IMPERIAL

LAIDLAW SCHOLARS

Optimising sensitivity and specificity in medical diagnostics

Student: Lana Wang
Supervisor: Dr. Reza Skandari

Introduction

The problem:
Post-treatment cancer surveillance is critical for early detection of recurrence, improving survival, and reducing healthcare burden. However, available tests vary in **sensitivity, specificity, and cost**, and current guidelines lack clarity on how to prioritise these trade-offs. A flexible, quantitative framework is needed to explore how different parameter combinations affect outcomes and values of tests.

Why this matters:
- **Sensitivity** enables early detection and better survival.
- **Specificity** prevents false positives, unnecessary procedures, and anxiety.
- **Cost** shapes affordability, equity, and system sustainability.
- These factors are interdependent—improving one often compromises another.

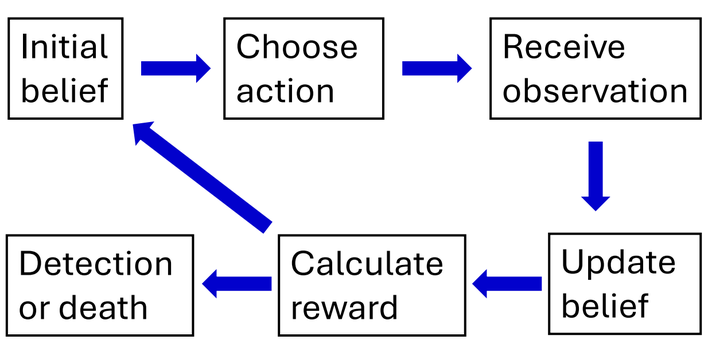
Objectives:
A Partially Observable Markov Decision Process (**POMDP**) model was applied to colorectal cancer surveillance to examine how sensitivity, specificity, and cost interact. The goal was to identify the **trends**, and to provide a flexible framework to guide **evidence-based guidelines** and **future test development**.

Method

Framework:
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 approach is suitable for cancer surveillance, where recurrence risk evolves over time, tests are error-prone, and decisions influence future steps. The model was implemented in Python using the *pomdp_py* library.

Model structure:
The model included four health states (cancer-free, recurrence, detected, dead), and three actions (Wait, CT, and ctDNA). These were chosen to represent the current standard (CT), the emerging test (ctDNA), and the realistic option of deferring testing.

Decision Process:
Surveillance was modelled in quarterly cycles over 5 years (20 steps). The model used a Monte Carlo tree search to choose the action. Each cycle:



Results

Dynamic surveillance decisions
Figure 1 shows, in early cycles, the preferred action was often to wait, reflecting low recurrence risk. As time progressed, CT scans became increasingly frequent and eventually dominated, while ctDNA tests remained a secondary choice.

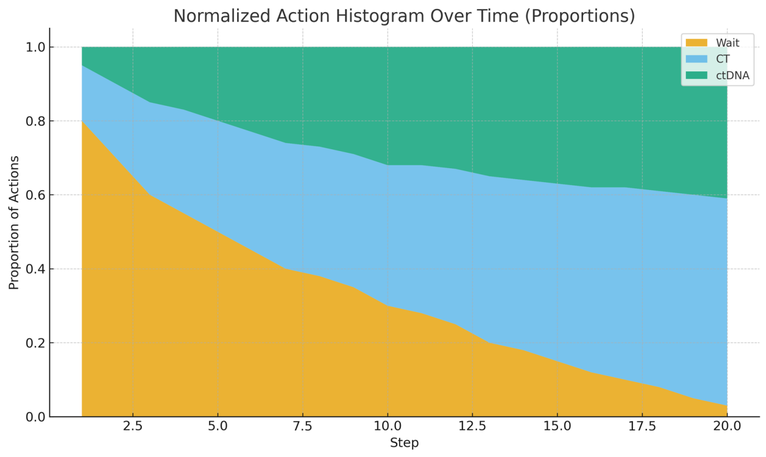


Figure 1: *surveillance choices over time with normalised percentages*

Effect of sensitivity and specificity on cost
Figure 2 shows that higher sensitivity lowers costs by reducing missed cancers, while higher specificity lowers costs by avoiding false positives. The lowest costs occur when both are high. Minor checkerboard patterns reflect Monte Carlo noise when strategies are nearly equivalent.

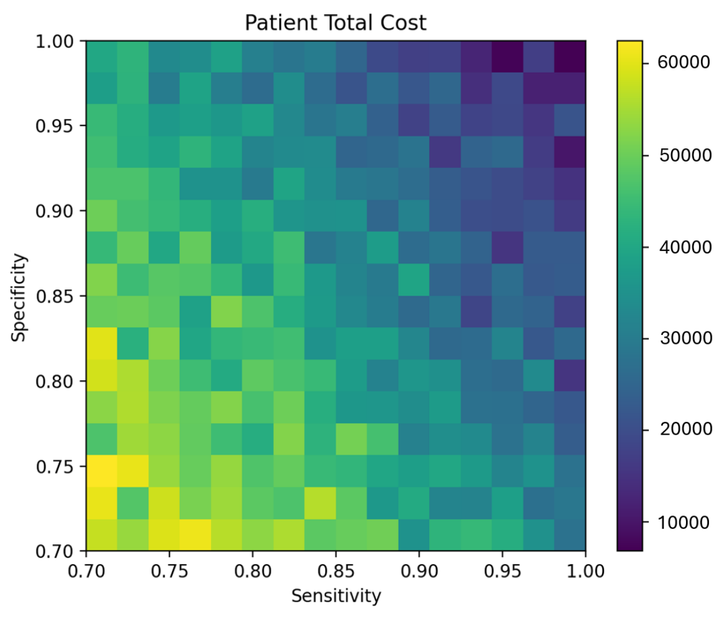


Figure 2: *Heatmap of patient total costs across sensitivity/specificity*

Effect of Sensitivity and Specificity on ctDNA
Figure 3 shows that ctDNA use rises with higher sensitivity and falls with lower specificity. The highest adoption occurs at high sensitivity and specificity, while small irregular patches reflect planning noise where CT and ctDNA are nearly tied.

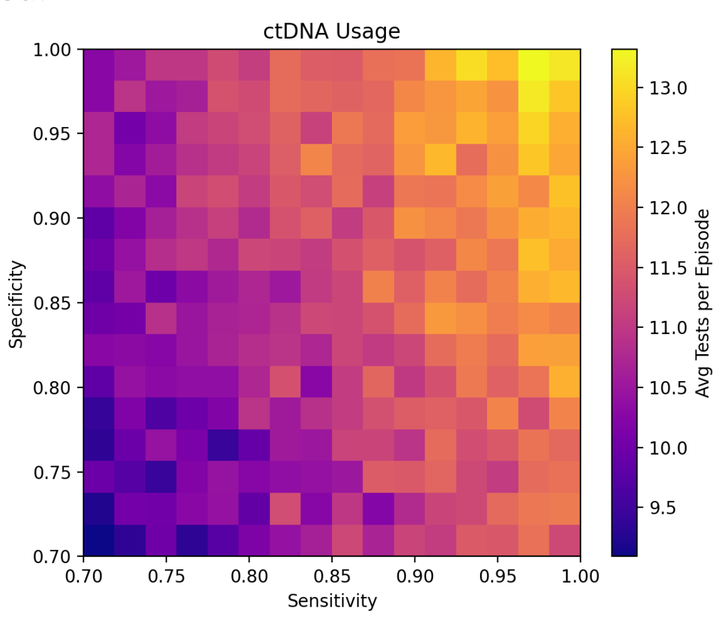


Figure 3: *Heatmap of ctDNA usage (average tests per episode) across sensitivity/specificity*

Discussion

- **Dynamic policy:** The POMDP framework produced adaptive surveillance strategies that aligned with intuition: wait when recurrence risk is negligible, test more as risk accumulates, and use cheaper CT once recurrence becomes likely.
- **Role of ctDNA:** ctDNA was most valuable in the low-to-moderate belief range, functioning as an early screening tool. At higher recurrence probabilities, CT dominated due to its lower cost, suggesting ctDNA may complement rather than replace CT.
- **Parameter sensitivity:** Cost behaved as expected — higher ctDNA cost reduced its use. Sensitivity and specificity effects were weaker than anticipated, reflecting that the current reward structure (low false-positive penalty, limited true-positive benefit) underweights accuracy improvements. Adjusting the undetected penalty and false-positive costs could sharpen these relationships.
- **Noise and thresholds:** Small deviations in monotonicity stemmed from planner stochasticity, finite episodes, and threshold switching between equally good policies. Increasing simulation runs or adjusting exploration constants could smooth results.
- **Implications:** These findings highlight that policy optimisation can uncover non-intuitive trade-offs. ctDNA is unlikely to dominate across all scenarios but may provide the greatest marginal value in intermediate-risk states or when priced competitively. The framework shows potential for informing guidelines by mapping where new surveillance technology adoption improves both economic and clinical outcomes.

Conclusion

This project developed a decision-analytic POMDP model to evaluate cancer surveillance strategies. Results show that ctDNA has the greatest potential when sensitivity and specificity are both high, lowering patient costs and encouraging adoption. The model highlights the trade-off between sensitivity, specificity, cost, and test usage, offering insights that can inform clinical guidelines and technology assessment.

Next steps

- **Model refinement**: incorporate manufacturing and implementation costs of higher-accuracy tests.
- **Validation**: compare model outputs with clinical datasets or trial results to confirm validity.
- **Extensions**: test the framework on high-risk subgroups, or include non-monetary costs (e.g., anxiety, quality of life).
- **Policy application**: use outputs to guide reimbursement thresholds and approval criteria for new surveillance technologies.