A European Cancer Image Platform Linked to Biological and Health Data for Next-Generation Artificial Intelligence and Precision Medicine in Oncology

Deliverable D6.5:

Tool for cost-effectiveness assessment of AI solutions

Reference	**D6.5_ EuCanImage_ONCO_28092023**
Lead Beneficiary	Radiomics
Author(s)	Radiomics
Dissemination level	Public
Type	Other
Official Delivery Date	28th September 2023
Date of validation of the WP leader	27th September 2023
Date of validation by the Project Coordinator	27th September 2023
Project Coordinator Signature	

EuCanImage is funded by the European Union's H2020 Framework
Under Grant Agreement No 952103

Version log

Issue Date	Version	Involved	Comments
14/07/2023	1	Radiomics	
28/08/2023	2	Radiomics	Revision of draft 1
25/09/2023	3	Radiomics	Revision of draft 2
27/09/2023	4	Xènia Puig, Oliver Díaz, Karim Lekadir	Revision of draft 3 & final approval

Executive Summary

The objective of this report is to describe the methods (health economic models and data input) used to assess the potential cost-effectiveness of AI in three different cancer use cases related to three different organs: liver, rectum and breast.

1. The detection and identification of liver lesions with AI improved Magnetic Resonance Imaging (MRI) in the surveillance for hepatocellular cancer (HCC) in patients with a cirrhotic liver;
2. The prediction of the level of response to neoadjuvant radio(chemo)therapy based on primary MRI in patients with locally advanced rectal cancer (LARC);
3. Improving the accuracy of mammograms in the screening of breast cancer by differentiating automatically and with high accuracy benign from malignant tumours.

For each use case, the current context and the potential role and benefits of AI are described. In patients with cirrhosis, at risk for HCC, AI can improve the accuracy of the surveillance for HCC, better detect small lesions and better distinguish malignant from non-malignant lesions. In patients with LARC, AI can contribute to a more personalized approach by better identifying which patients can benefit from neoadjuvant treatment and from surgical intervention, and which patients will need adjuvant therapy and more or less intense post-treatment follow up. In the screening for breast cancer, AI can improve accuracy of screening hence avoiding false positive and false negative results. For each of the use cases, the features of the health economic model that needs to be applied for assessing the cost-effectiveness of AI have been worked out, and overviews of the key clinical, epidemiological, cost and utility data (the latter to allow calculating quality adjusted life years) are provided. The report ends with a discussion on the data collection challenges and the required next steps.

Table of Contents

Acronyms

Name	Abbreviation
Alpha-fetoprotein	AFP
Artificial Intelligence	AI
Abbreviated Magnetic Resonance Imaging	AMRI
Barcelona Clinic Liver Cancer	BCLC
Contrast Enhanced Ultrasound	CEUS
Chemoradiotherapy	CRT
Disability Adjusted Life Years	DALYs
Digital Breast Tomosynthesis	DBT
Ductal Carcinoma in Situ	DCIS
Hepatocellular Carcinoma	HCC
Liver Imaging Reporting and Data System	LI-RADS
Locally Advanced Rectal Cancer	LARC
Low Anterior Resection Syndrome	LARS
Magnetic Resonance Imaging	MRI
Quality Adjusted Life Years	QALYs
Radiofrequency Ablation	RFA
Robotic Rectal Resection	RRR
Short-Course Radiotherapy	SCRT
Trans Arterial Chemoembolization	TACE
Total Mesorectal Excision	TME
Total Neoadjuvant Treatment	TNT
Ultrasound	US
Watch-and-Wait	WW

1 Introduction

The use of Artificial Intelligence (AI) in healthcare is increasing rapidly. AI finds more and more applications in diverse disease areas, and can contribute to better predicting, diagnosing, treating and monitoring patients (Davenport et al. 2019).

AI applications have also made enormous progress in the screening, diagnosis, treatment and follow up of cancer, and are already widely used in in several types of cancer (Mitsala et al, 2021).

As healthcare budgets are limited, new technologies are increasingly assessed on their cost-effectiveness. Indeed, in more and more jurisdictions, only if new technologies offer value for money, policymakers can decide to reimburse them with public money.

Within the EUCanImage consortium it is therefore of utmost importance to understand the potential cost-effectiveness of AI tools in a variety of applications in cancer.

The objective of this report is to describe the methods (health economic models and data input) that will be used to assess the potential cost-effectiveness of AI-based tools in different use cases. As such, the report represents an extensive study protocol describing those methods and their rationale.

Currently three use cases have been selected related to three different organs: liver, rectum and breast.

1. The detection and identification of liver lesions with AI improved Magnetic Resonance Imaging (MRI) in the surveillance for hepatocellular cancer (HCC) in patients with a cirrhotic liver;
2. The prediction of the level of response to neoadjuvant radio(chemo)therapy based on primary MRI in patients with locally advanced rectal cancer;
3. Improving the accuracy of mammograms in the screening of breast cancer by differentiating automatically and with high accuracy benign from malignant tumours.

These three use cases have been selected not only because of addressing different organs but also because of the different nature of the benefits of AI, as will become clear in the remainder of the report.

Obviously AI itself is undergoing continuous improvement and therefore the assessment of its potential cost-effectiveness in the three use cases will be subject to several assumptions (for instance about its accuracy). It should also be acknowledged that clinical diagnosis is both an art and a science, and is more challenging for AI to optimize than visual diagnostic interpretation, such as radiographic and pathologic diagnosis (Kulkarni and Singh, 2023).

The following chapters discuss the 3 use cases, their respective current context and the potential role of AI, the proposed modelling methods, and the required data inputs.

2 Detecting liver lesions in the cirrhotic liver

2.1 Current context and potential role of AI

Hepatocellular carcinoma (HCC) is one of the most common tumours in the world. It is the first primary malignant liver tumour and the second most common cause of cancer-related death (Renzulli et al. 2022).

The majority of HCCs occur in the setting of *liver cirrhosis*, which represents the final evolution stage of all chronic progressive liver diseases. Cirrhotic patients have indeed an annual risk to develop HCC of between 1 and 5% depending on the underlying aetiology and patient demographics and therefore surveillance of the liver is warranted in these patients (Villanueva, 2019, European Association for the Study of the Liver, 2018).

Different clinical practice guidelines have recommended performing surveillance in cirrhotic patients via ultrasound (US) in combination with serum alpha-fetoprotein (AFP) every 6 months, despite the fact that US is characterized by low and highly variable sensitivity. Indeed, the sensitivity of US in the detection of early-stage HCC has been reported to range from 40% to 80% and a large meta-analysis comprising 13,367 patients found the sensitivity of US alone for the detection of early-stage disease HCC to be only 47% (Barnard-Giustini et al. 2023).

Hence, the recommendations putting forward US are based rather on the rationale that it is more easily accessible and much cheaper than MRI than on its performance (Renzulli et al. 2022). (See Box 1 for basic explanations on sensitivity, specificity and related parameters).

Box 1: Basic explanation of test performance concepts

Test Result ↓/ Disease status →	Has the disease	Does not have the disease	TOTALS
Test positive	True Positives (TP)	False Positives (FP)	Total Positive tests
Test negative	False Negatives (FN)	True Negatives (TN)	Total Negative Tests
TOTALS	Total with disease	Total without disease	

Sensitivity = TP/(TP+FN) → If one has the disease, what is the probability to test positive for it?

Specificity = TN/(FP+TN) → If one does NOT have the disease, what is the probability to test negative for it?

Positive Predictive Value (PPV) = TP/(TP+FP) → if one tests positive, what is the probability to have the disease

Negative Predictive Value (NPV) = TN/(FN+TN) → if one tests negative, what is the probability to NOT have the disease

Source: https://uk.cochrane.org/news/sensitivity-and-specificity-explained-cochrane-uk-trainees-blog

Newer MRI techniques use shorter number of MRI sequences to simplify the interpretation of images and thus abbreviate the acquisition time to about 15 minutes in average. This abbreviated MRI (AMRI) is thus cheaper than traditional MRI and offers much better sensitivity than US (Barnard-Giustini et al. 2023).

In the meantime, progress has also been made with US, whereby nowadays contrast enhanced US (CEUS) is more and more used (Adeniji and Dhanasekaran, 2021). However, several studies have suggested that with CEUS, HCC could be hardly distinguished from some other non-HCC malignancies, thus leading to inappropriate clinical strategy (Li et al. 2021).

Therefore, the current state of the art is to apply contrast-enhanced pattern MRI in combination with the use of reporting diagnostic guidelines such as LI-RADS (Liver Imaging Reporting and Data System).

Another field of progress is to apply biomarkers either to better identify those patients with increased risk for developing HCC and therefore needing more intensive surveillance or to improve the diagnostic accuracy of the surveillance techniques (Barnard-Giustini et al. 2023).

The Figure 1 depicts well all the elements at stake in this field with an emphasis on the current gaps.

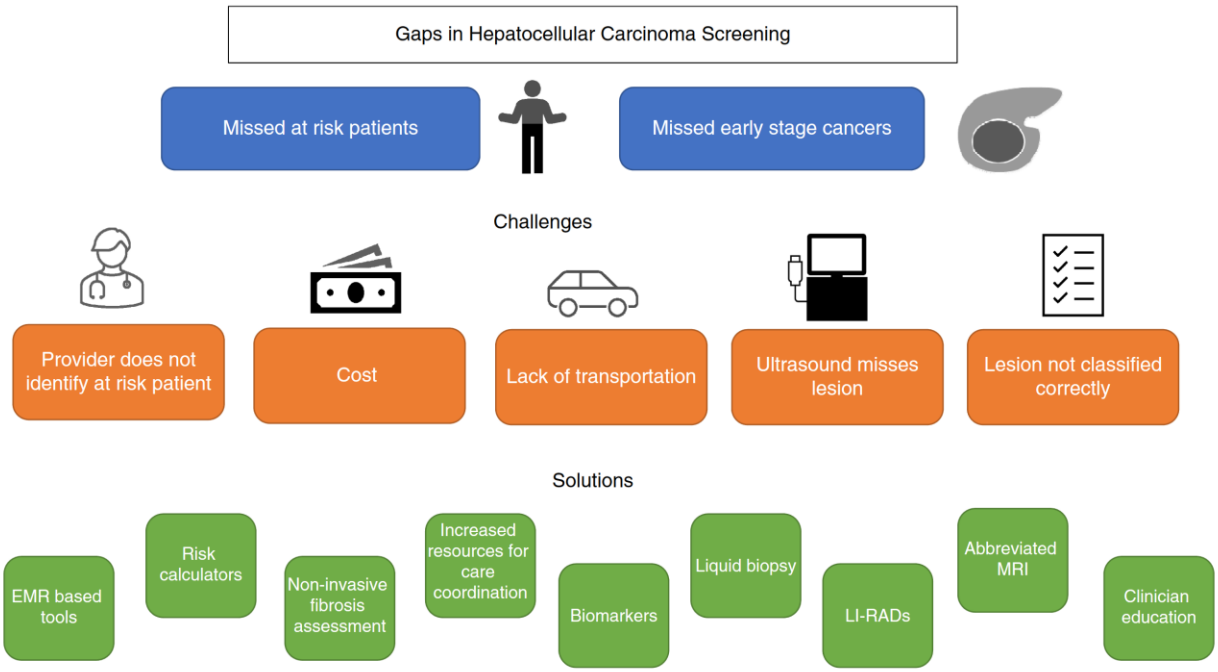


Figure 1: Illustration of the gaps in HCC surveillance/screening (Barnard-Giustini et al. 2023).

Of note, the presence of fibrous and regenerative tissue in the liver causes the distortion of normal liver parenchyma, changing the typical appearances of benign lesions and pseudolesions, which therefore may be misinterpreted as malignancies. Obviously, a correct distinction between pseudolesions and malignancy is crucial to allow appropriate targeted therapy and avoid treatment delays (Renzulli et al. 2022b). But despite the progress already made, the diagnosis of HCC – and especially the characterization of lesions smaller than 2 cm – continues to be a radiological challenge (Vogl et al. 2022).

The promise of AI in this setting is that when added to MRI the anticipated better accuracy will make it more likely to identify small lesions (<2cm) as well as to better make a distinction between pseudolesions and malignancies. This will then be leading to quicker and better decisions to treat patients, resulting in better outcomes in terms of quality of life and survival.

2.2 A model to assess the cost-effectiveness of AI in the surveillance of cirrhotic patients at risk for HCC

The proposed model as described here has been inspired on several publications that addressed cost-effectiveness of surveillance in patients at risk for HCC.

In cost-effectiveness analyses, the cost of an intervention is balanced with potential cost-offsets (for instance avoided hospitalisations) to result in a net cost of the intervention (called the 'incremental cost'). This incremental cost is then balanced with the health gain, mostly expressed in Quality Adjusted Life Years (QALYs). If the ratio between the incremental cost and the incremental effect is considered acceptable, then the intervention is called cost-effective. Some countries like the UK and The Netherlands have published official thresholds (i.e. the maximum willingness to pay for a QALY gained) to guide the interpretation of cost-effectiveness.

Farhang Zangneh et al. (2019) report a Markov model to assess the cost-effectiveness of biannual or annual HCC surveillance versus no surveillance in patients with Hepatitis C Virus (HCV) infection and advanced fibrosis after a Sustained Virologic Response (SVR) to therapy. In a Markov model, the population in the considered virtual cohort can be at any time in different health or disease states and can then make transitions over time (for instance from cirrhosis to HCC) with given probabilities per fixed period of time. The authors applied a USA healthcare payer perspective and a lifetime time horizon.

In the surveillance arm of the model, patients can transition every month between Markov states. Patients diagnosed with HCC receive treatment based on the Milan criteria and the Barcelona Clinic Liver Cancer staging systems (BCLC). The distribution of patients to different BCLC stages at diagnosis changes with each surveillance cycle to reflect tumour progression over time. The paper also considers patients who are not adherent to surveillance. These patients only undergo diagnostic tests when they develop HCC-related symptoms or have an HCC diagnosed incidentally through imaging for other purposes.

Patients with a false positive HCC diagnosis are managed identically to patients with a true-positive diagnosis for initial therapy before they are recognized as having a false-positive diagnosis.

Cadier et al. (2017) also constructed a Markov model in patients with compensated cirrhosis, comparing surveillance according to the published recommendations ("gold-standard monitoring") to what is observed in real life ("real-life monitoring"). The time horizon was 10 years and the study was conducted from the French and USA healthcare systems. Patients entered the model in the "Compensated cirrhosis" state. Upon being diagnosed with nodules or with liver cancer, they make a transition to respectively the "Nonmalignant nodules" and "HCC" states. Curative treatments of HCC included surgical Liver Resection, percutaneous radiofrequency ablation (RFA), and Liver Transplant (LT). Patients requiring palliative care received chemoembolization, systemic therapy (sorafenib), or other palliative care. After liver resection or radiofrequency treatment, patients could transit to "Successfully treated" and then to "Relapse". Patients in "Relapse" could only transit to one of two curative treatments:

percutaneous RFA or LT. Death was an absorbing state combining disease-specific mortality with age- and sex-specific mortality. Each Markov cycle duration was 3 months (hence 4 times 10 = 40 cycles in total). Figure 2 shows a simplified picture of the model.

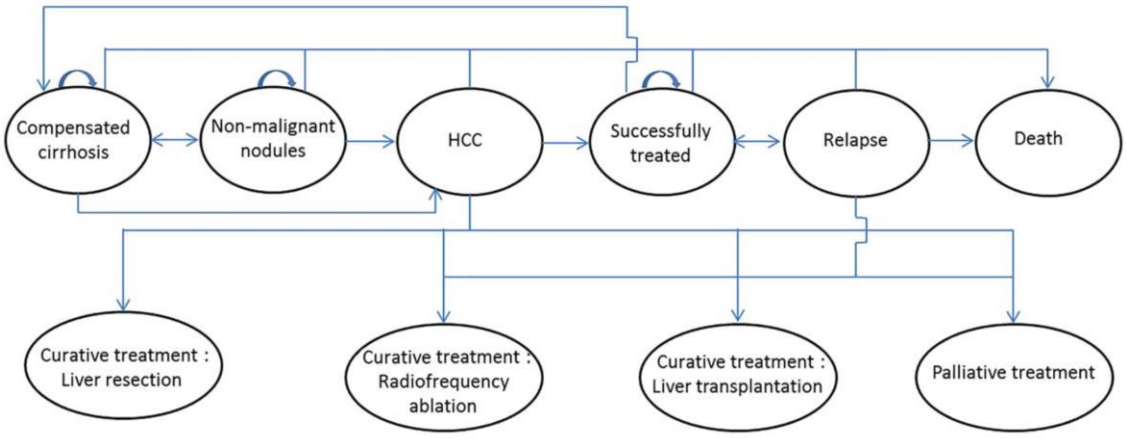


Figure 2: Illustration of the Markov model as applied by Cadier et al (2023).

Although this model is rather straightforward, and nodules are explicitly modelled, the main weaknesses are the absence of explicit modelling of false positives and false negatives and the lack of quality adjustment of the gained life years. A 10 years' time horizon also seems to be rather short giving the young starting age (50 years).

Taylor et al. (2017) conducted a study with a focus on the harms associated with HCC surveillance caused by false positives. The authors applied a simplified Markov model, not detailing the treatments but focussing on the harms due to false positive results leading to avoidable imaging and biopsies. The time horizon was only 5 years.

An interesting approach was reported by Goossens et al. (2017) whereby a Markov model was constructed to simulate *risk-stratified* HCC surveillance strategies in a cohort of 50-year-old subjects with compensated cirrhosis. Patients were stratified into high-risk, intermediate-risk, or low-risk groups by HCC risk biomarker–based scores and assigned to surveillance modalities tailored to HCC risk (2 non-risk-stratified and 14 risk-stratified strategies) and compared with non-stratified biannual ultrasound. The authors concluded that applying magnetic resonance imaging (MRI) and/or ultrasound only in high- and intermediate-risk patients, without screening in low-risk patients, was cost-effective.

Also, Carter et al (2021) investigated the cost-effectiveness of a risk-stratified approach in the Australian setting. The authors applied a societal perspective to account for productivity losses. Three scenarios were tested for patients with compensated cirrhosis: (1) risk-stratified screening for high-risk patients, (2) all-inclusive screening, and (3) no formal screening.

The authors modelled the risk stratification of patients with cirrhosis into those at high versus low to intermediate risk of developing HCC based on a risk scoring called LOS_HCC (Liver Outcomes Score_HCC) and applying a cut-off point of 8, which was found to have the highest accuracy in predicting 5-year HCC incidence, with a sensitivity of 90% and a specificity of

88%. The use of alternative cut-points was tested in a sensitivity analysis. Figure 3 shows the model as applied by Carter et al. (2021).

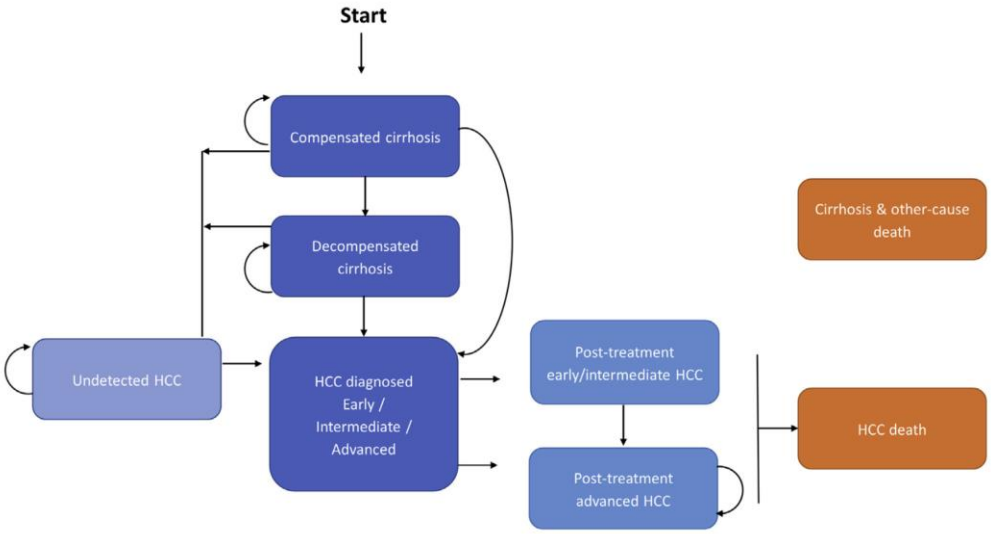


Figure 3: Illustration of the Markov model as applied by Carter et al (2021).

The authors simulated a virtual cohort of subjects with compensated cirrhosis (n = 10,000) followed up with a 6-month cycle for 30 years from a USA healthcare perspective, which seems more adequate than the earlier mentioned short time horizons. The model also distinguished between compensated and decompensated cirrhosis. False positive results were leading to additional imaging and/or biopsies. HCC risk was stable over time during the observation period (which is a weakness).

A more recent evaluation in France (Nahon et al. 2021) describes a Markov model to evaluate the cost-effectiveness of MRI versus US for the detection of very early HCC (Barcelona Clinic Liver Cancer [BCLC] 0) in patients with an annual HCC risk >3%. The authors applied a starting age of 55 years and a 20 years' time horizon, which seems a good compromise between too short and too long simulations. A scoring system was constructed to identify patients with an annual risk >3%. The model emphasizes explicitly the treatment options, i.e. RFA (radiofrequency ablation), TACE (trans arterial chemoembolization), Liver Resection and Liver Transplantation and the study reports their respective probabilities of use, depending on the HCC stage at diagnosis. The study used data from a randomized clinical trial (RCT) related to HCC surveillance and 3 French prospective cohort studies among adults with biopsy-proven compensated cirrhosis. All patients enrolled in those cohorts had periodic liver US surveillance according to international and French guidelines, with or without measurement of alpha-fetoprotein (AFP) serum levels. Sensitivity of US and MRI was accounted for. The model was validated by comparing the predicted numbers of HCC cases with the numbers observed in real life.

Based on the above we propose a Markov model with the following features, as shown in Table 1. It contains structural elements by Carter et al, an explicit modelling of non-malignant nodules, a risk stratification and an explicit modelling of the treatment options for HCC in function of its stage.

Table 1: proposed features of the Markov model to assess the C-Eff of AI in HCC surveillance

Design item	Description	Comment
Target population	*Patients with compensated and decompensated cirrhosis*	*Different prognosis*
Age of the target population	*Average 50 years but with age distribution applied*	*Typical age in the published literature*
Time horizon	*20-30 years*	*Five/ten years is too short. Lifetime requires too many assumptions*
Perspective	*Healthcare and society*	*Benefits to the healthcare system and an impact on productivity are envisaged given the average relatively young age.*
Comparators	*Suboptimal surveillance and golden standard surveillance with US/MRI*	*Additional benefits of AI explicitly modelled*
Risk stratification	*Yes*	*Will be more and more applied. Model needs to be future proof*
Accounting for false positives and negatives	*Yes, explicitly modelled*	*Is at the core of the anticipated benefits of AI*
Markov cycle	*3 months*	*Offers sufficient granularity*
Course of the disease	*Reflect tumour progression over time; explicitly include non-malignant nodules*	*More realistic representation of the real-life setting*
Treatments of HCC	*Modelled explicitly including probabilities of being applied, success rates and relapse rates*	*Enhances model completeness and transparency*

2.3 Clinical and epidemiological data inputs HCC surveillance in patients with liver cirrhosis

In order to populate the described Markov model to assess the potential cost-effectiveness of AI in the surveillance of patients with cirrhosis, several epidemiological and clinical data are required. Based on diverse publications and the proposed features of our Markov model, we established an overview of the key input data, as shown in table 2.

Table 2: key clinical and epidemiological data inputs for Markov model assessing the C-Eff of AI in HCC surveillance

Parameter	Possible sources	Comments
Starting age of the model	*Local epidemiological data*	*Preference for setting the starting age at 50*
Annual average HCC incidence in target group	*Local epidemiological data – cohort data*	*Make distinction dependent on cause of cirrhosis*
Annual average incidence of non-malignant nodules <20mm in target group	*Local epidemiological data – cohort data*	*Make distinction dependent on cause of cirrhosis*
Proportion at high/intermediate/low risk for HCC	*Published literature*	*Depends on applied biomarker*
HCC incidence in different risk groups	*Published literature*	*Will require additional calculations*
Sensitivity and Specificity of applied surveillance methods (US, MRI,…)	*Published literature*	*May require assumptions for AI*
HCC stages at the time of detection – distribution of early, intermediate and advanced HCC	*Published literature*	*Depends on performance of the detection method*
Expected distribution of treatments once HCC is diagnosed, depending on its stage	*Local epidemiological data*	*Can be country-dependent*
Markov transitions - *Compensated to decompensated cirrhosis* - *Decompensated cirrhosis to HCC* - *Early-/intermediate to advanced HCC*	*Published literature*	*Can be country-dependent*
Management options – treatment probabilities - *Resection/ablation for early HCC* - *Liver transplant for early HCC* - *Treatment options for unresectable HCC*	*Local epidemiological data*	*Can be country-dependent – success rates to be included as well*
Mortality rates related to different disease states	*Published literature*	
Natural mortality rates	*Local epidemiological data*	

2.4 Cost and utility data inputs related to HCC surveillance in patients with liver cirrhosis

Finally, Table 3 provides an overview of the key cost and utility data required for the Markov model regarding surveillance of HCC in cirrhotic patients.

Table 3: Key cost and utility data required for the Markov model regarding HCC surveillance in cirrhotic patients

Parameter	***Possible sources***	***Comment***
Unit cost of different imaging techniques	*Local official data*	*Also to include cost of biopsies*
Additional cost of applying AI	*To be assumed*	*Subject to what-if scenarios*
Annual cost of compensated cirrhosis	*Published literature*	*Possibly adjustments needed*
Annual cost of decompensated cirrhosis	*Published literature*	*Possibly adjustments needed*
Annual cost of nodule management	*Published literature*	*Possibly adjustments needed*
Annual cost of HCC	*Published literature*	*Depending on stage*
Annual cost of waitlist management	*Published literature*	*For patients on waitlist for transplant*
Cost of procedures and treatments *- Resection* *- Radiofrequency* *- Transplantation* *- Chemo-embolisation* *- Pharmaceutical treatment; Best supportive care*	*Local official data – Published literature*	*Probabilities of use depend on curative or non-curative (palliative) state (see Table 4).* *Also to include gollow-up costs*
Utilities *- compensated cirrhosis;* *- decompensated cirrhosis* *- HCC* *- on waitlist for liver transplant* *- post-liver transplant* *- chemo & palliative treatment for incurable HCC*	*Published literature*	

3 Prediction of the level of response to neoadjuvant radio(chemo)therapy in locally advanced rectal cancer (LARC)

3.1 Current context and potential role of AI

Colorectal cancer (CRC) is one of the most commonly diagnosed cancers and one of the leading causes of cancer-related death worldwide. Rectal cancers account for approximately 30–35% of all colorectal cancers, and about half of them are diagnosed at a locally advanced stage, i.e., locally advanced rectal cancer (LARC) (De Mattea et al. 2023). LARC is defined by a T3-T4 tumour and/or nodal involvement, with no metastatic sites.

For a while now the medical approach has been based on a multimodal approach of neoadjuvant therapy consisting of a long-course combination of chemotherapy and radiotherapy (CTRT) or short-course radiotherapy (SCRT) followed by resection of the rectum and possibly by adjuvant chemotherapy (Borelli et al. 2023). Since the effect of this approach on reducing metastatic disease has shown to be unsatisfactory, several alternative approaches were investigated moving the adjuvant chemotherapy to the neoadjuvant setting, the so-called "total neoadjuvant treatment (TNT)".

In recent years, TNT has seen tremendous progress (Bourbonne et al. 2023). Nowadays systemic chemotherapy is introduced at an earlier timepoint before surgery, with the main aims to downstage the tumour (Fleming et al. 2022) and to reduce the incidence of distant metastases (Aschele & Glynne-Jones, 2023). TNT represents a kind of "short-cut" in the treatment decision-making for LARC, as delivery of chemotherapy pre-operatively avoids the dilemma of selecting patients for postoperative adjuvant chemotherapy based on baseline MRI high risk features or based on post-treatment histopathology findings.

However, this approach limits the prospect of *personalized* treatment for LARC and it exposes all patients to the adverse events of the neoadjuvant chemotherapy (Bourbonne et al. 2023).

On the other hand, the use of this induction chemotherapy for all might modulate the need for radiotherapy in some groups of patients and could even allow non operative management in some patients, i.e. preserving the rectum (Borelli et al. 2023).

Then again, TNT does not appear to have any effect on local control. Long course chemoradiation (LCCRT) probably continues to be necessary for high-risk patients. Also, the optimal modalities of TNT are still under investigation (Borelli et al. 2023).

Multiple prospective studies have shown that *selective use* of early neoadjuvant treatment in patients with LARC is promising. Indeed, identifying patients as having favourable characteristics on preoperative MRI may enable those patients to undergo *surgery alone* with similar short- and long-term oncologic outcomes as those who undergo neoadjuvant chemoradiotherapy (nCRT) followed by surgery. However, the benefits of decreased radiation-related morbidity and lower costs need to be weighed against the potential increased risk of local recurrence (Mueller et al. 2022).

Therefore, today, it is still of utmost important to know, before or during therapy, which patients would respond and as such chose the best possible patient management pathway.

Borelli et al. summarize the advantages and disadvantages of TNT and organ preservation strategies very nicely in the following Figure (Figure 4).

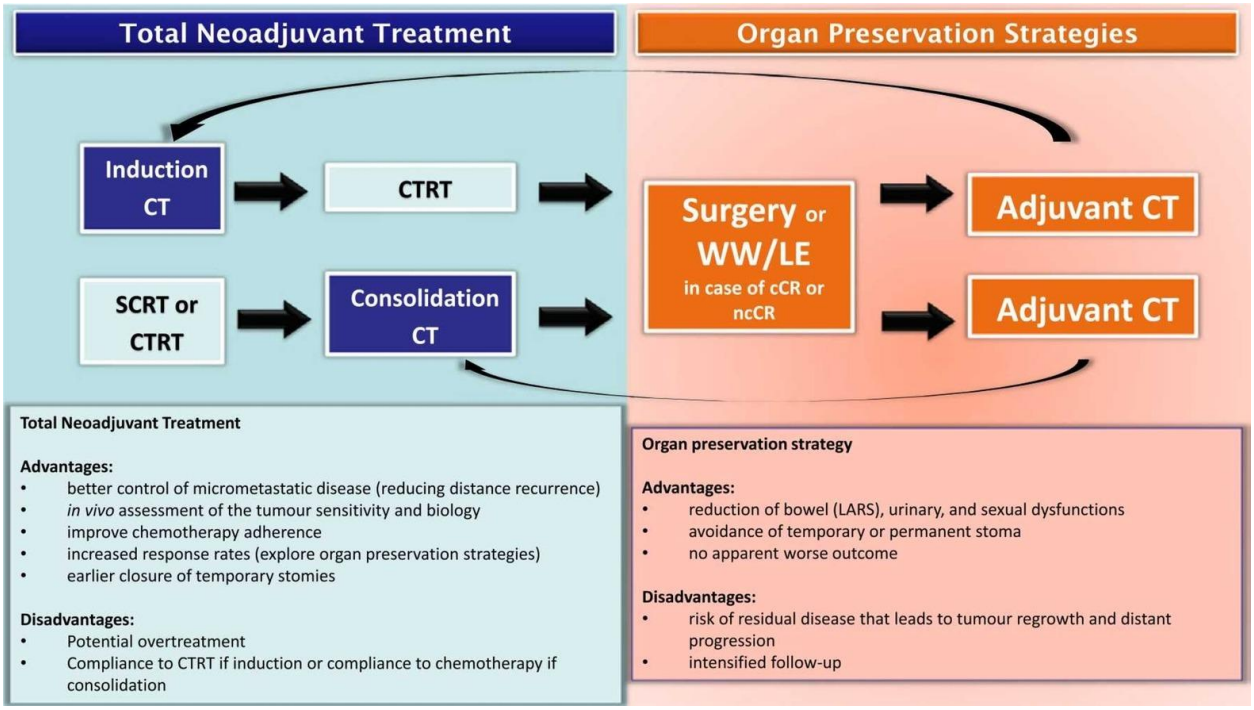


Figure 4: Advantages and disadvantages of TNT and Organ Preservation Strategies (Borelli et al. 2023).

Abbreviations: CT = chemotherapy; CTRT = long course chemo(radio)therapy; SCRT = Short Course Radiotherapy; WW = watchful waiting; LE : local excision; cCR = complete clinical response; ncCR = near-complete clinical response

Biomarker research is also ongoing in this regard (Slipsager et al. 2023). A well-known and promising biomarker in this field is KRAS (De Mattea et al. 2023). Indeed, KRAS mutations proved to be significantly associated with the risk of not achieving pathological complete response after preoperative treatment in LARC.

Moreover, after treatment, technologies are needed that can suggest more intense post-treatment surveillance due to a predicted high risk of a tumour recurrence for particular patients (Inchingolo et al. 2023)

Artificial Intelligence-based radiomics biomarkers have potential clinical implications for adaptive and personalized therapy (Ouyang et al. 2023). Indeed, by enhancing the performance of imaging and as such acting as a biomarker (Inchingolo et al. 2023), AI has the potential to better predict which patients will benefit from neoadjuvant treatment, which tumours are more likely to reduce in volume, and which patients will need adjuvant therapy and more or less intense post-treatment follow up.

3.2 A proposed model to assess the cost-effectiveness of AI in optimizing the management of patients with LARC

The proposed model as described here has been inspired on several publications that addressed cost-effectiveness of different treatment aspects of LARC.

Mueller et al. (2023) report on a model aiming at estimating the cost-effectiveness of a selective use of chemoradiation in patients with LARC. The study was conducted from the societal perspective in the USA. The target population consisted of adult patients with stage II and III rectal cancer. In the selective use arm, 60% of patients were assumed to meet the criteria for primary surgery and 40% requiring surgery after neoadjuvant chemoradiotherapy (nCRT). The selection was based on high resolution MRI staging.

Criteria for postoperative adjuvant therapy in the selective use group were according to the Quicksilver 2 study in which 7.3% of patients in the selective use arm who had upfront surgery required postoperative chemoradiation and adjuvant chemotherapy for poor pathologic features. Thirty percent required adjuvant chemotherapy for poor prognostic features.

Figure 5 shows the decision tree model.

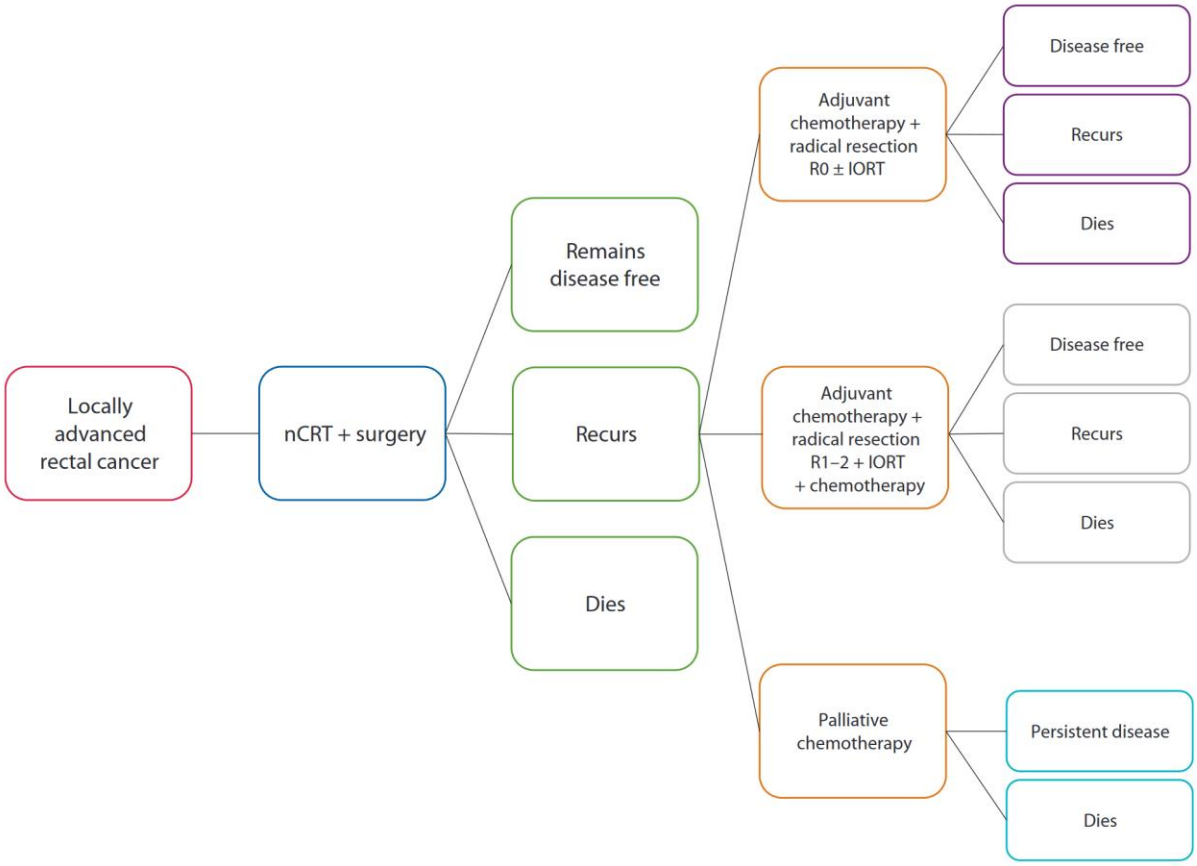


Figure 5: Illustration of the decision tree model as applied by Mueller et al (2023).

nCRT = neoadjuvant chemoradiotherapy; IORT = intra-operative radiotherapy

Applying a 5 years' time horizon the authors concluded that selective use reduces costs and increases QALYs as compared to blanket use.

This is a rather straightforward model paying attention to selective use of TNT, and its impact on the need for adjuvant therapy, but not considering the impact on radiotherapy, the need for surgery, nor the consequences of false positive or false negative outcomes of the imaging applied for selecting patients.

Chin et al. (2022) compared a short-course radiotherapy and total neoadjuvant therapy (SCRT-TNT) followed by total mesorectal excision (TME) with conventional long-course chemoradiotherapy (LCCRT) followed by TME with adjuvant chemotherapy. The analysis was conducted from a USA Medicare payer's perspective. A 5 years' time horizon and a combination of a decision tree and Markov model was applied (see Figure 6). SCRT-TNT was assumed to reduce the need for abdominoperineal resection by 5% and also to positively impact recurrence and metastasis.

The combination of a short-term decision tree with a long-term Markov model seems attractive. The weaknesses of the model are however the short time horizon and no consideration of a personalized strategy.

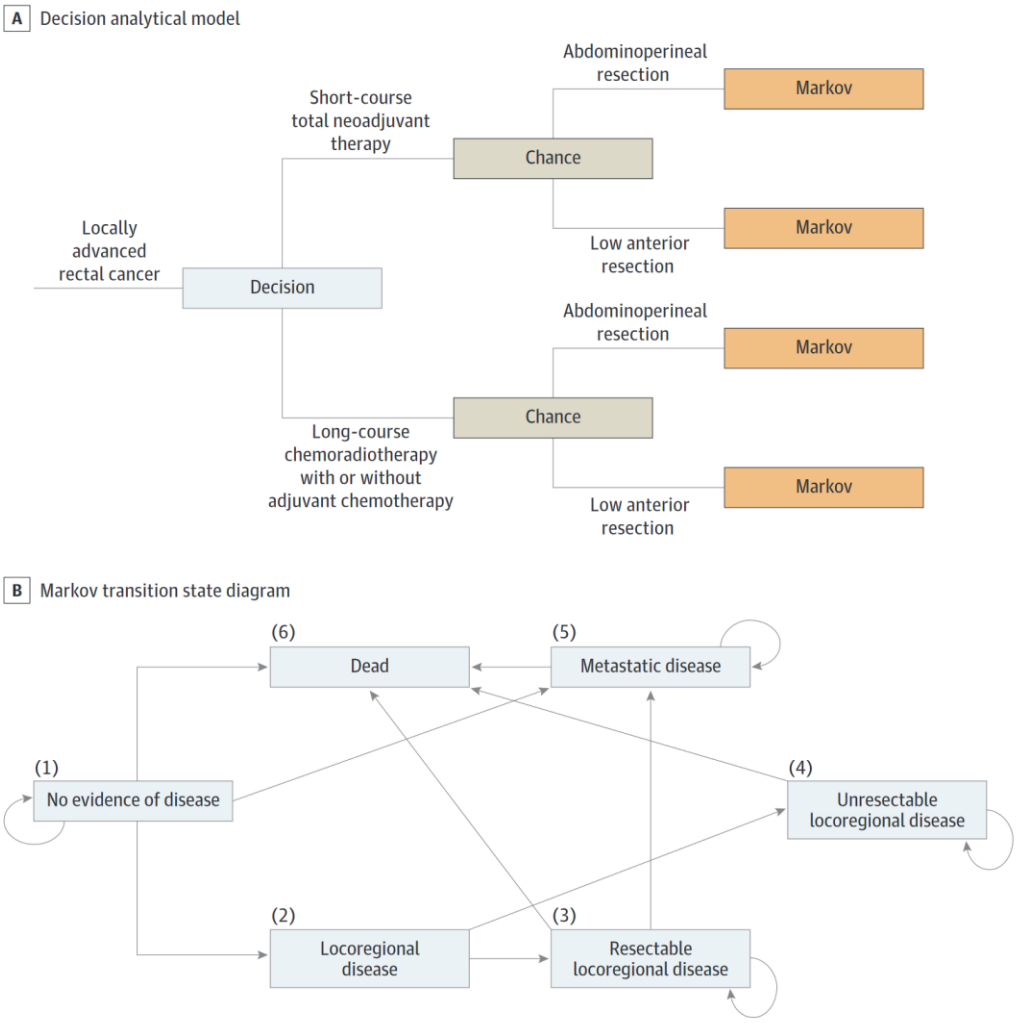


Figure 6: Illustration of the model as applied by Chin et al. (2022)

Rodriguez-Pascal et al. (2022) applied a Markov model to compare standard resection, Robotic Rectal Resection (RRR) and Watch-and-Wait (WW) strategies in patients who had reached clinical complete response to neoadjuvant chemoradiotherapy. A Spanish national Health Service perspective was used and a lifetime time horizon.

Figure 7 presents the structure of the model. Only the Markov chain for the WW arm is shown. The other arms have a similar structure. The advantage of applying a Markov model and especially a longer time horizon is that all relevant costs and outcomes can be captured. Weaknesses of the model are the absence of a personalized approach and the apparent isolated modelling of distant disease (lack of transition between locoregional recurrence and metastatic disease; see Figure 7).

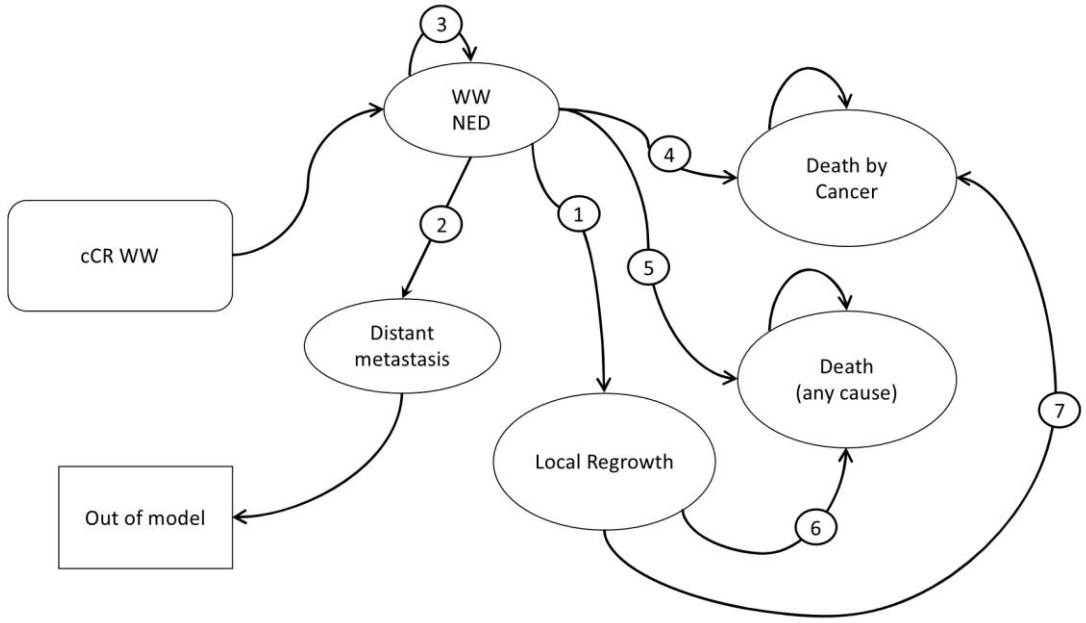
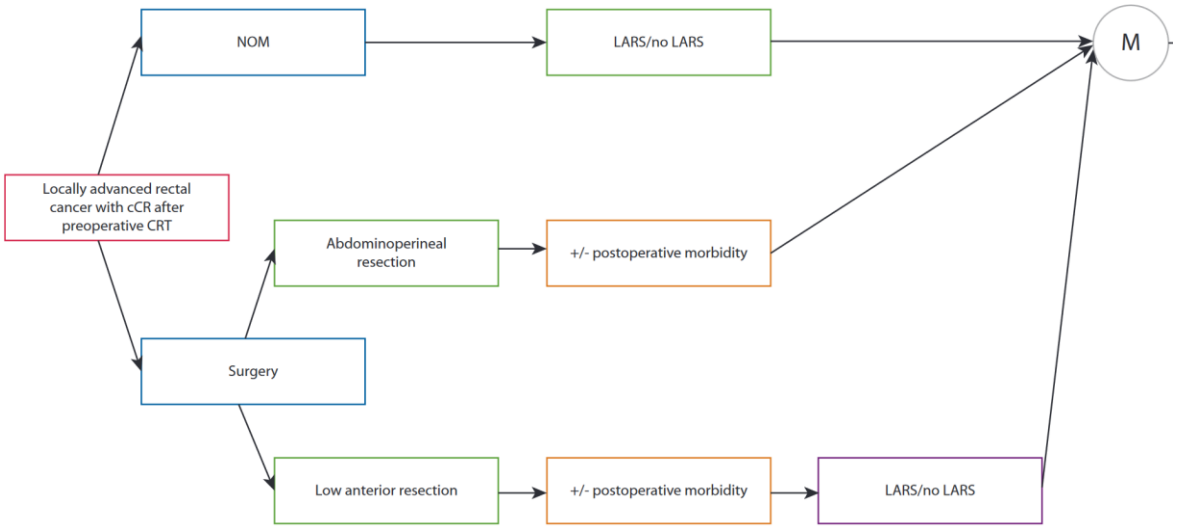


Figure 7: Illustration of the Markov model as applied by Rodriguez-Pascal et al. (2022)

de Buck van Overstraeten et al (2020) published a model comparing non-operative management with radical surgery after nCRT. The authors found about equal QALYs in both groups (better quality of life but slightly shorter life expectancy in the non-operative management group). The authors combined a decision tree with a Markov model as depicted in Figure 8. A Canadian healthcare perspective was used and a 10 years' time horizon.

No consideration to personalized approaches was given, but interestingly, the preferred treatment strategy changed with variations in the probability of local regrowth in nonoperative management, pointing to the possible benefit in better predicting patients at risk for local regrowth. Also, the authors give explicit attention to a well-known complication of surgical resection, the low anterior resection syndrome (LARS).

A. Decision tree



B. Markov states

Figure 8: Illustration of the model as applied by de Buck van Overstraeten et al. (2020).

NOM = Non-operative Management; LARS = low anterior resection syndrome

Based on the above we propose a combined decision tree- Markov model with the following features, as shown in Table 4.

Table 4: proposed features of the model to assess the C-Eff of AI in the management of LARC

Design item	Description	Comment
Target population	*Patients with LARC stages II and III*	*Consider the overall target population*
Age of the target population	*Average 65 years but with age distribution applied*	*Typical age in the published literature*
Model type	*Decision tree plus Markov*	*To better distinguish the short and long term outcomes*
Time horizon	*10 years*	*Given the older target population, 10 years can be justified. Lifetime requires too many assumptions*
Perspective	*Healthcare and society*	*Benefits are not restricted to the healthcare system; also impact on productivity envisaged..*
Comparators	*Standard MRI to guide decision making compared with AI enhanced MRI*	*AI is expected to serve as biomarker*
Risk stratification	*Yes*	*Will be more and more applied. Model needs to be future-proof*
Accounting for false positives and false negatives	*Yes*	*Is at the core of the anticipated benefits of AI*
Markov cycle	*6 months*	*Offers sufficient granularity*
Impact of resection	*Explicitly included*	*Impact on QoL to be accounted for*
Treatments of local and distant recurrence	*Not modelled explicitly*	*Focus is on the course of the disease, i.e. locoregional recurrence and metastasis*

3.3 Key clinical and epidemiological data inputs related to optimizing the management of patients with LARC

In order to populate the combined decision tree-Markov model to assess the potential cost-effectiveness of AI in optimizing the management of LARC patients, several epidemiological and clinical data are required. The following table is inspired by several papers, such as Chin et al. (2022), Mueller et al. (2023) and Rodriguez-Pascual et al. (2022).

Table 5: key epidemiological data required for the combined decision tree - Markov model regarding personalizing LARC management

Parameter	Possible sources	Comments
% of patients undergoing nCRT	*Local epidemiological data*	*Will depend on imaging/AI based risk assessment*
% of patients undergoing operative resection	*Local epidemiological data – Published literature*	*Will depend on success of TNT if applied*
Mortality after nCRT plus surgery	*Published literature*	
Types of resection (abdominoperineal vs low anterior)	*Local epidemiological data – Published literature*	*Can be country-dependent*
Probability of anastomotic leak after resection	*Local epidemiological data – Published literature*	"
Risk of LARS - *After non operative management* - *After LAR* - *After LAR + anastomotic leak*	*Local epidemiological data – Published literature*	"
Transition to locoregional relapse - *After NOM* - *After resection*	*Local epidemiological data – Published literature*	*Depends on initial risk profile*
Management options for locoregional recurrence - *chemo + radical resection + IORT* - *palliative chemo*		
Transition to metastatic disease	*Local epidemiological data – Published literature*	"
Transition from locoregional relapse to metastasis	*Local epidemiological data – Published literature*	*Depends on whether locoregional recurrence is resectable*
Mortality - *From chemoradiotherapy* - *From resection* - *From locoregional recurrence* - *From metastasis* *Natural mortality*		

3.4 Cost and utility data inputs related to optimizing the management of patients with LARC

Table 6 provides an overview of key cost and utility data required for the decision tree-Markov model combination regarding the management of patients with LARC.

Table 6: key cost and utility data required for the combined decision tree - Markov model regarding personalizing LARC management

Parameter	Possible sources	Comment
Cost of risk stratification	*Local official data*	*Extra cost of AI to be applied*
Cost of neoadjuvant chemotherapy	*Local official data – Published literature*	
Cost of adjuvant chemotherapy	*Local official data – Published literature*	
Cost of resection (abdominoperineal and low anterior)	*Local official data – Published literature*	
Annual cost of follow-up	*Published Literature*	
Cost of resection for recurrence	*Local official data – Published literature*	*Including rehabilitation and Home health care*
Cost of intraoperative radiation	*Local official data*	
Duration and cost of productivity loss	*Local official data – Published literature*	*Focus on resection and mortality*
Cost of dying	*Local official data – Published literature*	*Relates to last weeks/months of life*
Utility associated with LAR	*Published Literature*	
Utility associated with APR	*Published Literature*	
Utility associated with NOM	*Published Literature*	
Disutility associated with a stoma	*Published Literature*	*= loss in utility*
Disutility associated with chemotherapy	*Published Literature*	*= loss in utility*
Disutility associated with LARS	*Published Literature*	*= loss in utility*
Utility of postoperative management	*Published Literature*	*Depends on response status and morbidity*
Utility associated with local recurrence	*Published Literature*	
Utility associated with distant recurrence	*Published Literature*	

4 Improving the accuracy of mammograms in the screening of breast cancer

4.1 The context and the potential role of AI

Every year, there are more than 2.26 million new breast cancer cases worldwide (Srinath et al, 2023). As of 2020, breast cancer has become the most diagnosed cancer globally, overtaking lung and prostate cancers. The countries with the highest breast cancer incidence are Belgium, Luxemburg, The Netherlands and France (Arzanova and Mayrovitz, 2022).

Substantial progress has been made in the treatment of breast cancer, resulting in consistently declining breast cancer mortality rates and an improvement in quality of life (El Masri and Phadke, 2022).

Early detection and diagnosis through screening is still needed in order to increase the chances of treatment and therefore of survival (Srinath et al, 2023). In many jurisdictions biannual screening is applied between the age of 50 and 70. Currently there is debate about the effectiveness and cost-effectiveness of screening among the 40-49 years age group.

Digital mammography and digital breast tomosynthesis (DBT) are the cornerstones of breast imaging, especially for breast cancer screening. Randomized clinical trials, systematic reviews, and observational studies have demonstrated that screening mammography reduces breast cancer–related mortality by 20%–50% (Yoon et al. 2023). A study by Molassiotis et al. (2021) based on the global burden of disease study estimated a lower effect on mortality but a much higher effect on disability adjusted life years (DALYs).

On the negative side, screening is still associated with false negative results and therefore not detecting all cancers, and with false positive results having an impact on the quality of life of patients and requiring additional investigations (Flemban, 2023).

Computer-aided detection of cancer has reduced the number of missed cases but at the cost of an increase in false-positive interpretations. Estimates derived from trials suggest that 11–22% of the breast cancer cases detected by screening might be overdiagnosed (Mühlberger et al. 2021).

There is also an ongoing trend in increased use of risk-stratified screening approaches whereby women at increased risk are subject to more intensified screening and those at low risk to less intensive or no screening (Mühlberger et al. 2021).

A recent meta-analysis of studies on the stand-alone performances of AI for interpretation of digital mammography and DBT shows that current algorithms perform on par with, if not better than, the average performance of breast radiologists (Yoon et al. 2023). AI can also identify discriminative image patterns from full-field mammograms to categorize a woman's risk of developing breast cancer in the future (Mital and Nguyen, 2022)

It is also expected that AI will contribute to decreased false positive and false negative results, hence improving the screening and management of breast cancer.

4.2 A model to assess the cost-effectiveness of AI in the screening for breast cancer

In an extensive review, Mühlberger et al (2021) found that existing evaluations suggest risk-adapted screening for breast cancer should be more effective and efficient than conventional screening. The authors recommend that future evaluations of breast cancer screening should more strongly focus on risk-adapted strategies.

Several studies have indeed been published regarding the cost-effectiveness of breast cancer screening. In general, these models have higher complexity than the models in the two other case studies.

Mital and Nguyen (2022) concentrated on the use of AI in deciding the optimal risk-based screening approach in the 40-49 age group. The authors simulated 100,000 white women aged 40 years with no previous history of breast cancer. Each woman had an underlying risk of developing breast cancer based on a risk distribution estimated for USA white females using a comprehensive set of genetic and other non-modifiable and modifiable breast cancer risk factors. Women were classified into three categories: (i) 'true' low risk, defined as those with an underlying risk of breast cancer less than 1.1 times the average risk in the population of 40 year old women (that is, relative risk (RR) is lower than 1.1); (ii) 'true' high risk, defined as those with RR between 1.1 and 4; and (iii) 'true' very high risk, defined as those with RR of 4 or higher. With these RR thresholds, 1% of the hypothetical study cohort was classified as 'true' very high risk, 42% as 'true' high risk and the remaining 57% as 'true' low risk.

The authors developed a hybrid decision tree/microsimulation model to estimate the costs and effectiveness of eight possible screening strategies. The analysis was conducted from the health care system's perspective, the cycle length was 1 year and a lifetime horizon was used. The authors accounted for false positive and negative results and simulated, based on the screening result, the further prognosis of patients going from in situ cancer to local, regional and distant cancer.

Tollens et al (2021) focused on women with dense breast and compared the cost-effectiveness of Digital breast tomosynthesis (DBT) and abbreviated breast MRI (AB-MRI). Decision analysis and Markov simulations were used to model the cumulative costs and quality-adjusted life-years (QALYs) over a time horizon of 30 years. First, a decision analytic model including DBT and AB-MRI was constructed, and the respective outcomes true-positive, false-positive, true negative, and false-negative were defined for each diagnostic strategy. Next, a Markov model was constructed with the following states: absence of cancer, undetected breast cancer, detected malignancy, post-treatment states, and death. The cycle length was set to one year and the screening interval was set to 2 years. A true-positive finding resulted in a timely treatment, whereas a false-negative finding consecutively resulted in delayed treatment and a higher probability of progressive disease with more extensive and costly therapy. In case of false-positive findings, unnecessary follow-ups with associated costs and impairment of quality of life (QoL) were assumed. Positive findings of DBT resulted in biopsy, whereas positive findings of AB-MRI resulted in a full protocol breast MRI examination followed by a biopsy only in case of a confirmed finding.

Both models can be considered to do the job of analysing the cost-effectiveness of AI in the risk stratification of women at risk for breast cancer as well as in improving the accuracy of MRI in women with dense breasts. Based on these examples, the following table describes the key characteristics of the proposed breast cancer model.

Table 7: proposed features of the Markov model to assess the C-Eff of AI in breast cancer screening

Design item	Description	Comment
Target population	*Women eligible for breast cancer screening*	*Including women with dense breasts*
Age of the target population	*40 to 75 years*	*Explore widening the age range*
Time horizon	*30 years*	*Five/ten years is too short. Lifetime requires too many assumptions*
Perspective	*Healthcare and society*	*Benefits to the healthcare system and an impact on productivity are envisaged given the average relatively young age.*
Comparators	*Current state of the art screening without AI vs with AI*	
Risk stratification	*Yes*	*Will be more and more applied. Model needs to be future-proof*
Accounting for false positives and false negatives	*Yes*	*Also to be applied for risk stratification; Is at the core of the anticipated benefits of AI*
Accounting for tumour size	*Yes*	*Better identification of small tumours anticipated with AI*
Markov cycle	*6 months*	*Offers sufficient granularity*
Course of the disease	*Yes*	*Diagnosis at earlier stage means better prognosis*
Treatments of breast cancer	*Modelled implicitly*	*Affects costs and outcomes per cancer stage*

Figure 9 shows how the structure of this model can be built up (adapted from Fobelets et al. 2015)

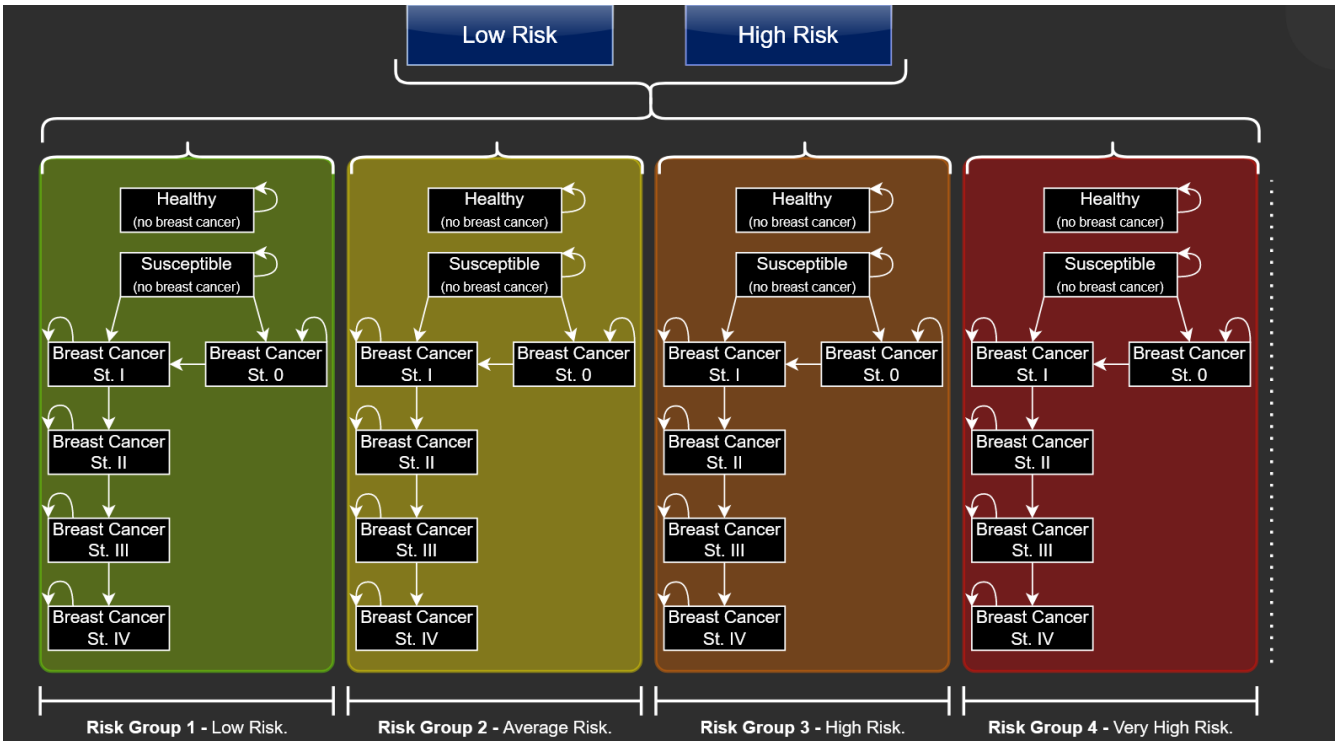


Figure 9: Illustration of the proposed structure of the model to assess the cost-effectiveness of AI in breast cancer screening.

Women can be considered as Low Risk or High Risk for developing breast cancer, which can further be finetuned into Low, Average, High and very High risk. Without screening, in women who are affected by cancer, the cancer will naturally progress until it is diagnosed based on symptoms. With screening, the cancer will be detected earlier and hence on average in an earlier stage. Implicitly in the model is the size of the tumour, the nodes affected etc… to determine the risk of further progression.

4.3 Key clinical and epidemiological data related to screening for breast cancer

In order to populate the Markov model to assess the potential cost-effectiveness of AI in the screening for breast cancer several epidemiological and clinical data are required. Based on several publications and the proposed features of our Markov model, we established an overview of the key input data. These are summarized in table 8.

Table 8: key clinical and epidemiological data required for the Markov model regarding screening for breast cancer

Parameter	*Possible sources*	*Comments*
Starting age of the model	*Local epidemiological data*	*Preference for setting a range from 40 to 75 years*
Incidence of breast cancer	*Local epidemiological data*	*Age dependent*
Sensitivity and specificity for risk assessment	*Published literature*	*Better accuracy of AI anticipated*
Sensitivity and specificity of mammography	*Published literature*	*Better accuracy of AI anticipated*
Stage distributions upon diagnosis - *No screening* - *Annual screening* - *Biannual screening*	*Local epidemiological data – cohort data*	
Types of breast cancer depending on Estrogen Receptor (ER) and HER2 status upon diagnosis	*Published literature*	*Additional finetuning may be needed*
Markov transitions - *From local to regional* - *From local to distant* - *From regional to distant*	*Published literature*	*Will require additional calculations*
Probability of resection	*Local epidemiological data*	*Depending on tumour size - Can be country-dependent*
Success of resection	*Published literature*	*Depending on tumour size*
Breast cancer mortality	*Local epidemiological data and published literature*	*Specific to age and stage at diagnosis as well as (ER) and human epidermal growth factor 2 (HER2) status*
Natural mortality rates	*Local epidemiological data*	

4.4 Cost and utility data inputs related to screening for breast cancer

Table 9 provides an overview of key cost and utility data required for the decision tree-Markov model combination regarding the screening for breast cancer.

Table 9: key cost and utility data required for the Markov model regarding screening for breast cancer

Parameter	***Possible sources***	***Comment***
Cost of genetic testing and counselling for risk stratification	*Published literature*	*Extra cost of AI to be applied*
Cost of mammography	*Local official data – Published literature*	
Additional diagnostic costs in case of true and false positives	*Local official data – Published literature*	*Including biopsy*
Cost of treatment for tumour <1cm	*Local official data – Published literature*	*Including follow-up costs*
Cost of treatment for tumour >1cm	*Local official data – Published literature*	*Including follow-up costs*
Cost of treating locoregional recurrence	*Local official data – Published literature*	*Including follow-up costs*
Cost of distant metastasis One site	*Local official data – Published literature*	*Accounting for most recent pharmacotherapy*
Cost of distant metastasis Multiple sites	*Local official data – Published literature*	*Accounting for most recent pharmacotherapy*
Cost of end-of life management	*Local official data – Published literature*	
Disutility of false positive results	*Published Literature*	
Utility of detected tumor < 1 cm	*Published Literature*	
Utility of detected tumor > 1 cm	*Published Literature*	
Utility of regional recurrence	*Published Literature*	
Utility of metastasis one site	Published Literature	
Utility of metastasis multiple sites	Published Literature	

5 Challenges and next steps

5.1 Data challenges

For all described models, the majority of the clinical data and utilities (the latter needed to calculate QALYs) can be found in the available published literature. Epidemiological data however can be quite country specific. For instance, countries such as Sweden and Italy generally have good data availability with this regard, as have some regions in Spain. For Poland and Lithuania more challenges are anticipated.

More specifically for the first use case, surveillance for HCC in patients with cirrhosis, observational cohort data are needed to estimate the annual probability of HCC in patients with cirrhosis. To illustrate this, a literature search looking for the combination "(Poland OR Lithuania) AND cirrhosis AND QALY" did not result in any hits.

Also, the management options for the different stages of HCC will be country dependent and therefore more challenging to obtain in some countries. Nevertheless, the Global Burden of Disease Liver Cancer Collaboration (2017) included all the above countries. Also, Herman et al (2015) studied the variation in treatment modalities, costs and outcomes of rectal cancer patients in Poland.

For the second use case, clinical data and utilities are again available in the literature. But the management of LARC, more specifically the current use of neoadjuvant treatment, and the types of surgery applied will be country dependent and will pose a challenge to collect in some countries. The same counts for the management and the cost of locoregional relapse and metastatic disease.

Also, some of the transitions in the model can be country dependent, for instance the transition to locoregional and distant disease or the risk of LARS.

Finally, for the breast cancer model, we expect less challenges, since this disease has been subject to many more publications than the other two use cases. For instance, in Lithuania, Ivanauskienė et al. (2015) studied the cost of newly diagnosed breast cancer.

5.2 Next steps

In this report, we described the protocol for assessing the cost-effectiveness of AI-based tools in three different 'use cases'. This involves the anticipated benefits of AI, the design of the model to assess the cost-effectiveness and the required clinical, epidemiological, cost and utility data to complete the analyses.

It is proposed to work out the 3 models in Microsoft Excel for the sake of transparency. Indeed, for a model to be validated, not only the structure of the model needs validation, but also the content (the link between the different variables and values in the model). This is much facilitated if all the elements in the model are transparent and verifiable.

It is proposed to work out the draft model for each use case for one country (to be selected) which will allow to have a first assessment of the potential cost-effectiveness of AI in the 3 use cases.

For instance, a typical output of the breast cancer model is shown in Figure 10, illustrating the predicted stage of breast cancer at diagnosis, with or without screening.

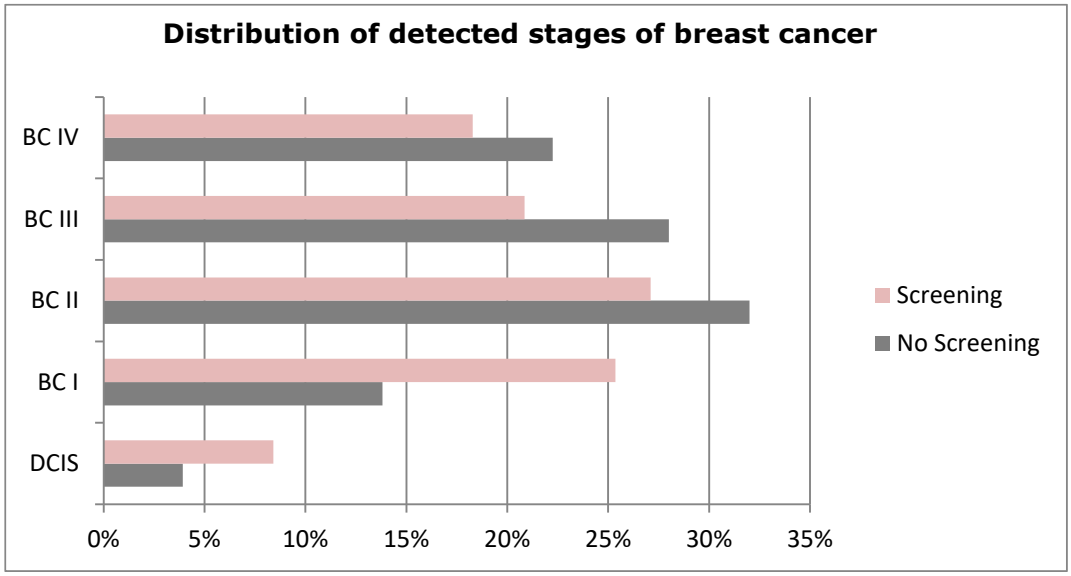


Figure 9: Illustration of the intermediate outcomes the model to assess the cost-effectiveness of AI in breast cancer screening

Also, this step will allow to better and more concretely identify the data gaps, to discuss assumptions made and to assess, via sensitivity analysis (such as Tornado Diagrams) which variables affect the results most, hence assessing the drivers of the cost-effectiveness.

This first version will then also be subject to clinical validation of the models in the 3 use cases, with advisory boards of specialists in the respective use cases.

Based on the above the models will be revised and made ready for adaptation to other countries. This means that for all country dependent variables placeholders will be foreseen such that local researchers can assist in the completion of the models. An instruction guide will be prepared to assist for that purpose.

The aim is to publish the final results in peer reviewed journals and contribute to building the knowledge on the value and value for money of AI in healthcare.

LITERATURE REFERENCES

Adeniji N, Dhanasekaran R. Current and Emerging Tools for Hepatocellular Carcinoma Surveillance. Hepatol Commun. 2021 Dec;5(12):1972-1986.

Arzanova E, Mayrovitz HN. The Epidemiology of Breast Cancer. In: Mayrovitz HN, editor. Breast Cancer [Internet]. Brisbane (AU): Exon Publications; 2022 Aug 6. Chapter 1.

Aschele C, Glynne-Jones R. Selecting a TNT Schedule in Locally Advanced Rectal Cancer: Can We Predict Who Actually Benefits? Cancers (Basel). 2023 Apr 30;15(9):2567.

Barnard Giustini A, Ioannou GN, Sirlin C, Loomba R. Review article: Available modalities for screening and imaging diagnosis of hepatocellular carcinoma-Current gaps and challenges. Aliment Pharmacol Ther. 2023 May;57(10):1056-1065.

Bartolotta TV, Randazzo A, Bruno E, Taibbi A. Focal liver lesions in cirrhosis: Role of contrast-enhanced ultrasonography. World J Radiol. 2022 Apr 28;14(4):70-81.

Borelli B, Germani MM, Carullo M, Mattioni R, Manfredi B, Sainato A, Rossi P, Vagli P, Balestri R, Buccianti P, Morelli L, Antoniotti C, Cremolini C, Masi G, Moretto R. Total neoadjuvant treatment and organ preservation strategies in the management of localized rectal cancer: A narrative review and evidence-based algorithm. Crit Rev Oncol Hematol. 2023 Jun;186:103985.

Bourbonne V, Schick U, Pradier O, Visvikis D, Metges JP, Badic B. Radiomics Approaches for the Prediction of Pathological Complete Response after Neoadjuvant Treatment in Locally Advanced Rectal Cancer: Ready for Prime Time? Cancers (Basel). 2023 Jan 9;15(2):432.

Cadier B, Bulsei J, Nahon P, Seror O, Laurent A, Rosa I, Layese R, Costentin C, Cagnot C, Durand-Zaleski I, Chevreul K; ANRS CO12 CirVir and CHANGH groups. Early detection and curative treatment of hepatocellular carcinoma: A cost-effectiveness analysis in France and in the United States. Hepatology. 2017 Apr;65(4):1237-1248

Carter HE, Jeffrey GP, Ramm GA, Gordon LG. Cost-Effectiveness of a Serum Biomarker Test for Risk-Stratified Liver Ultrasound Screening for Hepatocellular Carcinoma. Value Health. 2021 Oct;24(10):1454-1462.

Chin RI, Otegbeye EE, Kang KH, Chang SH, McHenry S, Roy A, Chapman WC Jr, Henke LE, Badiyan SN, Pedersen K, Tan BR, Glasgow SC, Mutch MG, Samson PP, Kim H. Cost-effectiveness of Total Neoadjuvant Therapy With Short-Course Radiotherapy for Resectable Locally Advanced Rectal Cancer. JAMA Netw Open. 2022 Feb 1;5(2)

Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. Future Healthc J. 2019 Jun;6(2):94-98.

de Buck van Overstraeten A, Khorasani S, Kennedy E, Look Hong NJ. Nonoperative Management Versus Radical Surgery of Rectal Cancer After Neoadjuvant Therapy-Induced Clinical Complete Response: A Markov Decision Analysis. Dis Colon Rectum. 2020 Aug;63(8):1080-1089

De Mattia E, Polesel J, Mezzalira S, Palazzari E, Pollesel S, Toffoli G, Cecchin E. Predictive and Prognostic Value of Oncogene Mutations and Microsatellite Instability in Locally-Advanced Rectal Cancer Treated with Neoadjuvant Radiation-Based Therapy: A Systematic Review and Meta-Analysis. Cancers (Basel). 2023 Feb 25;15(5):1469

El Masri J, Phadke S. Breast Cancer Epidemiology and Contemporary Breast Cancer Care: A Review of the Literature and Clinical Applications. Clin Obstet Gynecol. 2022 Sep 1;65(3):461-481.

European Association for the Study of the Liver. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. J Hepatol. 2018 Jul;69(1):182-236.

Farhang Zangneh H, Wong WWL, Sander B, Bell CM, Mumtaz K, Kowgier M, van der Meer AJ, Cleary SP, Janssen HLA, Chan KKW, Feld JJ. Cost Effectiveness of Hepatocellular Carcinoma Surveillance After a Sustained Virologic Response to Therapy in Patients With Hepatitis C Virus Infection and Advanced Fibrosis. Clin Gastroenterol Hepatol. 2019 Aug;17(9):1840-1849.e16

Flemban AF. Overdiagnosis Due to Screening Mammography for Breast Cancer among Women Aged 40 Years and Over: A Systematic Review and Meta-Analysis. J Pers Med. 2023 Mar 14;13(3):523.

Fleming C, Vendrely V, Rullier E, Denost Q. Organ preservation in rectal cancer: review of contemporary management. Br J Surg. 2022 Jul 15;109(8):695-703.

Fobelets M. et al. [Dutch] https://borstkanker.bevolkingsonderzoek.be/sites/default/files/2022-03/Kosteneffectiviteitsstudie%20Bevolkingsonderzoek%20BK%205%20oktober%202015.pdf

Global Burden of Disease Liver Cancer Collaboration. The Burden of Primary Liver Cancer and Underlying Etiologies From 1990 to 2015 at the Global, Regional, and National Level: Results From the Global Burden of Disease Study 2015. JAMA Oncol. 2017 Dec 1;3(12):1683-1691.

Goossens N, Singal AG, King LY, Andersson KL, Fuchs BC, Besa C, Taouli B, Chung RT, Hoshida Y. Cost-Effectiveness of Risk Score-Stratified Hepatocellular Carcinoma Screening in Patients with Cirrhosis. Clin Transl Gastroenterol. 2017 Jun 22;8(6):e101.

Herman KJ, Komorowski AL, Wysocki WM, Tabor J, Herman RM, Śliwczyński A. Variation in treatment modalities, costs and outcomes of rectal cancer patients in Poland. Contemp Oncol (Pozn). 2015;19(5):400-9

Inchingolo R, Maino C, Cannella R, Vernuccio F, Cortese F, Dezio M, Pisani AR, Giandola T, Gatti M, Giannini V, Ippolito D, Faletti R. Radiomics in colorectal cancer patients. World J Gastroenterol. 2023 May 21;29(19):2888-2904.

Ivanauskienė R, Domeikienė A, Kregždytė R, Milašauskienė Ž, Padaiga Ž. The cost of newly diagnosed breast cancer in Lithuania, 2011. Medicina (Kaunas). 2015;51(1):63-8.

Kulkarni PA, Singh H. Artificial Intelligence in Clinical Diagnosis: Opportunities, Challenges, and Hype. JAMA. 2023;330(4):317–318.

Li L, Hu Y, Han J, Li Q, Peng C, Zhou J. Clinical Application of Liver Imaging Reporting and Data System for Characterizing Liver Neoplasms: A Meta-Analysis. Diagnostics (Basel). 2021 Feb 17;11(2):323.

Mitsala A, Tsalikidis C, Pitiakoudis M, Simopoulos C, Tsaroucha AK. Artificial Intelligence in Colorectal Cancer Screening, Diagnosis and Treatment. A New Era. Curr Oncol. 2021 Apr 23;28(3):1581-1607.

Molassiotis A, Tyrovolas S, Giné-Vázquez I, Yeo W, Aapro M, Herrstedt J. Organized breast cancer screening not only reduces mortality from breast cancer but also significantly decreases disability-adjusted life years: analysis of the Global Burden of Disease Study and screening programme availability in 130 countries. ESMO Open. 2021 Jun;6(3):100111

Mueller AN, Torgersen Z, Shashidharan M, Ternent CA. Cost-Effectiveness Analysis: Selective Use of Neoadjuvant Chemoradiation in Locally Advanced Rectal Cancer. Dis Colon Rectum. 2023 Jul 1;66(7):946-956.

Nahon P, Najean M, Layese R, Zarca K, Segar LB, Cagnot C, Ganne-Carrié N, N'Kontchou G, Pol S, Chaffaut C, Carrat F, Ronot M, Audureau E, Durand-Zaleski I. Early hepatocellular carcinoma detection using magnetic resonance imaging is cost-effective in high-risk patients with cirrhosis. JHEP Rep. 2021 Nov 4;4(1):100390.

Ouyang G, Chen Z, Dou M, Luo X, Wen H, Deng X, Meng W, Yu Y, Wu B, Jiang D, Wang Z, Yao Y, Wang X. Predicting Rectal Cancer Response to Total Neoadjuvant Treatment Using an Artificial Intelligence

Model Based on Magnetic Resonance Imaging and Clinical Data. Technol Cancer Res Treat. 2023 Jan-Dec;22

Renzulli M, Braccischi L, D'Errico A, Pecorelli A, Brandi N, Golfieri R, Albertini E, Vasuri F. State-of-the-art review on the correlations between pathological and magnetic resonance features of cirrhotic nodules. Histol Histopathol. 2022 Dec;37(12):1151-1165.

Renzulli M, Brandi N, Argalia G, Brocchi S, Farolfi A, Fanti S, Golfieri R. Morphological, dynamic and functional characteristics of liver pseudolesions and benign lesions. Radiol Med. 2022 Feb;127(2):129-144. (b)

Rodriguez-Pascual J, Nuñez-Alfonsel J, Ielpo B, Lopez M, Quijano Y, de Vicente E, Cubillo A, Saborido CM. Watch-and-Wait policy versus robotic surgery for locally advanced rectal cancer: A cost-effectiveness study (RECCOSTE). Surg Oncol. 2022 May;41:101710.

Slipsager A, Henrichsen SN, Falkmer UG, Dybkær K, Belting M, Poulsen LØ. Predictive biomarkers in radioresistant rectal cancer: A systematic review. Crit Rev Oncol Hematol. 2023 Jun;186:103991.

Srinath A, van Merode F, Rao SV, Pavlova M. Barriers to cervical cancer and breast cancer screening uptake in low- and middle-income countries: a systematic review. Health Policy Plan. 2023 Apr 11;38(4):509-527

Taylor EJ, Jones RL, Guthrie JA, Rowe IA. Modeling the benefits and harms of surveillance for hepatocellular carcinoma: Information to support informed choices. Hepatology. 2017 Nov;66(5):1546-1555.

Villanueva A. Hepatocellular carcinoma. N Engl J Med. 2019;380(15):1450–62

Vogl TJ, Mader C, Michalik S, Hammerstingl R, Koch V. Das kleine hepatozelluläre Karzinom : Leitliniengerechte und klinisch etablierte Diagnostik [Small hepatocellular carcinoma : Diagnostics according to guidelines and established in the clinical setting]. Radiologe. 2022 Mar;62(3):239-246.

Yoon JH, Strand F, Baltzer PAT, Conant EF, Gilbert FJ, Lehman CD, Morris EA, Mullen LA, Nishikawa RM, Sharma N, Vejborg I, Moy L, Mann RM. Standalone AI for Breast Cancer Detection at Screening Digital Mammography and Digital Breast Tomosynthesis: A Systematic Review and Meta-Analysis. Radiology. 2023 Jun;307(5):e222639.