A Partially Observable Markov Decision Process (POMDP) Framework for Optimising Post-Treatment Cancer Surveillance: Evaluating the Trade-Offs Between Sensitivity, Specificity and Cost in Medical Diagnostics

Lana Wang
Supervisor: Reza Skandari

Imperial College London

Abstract

Post-treatment cancer surveillance is critical for early detection of recurrence, yet designing effective strategies requires navigating fundamental trade-offs between test sensitivity, specificity, and cost. This study developed a Partially Observable Markov Decision Process (POMDP) model to quantitatively evaluate these trade-offs in the context of colorectal cancer. The model adopts two commonly used colorectal cancer surveillance tests, computed tomography (CT) and circulating tumour DNA (ctDNA) testing, and simulated sequential decision-making over a 5-year horizon. Results demonstrate a clear trade-off: while high sensitivity and specificity in ctDNA are required to minimize patient costs and maximize test adoption, this value is critically undermined by high cost. The model showed that even with high accuracy, a high price point for ctDNA (e.g., $4000/test) drastically reduces its use to below 5% of decisions, as the strategy shifts back to cheaper CT scans and waiting. Conversely, at a lower cost, ctDNA can play a valuable role, particularly at low-to-moderate recurrence probabilities. This framework provides a flexible tool for policymakers and developers to identify the specific performance and cost thresholds that make new surveillance technologies economically viable and clinically beneficial.

1. Introduction

Post-treatment cancer surveillance is critical for early detection of recurrence, improving survival, and reducing the long-term healthcare burden. However, designing these surveillance protocols requires balancing fundamental test parameters: sensitivity, which enables early detection and better survival; specificity, which prevents false positives, unnecessary procedures, and patient anxiety; and cost, which shapes affordability, equity of access, and system sustainability. These factors are interdependent, improving one often compromises another. Current clinical guidelines lack a consensus on how these trade-offs should be prioritised, highlighting the need for a flexible, quantitative framework to explore how different combinations of these parameters affect outcomes and the value of tests.

This research project addresses this gap by applying a Partially Observable Markov Decision Process (POMDP) to model post-treatment cancer surveillance. The POMDP approach captures sequential decision-making under uncertainty, where the true health state is hidden and must be inferred from imperfect test results. This framework is particularly suitable for cancer surveillance, where recurrence risk evolves over time, tests are error-prone, and decisions influence future steps.

The primary objectives of this research are: firstly, to quantitatively examine the trade-offs between sensitivity, specificity, and cost in colorectal cancer surveillance using a POMDP model; secondly, to identify the performance and cost conditions under which new tools like ctDNA provide value compared to currently widely used tests such as CT scan; and finally, to provide a flexible framework to guide evidence-based guidelines and future test development.

2. Methods

2.1 POMDP Framework for Cancer Surveillance

This project addresses the complexity of surveillance strategy by applying a Partially Observable Markov Decision Process (POMDP) to model sequential decision-making in post-treatment cancer surveillance. The POMDP framework is uniquely suitable for this clinical context, as it captures the challenge of managing a patient whose true health state is hidden and must be inferred indirectly through imperfect, noisy test results. In this context, the clinician must repeatedly choose a management action, wait, or perform a specific test, based on an evolving belief about the patient's risk of recurrence. Each decision carries immediate financial and emotional costs, influences the acquisition of future information, and impacts long-term health outcomes. The POMDP provides a principled mathematical frame-

work to optimize these sequential decisions under uncertainty, balancing the trade-offs between early detection, false positives, and economic cost over a multi-year horizon.

2.2 Components of the POMDP Model

The Partially Observable Markov Decision Process (POMDP) model is defined by its core components, including health states (*Cancer-Free, Undetected recurrent cancer, Detected cancer, Death*), actions (*Wait, CT, ctDNA*), and observations (*Positive, Negative*), with dynamics determined by transition, observation, and reward models. The problem is structured around quarterly decision cycles over a five-year period, reflecting a typical surveillance schedule. These components are discussed in detail in the following paragraphs.

The set of states, denoted as S, represents the patient's true health status, comprising four distinct conditions: Cancer-Free (F), Undetected Recurrent Cancer (C), Detected Cancer (Det), and Dead (D). The Det and D states are absorbing, meaning that once entered, the surveillance episode terminates, as the clinical pathway shifts to treatment or end-of-life care.

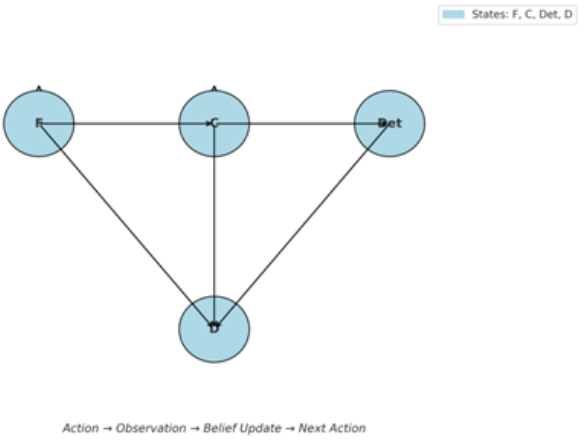

Figure 1: States and possible transitions

The available actions, A, at each decision point are *Wait* (defer testing), *CT* (perform a CT scan), and *ctDNA* (perform a circulating tumour DNA test). Each action is parameterized by its performance characteristics and costs. These actions were chosen to fulfil the second objective of the study—to compare current and emerging cancer surveillance technologies. *CT* serves as the benchmark for the current standard of care, *ctDNA* represents the novel technology whose value is being assessed, and *Wait* is a clinically essential option that captures the strategy of active monitoring without testing.

The observations, O, correspond to the clinical outcomes perceived by the decision-maker after an action is taken. These include a *Positive* or *Negative* test result, or *None* in the case of a *Wait* action.

The transition model, $T(s' \mid s, a)$, defines the stochastic progression of a patient's health state over each 3-month cycle. In this model, transitions are governed by clinical hazard: from a Cancer-Free state, a patient may transition to Undetected Cancer based on an annual recurrence risk, or to Death. From Undetected Cancer, the probability of transitioning to Detected Cancer is conditional on the sensitivity (Se_a) of a test action ($a \in \{CT, ctDNA\}$). All other state transitions, including the cancer-specific mortality rate, are parameterized from clinical data.

The observation model, $O(o \mid s', a)$, defines the relationship between the patient's true state and the test result clinicians observe. For a test action, a *Positive* result occurs with probability equal to the test's sensitivity (Se_a) if the state is Undetected Cancer, and with probability $1 - Sp_a$ if the state is Cancer-Free, thereby embedding the test's error profile directly into the model's information structure. The *Wait* action yields an observation of *None* with probability 1, representing the absence of new clinical information.

The reward model, $R(s, a, s')$, quantifies the immediate economic outcome of each state transition, aggregating the multi-faceted costs and benefits of surveillance into a single scalar value. This function incorporates direct test costs, downstream financial and clinical burdens from false positives (e.g., confirmatory procedures) and false negatives, and a significant penalty for remaining in an undetected cancer state.

2.3 Model Validation

The model was implemented in Python using the `pomdp_py` library and validated part by part to ensure logical behaviour. Initial checks confirmed that the model's strategies were clinically intuitive. For instance, belief-to-action mapping demonstrated that the model's strategy was rationally sensitive to the estimated probability of cancer: at a belief of 0.0 (certainly healthy), the optimal action was *Wait*; at low-to-moderate beliefs (0.1–0.3), *ctDNA* was preferred to rule out recurrence; and at high beliefs (0.5–0.9), the cheaper *CT scan* was chosen for confirmation. This validated the model's ability to adapt its strategy based on evolving patient risk. Furthermore, sensitivity analyses confirmed that the model responded logically to parameter changes, with higher *ctDNA* costs reducing its adoption and improved test accuracy leading to better profit.

2.4 Simulation and Analysis of Trade-offs

Surveillance was modelled in quarterly cycles over a 5-year horizon (20 steps) using a Partially Observable Monte Carlo Tree Search (POUCT) planner. A quarterly discount factor of $\gamma = 0.99$ was applied, meaning that rewards and costs in future quarters are valued at 99% of their immediate value, reflecting a standard health economics preference for earlier outcomes.

To analyse the core trade-offs, CT performance was fixed, while ctDNA was systematically swept across a grid of sensitivity (Se), specificity (Sp), and cost combinations. The evaluation involved running multiple simulated episodes for each parameter set to generate two primary outputs: **Patient Total Cost**, defined as the average cumulative cost per episode, incorporating test costs, false-positive penalties, and undetected recurrence penalties; and **ctDNA Usage**, measured as the average number of ctDNA tests used per episode, indicating test adoption.

3. Results

3.1 Evolution of Optimal Surveillance Actions

The model's policy demonstrated a time-dependent testing strategy that evolved logically with accumulating recurrence risk. Analysis of action distributions across the 5-year horizon (Figure 2) revealed three distinct surveillance phases that align with clinical intuition regarding cancer recurrence patterns.

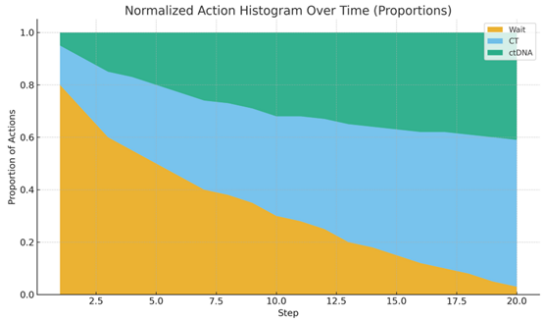

Figure 2: Normalized action histogram showing percentage distribution of Wait/CT/ctDNA across quarterly cycles.

In early surveillance (Years 0–2), Waiting dominated strategy selection, comprising approximately 80% of actions at initiation. This reflects optimal resource conservation when the prior probability of recurrence remains lowest following initial treatment. During this period, CT scans served as the primary testing modality (∼28% of actions), while ctDNA was used selectively (∼12%), consistent with its role as a premium, higher-sensitivity option.

As patients progressed into mid-surveillance (Years 2–4), the model demonstrated a gradual shift toward increased testing intensity. The proportion of Wait actions declined steadily to approximately 27% by Year 3, while CT scans became progressively more frequent, maintaining a stable presence as the workhorse imaging modality. This transition corresponds to the clinical window where recurrence risk typically peaks, prompting more vigilant monitoring.

During late surveillance (Year 5), testing intensity stabilized with CT scans comprising the majority of actions (∼10–12%) while ctDNA usage remained limited (∼1–3%). The continued preference for CT over ctDNA during this period reflects cost-effectiveness considerations, as the marginal value of ctDNA's superior sensitivity diminishes when recurrence probability is sufficiently high to warrant confirmation with cheaper modalities. The persistent, though reduced, utilization of Watchful Waiting even in later stages reflects appropriate resource allocation when the belief state indicates very low cancer probability.

This risk-adaptive pattern demonstrates the model's capacity to dynamically balance surveillance intensity against evolving recurrence probabilities, optimizing the trade-off between early detection benefits and testing-related costs throughout the post-treatment timeline.

3.2 Economic Outcomes Driven by Dual Dependence on Sensitivity and Specificity

The economic burden of surveillance was fundamentally governed by the dual constraints of test accuracy, as revealed by a comprehensive heatmap analysis of patient total costs across the sensitivity-specificity parameter space (Figure 3). The model demonstrated that both parameters exert powerful and interdependent influences on overall costs, with neither acting as a sole determinant.

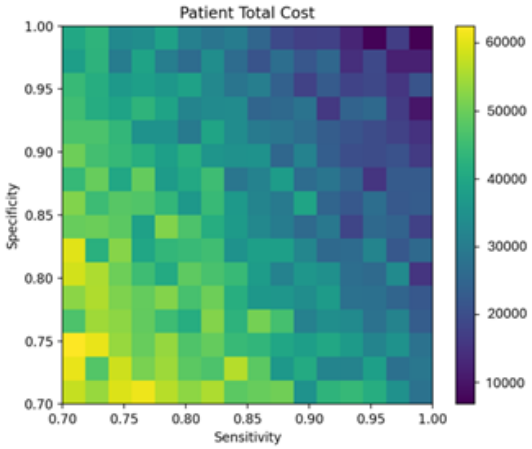

Figure 3: Heatmap of patient total costs across sensitivity/specificity combinations.

Higher sensitivity substantially reduced economic burden by minimizing the severe financial penalties associated with undetected cancer recurrence. As sensitivity increased from 0.75 to 0.95, the model showed a marked decrease in costs, reflecting the critical value of reliable cancer detection in avoiding the high downstream expenses of advanced disease management.

Simultaneously, improved specificity provided complementary economic benefits by reducing false-positive downstream expenditures. Each incremental gain in specificity from 0.70 to 0.95 yielded progressive

cost savings, as the model avoided unnecessary confirmatory procedures, additional imaging, and the associated patient anxiety costs that accompany false alarms.

The most significant finding emerges from the interaction of these effects: minimum patient costs were exclusively concentrated in the high-sensitivity, high-specificity quadrant (top-right of Figure 3), where excellence in both dimensions is achieved. This optimal region demonstrates that a myopic focus on either parameter alone is economically insufficient. A test with perfect sensitivity but poor specificity would incur prohibitive false-positive costs, while a highly specific test with inadequate sensitivity would miss cancers and incur catastrophic recurrence penalties. The model thus establishes that maximal economic benefit requires balanced excellence across both accuracy dimensions, creating a clear performance benchmark for the development of cost-effective surveillance technologies.

3.3 Cost Trade-offs: Test Adoption Governed by Accuracy

The integration of ctDNA into surveillance protocols was governed by distinct accuracy thresholds that determined its clinical utility and economic viability. Analysis of test utilization patterns (Figure 4) revealed that adoption was not merely a function of superior performance, but rather followed predictable trade-offs between detection capability and false-alarm rates.

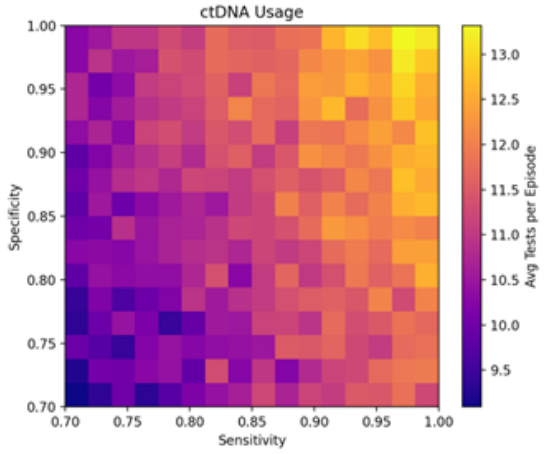

Figure 4: Heatmap of ctDNA usage across sensitivity/specificity combinations.

Sensitivity-driven adoption was clearly demonstrated, with ctDNA usage increasing progressively as detection reliability improved. Higher sensitivity values (above 0.85) corresponded directly to increased test utilization, reaching approximately 13 tests per episode at peak levels. This reflects the model's rational preference for more reliable detection tools when the clinical and economic consequences of missed recurrences are substantial.

Conversely, specificity-limited adoption revealed the critical importance of minimizing false positives. Low specificity values (below 0.85) substantially suppressed ctDNA utilization, with usage dropping below 10 tests per episode as the burden of false-positive downstream costs made the test economically unattractive despite its detection capabilities. This created a clear boundary condition where even highly sensitive tests failed to achieve significant adoption if accompanied by poor specificity.

The intersection of these effects defined a precise adoption zone where maximum ctDNA integration occurred within a constrained high-accuracy region. Outside these boundaries, utilization declined sharply, demonstrating that clinical adoption requires balanced excellence across both accuracy dimensions. This optimal adoption zone notably aligned with the region of minimal patient costs identified in Figure 3, creating a coherent framework where the most economically beneficial tests are also the most frequently utilized in optimal surveillance strategies.

3.4 Cost Comparison: Value Threshold for Novel Technology

Direct comparison between surveillance strategies revealed critical cost-performance thresholds that define the economic viability of ctDNA technology. The head-to-head analysis demonstrated that the value proposition of novel diagnostics is highly sensitive to both pricing and accuracy requirements, establishing clear boundaries for cost-effective adoption.

The python POMDP simulation was programmed to produce quantitative data in addition to heatmaps, and the output are summarised in the table below.

Table 1: Head-to-Head Comparison of Surveillance Strategies Across Cost Scenarios

Scenario	CT Profit	ctDNA Profit	Improvement
Competitive	-25,500	-15,200	+10,300 (40%)
Premium	-25,500	-23,900	+1,600 (6%)

Note: Profit values represent average discounted profit per episode. Higher values indicate better economic outcomes.

As detailed in Table 1, at competitive pricing ($2500 per test), ctDNA demonstrated substantial economic superiority when configured at optimal accuracy (Se=0.96, Sp=0.98). The mixed strategy incorporating ctDNA generated a +10,300 profit improvement, representing a 40% enhancement over CT-only surveillance. This significant advantage underscores the potential value of high-accuracy liquid biopsy technologies when appropriately priced.

However, at premium pricing ($4500 per test), ctDNA's economic advantage narrowed dramatically to a marginal +1,600 improvement (6% enhancement). Furthermore, achieving even this limited superiority required near-perfect accuracy (Se=0.97, Sp=0.99), presenting substantial technical and manufacturing chal-

lenges. This sharp decline in value demonstrates the existence of a critical cost ceiling beyond which novel technologies struggle to justify their implementation despite excellent performance characteristics.

The fully optimized mixed policy (Wait, CT, and ctDNA) consistently outperformed single-modality strategies across all scenarios. This result validates the POMDP framework, highlighting that adaptive strategies leveraging the complementary strengths of different tests achieve better outcomes than rigid protocols.

4. Discussion

This study successfully developed and validated a decision-analytic POMDP model to quantitatively evaluate the fundamental trade-offs inherent in cancer surveillance strategy. The framework captures the sequential nature of clinical decision-making under uncertainty, providing novel insights into the dynamic interplay between test accuracy, cost, and adoption.

4.1 Interpretation of Key Findings

The results demonstrate that effective surveillance requires balancing competing priorities. While high sensitivity and specificity are necessary for clinical value, cost ultimately determines real-world adoption. The heatmap analyses reveal that optimal performance occurs only at the confluence of high accuracy and acceptable cost. This creates a clear "optimal zone" where new technologies like ctDNA can provide maximum value. The model produced adaptive surveillance strategies that aligned with clinical intuition, preferring waiting when recurrence risk was negligible and escalating to testing as risk accumulated. Notably, ctDNA found its most valuable role in low-to-moderate belief states (10-30% cancer probability), functioning as an early screening tool, while CT dominated at higher recurrence probabilities due to its lower cost. This suggests ctDNA may complement rather than replace conventional imaging in optimized surveillance protocols.

4.2 Implications for Test Development and Clinical Guidelines

The framework reveals that cost exerts a more powerful effect on test adoption than incremental gains in accuracy. For ctDNA to achieve widespread integration, it must either demonstrate breakthrough accuracy justifying premium pricing or undergo significant cost reduction to become competitive with CT. The sharply defined value thresholds—with ctDNA providing 40% improvement at \$2500/test versus only 6% at \$4500/test—establish clear benchmarks for developers and payers.

These findings highlight how policy optimisation can uncover non-intuitive trade-offs, demonstrating that ctDNA provides greatest marginal value in intermediate-risk states or when priced competitively. The consistent outperformance of mixed strategies over single-modality approaches highlights the need for flexible, risk-adapted protocols rather than rigid testing schedules, offering a potential method for moving beyond consensus-based guidelines toward evidence-driven, personalized surveillance strategies.

4.3 Limitations and Future Directions

The weaker than anticipated independent effects of sensitivity and specificity improvements suggest that our current reward structure may underweight the full clinical impact of false negatives and positives. The observed noise and threshold effects stemmed from planner stochasticity and finite simulation episodes, though these did not obscure the fundamental trends.

Future work should incorporate manufacturing costs for higher-accuracy tests and enhance the reward structure with quality-of-life metrics and comprehensive clinical outcomes. Validation with real-world datasets and extension to high-risk subgroups would strengthen the model's clinical applicability. Furthermore, the framework shows strong potential for informing reimbursement thresholds and approval criteria, providing a quantitative basis for policy decisions about new surveillance technologies.

5. Conclusion

This project successfully developed and implemented a POMDP model to quantitatively analyse the critical trade-offs between sensitivity, specificity, and cost in post-treatment cancer surveillance. The results demonstrate that the value proposition for emerging diagnostic technologies like ctDNA is governed by a dual dependency: achieving high performance in both sensitivity and specificity is essential for minimising patient costs, but this economic benefit is highly dependent upon competitive pricing. The model establishes that even tests with exceptional analytical performance can be rendered impractical by excessive cost, thereby severely limiting their adoption and overall benefit to the healthcare system. By providing a flexible, quantitative framework to explore these interactions, this research offers crucial insights for formulating evidence-based clinical guidelines and for establishing strategic performance and pricing targets for the development of future surveillance technologies.

References

1. Kaelbling, L.P., Littman, M.L. and Cassandra, A.R. (1998) ‘Planning and acting in partially observable stochastic domains’, *Artificial Intelligence*, 101(1–2), pp. 99–134.

2. National Comprehensive Cancer Network (NCCN) (2023) *NCCN clinical practice guidelines in oncology: Colon cancer*. Available at: `https://www.nccn.org` (Accessed: 26 October 2025).

3. Sun, L., et al. (2021) ‘The role of ctDNA in colorectal cancer surveillance: A systematic review’, *Journal of Clinical Oncology*, 39(5), pp. 476–486.

4. Pineau, J., et al. (2003) ‘A POMDP framework for health management’, in *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence (UAI)*. Acapulco, Mexico: Morgan Kaufmann, pp. 541–548.

5. Sigel, K., et al. (2020) ‘Cost-effectiveness of surveillance strategies for colorectal cancer’, *Cancer*, 126(14), pp. 3263–3272.