-10 Years
- ◯ >10 Years

7.

Experience in Health Care (optional):

- ◯ <1 Year
- ◯ 1-2 years
- ◯ 2-5 years
- ◯ 5-10 years
- ◯ >10 Years

8. On a scale of 1-10 (10 being very positive), how positive do you feel towards the multimodal AI? *

1	2	3	4	5	6	7	8	9	10
◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

9. What drove the positive/negative reaction? If negative, how could it be rectified? *

10. What do you think about the aims of having the multimodal AI? *

	positive	neutral	negative
Patient safety	◯	◯	◯
Patient outcomes	◯	◯	◯
Efficiency of multimodal AI	◯	◯	◯
Teamwork	◯	◯	◯
Communication	◯	◯	◯

11. Would you like to expand on your choices above?

12. On a scale of 1-10 (10 being the model will improve the safety of cancer diagnosis), do you think the AI model is likely to improve the safety of prostate cancer diagnosis?

1	2	3	4	5	6	7	8	9	10
◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

13. When thinking about how the multimodal AI will be introduced, are there ways that could help to make it easier/better for you?

14. What do you see as the main issues around actually using the multimodal AI? *

15. What are the barriers to using the multimodal AI? *

16. What are the enablers to using the multimodal AI? *

17. On a scale of 1-10, (10 being most comfortable) how comfortable would you be with using the multimodal AI? *

1 2 3 4 5 6 7 8 9 10

18. How would you make it easier to use/implement? *

19. What would you say are the most important issues you would like to express about this multimodal AI? *

20. Would you be happy to discuss your responses further? If yes, please leave your email below. *

Appendix 6 - Multimodal AI analysis of Prostate Cancer Patient Questionnaire

The use of Artificial Intelligence to support the diagnosis of Prostate Cancer.

Professor Chris Hopkins
Email : chris.hopkins@wales.nhs.uk

Introduction:
Artificial intelligence (AI) and machine learning solutions are transforming the way healthcare is being delivered. Health boards across Wales have accumulated vast data sets in the form of health records and images and AI technologies are well suited to analyse this data and uncover patterns and insights that humans could not find on their own.

This survey intends to assess the reaction of those patients with prostate cancer to our multimodal AI analysis content, format, and use. Multimodal AI in prostate cancer is a software platform that looks at various data points on an MRI image to support a diagnosis of prostate cancer. The survey will be used to explore:

- usefulness of the multimodal AI, and
- perceptions of the multimodal AI

* Required

1. Name

2. Age *

3. On a scale of 1-10 (10 being very positive), how positive do you feel towards the idea of using a computer software platform (Multimodal AI) to support decision making on prostate cancer ? *

1	2	3	4	5	6	7	8	9	10
◯	◯	◯	◯	◯	◯	◯	◯	◯	◯

4. Why do you say this? If negative, how could it be improved? *

5. What do you think about the aims of having the multimodal AI? *

	positive	neutral	negative
Safety	◯	◯	◯
Treatment outcomes	◯	◯	◯
Diagnosis	◯	◯	◯
Monitoring	◯	◯	◯

6. Would you like to expand on your choices above?

7. On a scale of 1-10 (10 being the model will improve the safety of cancer diagnosis), do you think the AI model is likely to improve the safety of prostate cancer diagnosis?

1 2 3 4 5 6 7 8 9 10

8. When thinking about how the multimodal AI will be introduced, are there ways that could help to make it easier/better for you?

9. Do you see any issues around actually using the multimodal AI? *

10. Would you want AI to be used as part of your care? Why is this? *

11. On a scale of 1-10, (10 being most comfortable) how comfortable are you with your clinician using the multimodal AI in your care? *

1 2 3 4 5 6 7 8 9 10

12. What would you say are the most important issues you would like to express about this multimodal AI? *

13. Would you be happy to discuss your responses further? If yes, please leave your email below. *

ID	Age	On a scale of 1-10 (10 being very positive), how positive do you feel towards the idea of using a computer software platform (Multimodal AI) to support decision making on prostate cancer ?	Why do you say this? If negative, how could it be improved?	Safety	Treatment outcomes	Diagnosis	Monitoring	Would you like to expand on your choices above?	On a scale of 1-10 (10 being the model will improve the safety of cancer diagnosis), do you think the AI model is likely to improve the safety of prostate cancer diagnosis?	When thinking about how the multimodal AI will be introduced, are there ways that could help to make it easier/better for you?	Do you see any issues around actually using the multimodal AI?	Would you want AI to be used as part of your care? Why is this?	On a scale of 1-10, (10 being most comfortable) how comfortable are you with your clinician using the multimodal AI in your care?	What would you say are the most important issues you would like to express about this multimodal AI?
1	66	8	I am fine with it's use	positive	positive	positive	positive	Mo	8		No	Yes it can only be positive	8	Who has the final say in my treatment.
2	43	9	Quickly identify patterns and correlations based on vast data resources, to augment physician's decision making, to provide early diagnosis, and support in an optimal personalized prognosis for a particular patient's demographics, health condition and medical history.	positive	positive	positive	positive		9		Transparency and an ability to trace-back and explain how a certain outcome was calculated.	Yes, personalized medicine - find the best modality which has the highest probability of producing results for my particular health and bodily condition	9	comments same as in #9
3	76	10	I am sure using AI can only help decision making	positive	positive	positive	positive	Nothing to add	10	Nothing at the moment - I would need more information on the process to answer this question	I'm not aware of any issues at the moment	Yes, if it could add to my care.	10	I have no issues at the moment with my limited knowledge of multi AI

References

1. 1.https://gco.iarc.fr/tomorrow/en/dataviz/bars?types=0&sexes=1&mode=population&group_populations=0&multiple_populations=1&multiple_cancers=1&cancers=27&populations=826&group_cancers=1&years=2030

2. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(16)32401-1/fulltext

3. ProtecT study https://www.baus.org.uk/_userfiles/pages/files/13-30%20BAUS%20ProtecT%20Nov%202016.pdf

4. NICE guidance NG131 https://www.nice.org.uk/guidance/ng131/resources/prostate-cancer-diagnosisand-management-pdf-66141714312133

5. Overdiagnosis and Overtreatment of Prostate Cancer https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4113338/

6. RCR Overdiagnosis and Overtreatment of Prostate Cancer https://www.rcr.ac.uk/sites/default/files/rcrcqc-radiology-review-where-are-we-now.pdf

7. RCR Clinical Radiology Workforce Report https://www.rcr.ac.uk/system/files/publication/field_publication_files/clinical-radiology-uk-workforcecensus-report-2018.pdf

8. AI in medical imaging may make the biggest impact in healthcare https://healthmanagement.org/c/hospital/issuearticle/ai-in-medical-imaging-may-make-the-biggestimpact-in-healthcare

9. A systematic review of the diagnostic accuracy of prostate specific antigen https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2753579/

10. Coronavirus: Fears 2,300 people a week may be missing cancer diagnoses and put off seeking help https://inews.co.uk/news/coronavirus-cancer-patients-referrals-fall-treatment-delays-420957

11. NICE guideline NG131 Prostate cancer: diagnosis and management – intervention comparisons https://www.nice.org.uk/guidance/ng131/evidence/g-active-surveillance-radical-prostatectomy-or-radicalradiotherapy-in-people-with-localised-prostate-cancer-pdf-6779081780

12. Prediction models in cancer care https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3189416/

13. A systematic review on multiparametric MR imaging in prostate cancer detection https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5663098/

14. New data shows men wait four times longer to be diagnosed with prostate cancer than women with breast cancer https://prostatecanceruk.org/about-us/news-and-views/2017/12/new-data-shows-menwait-four-times-longer-to-be-diagnosed-with-prostate-cancer-than-women-with-breast-cancer

15. Artificial intelligence at the intersection of pathology and radiology in prostate cancer https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6521904

16. Recommendations | Prostate cancer: diagnosis and management | Guidance | NICE. Available at: https://www.nice.org.uk/guidance/ng131/chapter/Recommendations

17. Ahmed, H.U. et al., 2017. Diagnostic accuracy of multi-parametric MRI and TRUS biopsy in prostate cancer (PROMIS): a paired validating confirmatory study. The Lancet, 389(10071), pp.815–822.

18. Faria, R. et al., 2018. Optimising the Diagnosis of Prostate Cancer in the Era of Multiparametric Magnetic Resonance Imaging: A Cost-effectiveness Analysis Based on the Prostate MR Imaging Study (PROMIS). European urology, 73(1), pp.23–30.

19. Aydın, Abdullatif & Kahir, C. & Patel, M. & Allen, Clare & Dasgupta, P.. (2019). Detection of clinically relevant prostate cancer from multi-parametric 3Tesla MRI scans using artificial intelligence. European Urology Supplements. 18. e2722. 10.1016/S1569-9056(19)32819-https://www.sciencedirect.com/science/article/abs/pii/S1569905619328192?via%3Dihub

TRITECH

Sefydliad | Institute

Hywel Dda University Health Board

TriTech Institute
Unit 2 Dura Park
Bynea
SA14 9TD

0300-303-6115
Tritech.HDD@wales.nhs.uk

tritech.nhs.wales @tritech_hd Tritech Institute