



ISE432 Decision Analysis

Project Report

A Decision-Analytic Re-Examination of “A Decision Support Model for Colorectal Cancer Screening” (Corrales, Ríos Insua & González, 2025)

Hikmet Alp Babataş - [REDACTED]

Tan Ergin - [REDACTED]

Selin Koçyiğit - [REDACTED]

Şeyda Nur Budan [REDACTED]

Instructor: Assoc. Prof. Pelin Gülşah Canbolat

**Yeditepe University, Faculty of Engineering - Industrial Engineering
Department**

TABLE OF CONTENTS

(a) Description of the Decision Problem

1. Problem Classification and Context
2. Decision Maker and Context
3. Decision Variables, State Space and Outcomes
4. Notation and Parameters
5. The Influence Diagram
6. Objectives and Value Attributes
7. Uncertainties
8. Decision Criterion and Solution Approach

(b) Summary of the Paper's Main Results

1. Model Construction
2. Preference Elicitation
3. Personalized Policy Recommendations
4. Referral-Decision Sensitivity
5. Overall Qualitative Conclusion

(c) Discussion of the Model and Results

1. Strengths
2. Modeling Assumptions and Limitations
3. Discussion of the Reported Results

(d) Systematic Numerical Experiments:

- A Turkey vs. France Cost-Structure Comparison

1. Motivation and Research Question
2. Re-implementation of the Model
3. The Turkey Cost Scenario
4. Results
5. Discussion of the Experiment
6. Limitations of the Experiment

References

(a) Description of the Decision Problem

1. Problem Classification and Context

Colorectal cancer (CRC) is among the most common and most preventable cancers because the great majority of cases develop slowly from detectable precursor lesions (adenomatous polyps). Population screening programs can substantially reduce CRC incidence and mortality by catching these lesions (or early-stage cancers) before they progress. However, screening programs face a basic tension: the most diagnostically informative procedure “colonoscopy” is also the most expensive and the least comfortable, while cheap and convenient stool-based tests (gFOBT, FIT) are far less sensitive. National guidelines therefore typically recommend a single “one-size-fits-all” test (commonly FIT every one to two years for the average-risk population), even though individuals differ enormously in their personal risk of CRC, their tolerance of invasive procedures and their sensitivity to out of pocket cost.

Corrales, Ríos Insua and González (2025) address this tension by formulating CRC screening as a **personalized sequential decision problem under uncertainty**: for a given individual, which screening test (if any) should be recommended first, and depending on its result, should the individual be referred for a diagnostic colonoscopy? The paper's central contribution is a **Bayesian decision-analytic model**, implemented as an **Influence Diagram (ID)**, a Bayesian network augmented with decision nodes (which test to take; whether to refer to colonoscopy) and a value node (the individual's utility), that can be personalized to a specific patient's risk profile and preferences and solved by maximizing expected utility (EU). This places the paper squarely within the classical Decision Analysis toolkit covered in our lecture: a sequential decision under uncertainty, multi-attribute value/utility elicitation, Bayesian updating of beliefs and expected-utility maximization as the normative decision rule.

2. Decision Maker and Context

The **decision maker (DM)** is, jointly, a screening program / clinician and the patient: the model is intended as a *decision-support* tool that a clinician would use, in consultation with the patient, to arrive at a personalized screening recommendation. The patient is characterized by a vector of observable **risk factors** : age, sex, body-mass index (BMI), smoking and alcohol habits, physical activity level, sleep, family history of CRC, and comorbidities such as hypertension and diabetes. These risk factors are the *inputs* to a Bayesian network (built from epidemiological evidence) that outputs the patient's prior probability of having CRC, $p(CRC)$.

Crucially, the DM in this model is not a national health authority choosing a single population-wide policy (as is implicitly the case for guideline-based “everyone gets FIT” recommendations), but an individual decision unit re-solving the *same* decision model for each patient with that patient's own $p(CRC)$ and own elicited preferences (comfort weights λ_k and CARA risk-aversion parameters a, b, ρ). This personalization is what distinguishes the paper from a simple cost-effectiveness comparison of screening tests at the population level.

3. Decision Variables, State Space and Outcomes

Decision variables. The DM must choose, sequentially:

1. A first-stage screening test $r_s \in$

{No screening, gFOBT, FIT, Blood-based test, sDNA (stool DNA), CT colonography (CTC), Colon capsule, eight options in total. Each test is characterized by a sensitivity Se , specificity Sp , monetary cost and a *comfort level* $comf \in \{1,2,3,4\}$ (Table 2 of the paper groups the eight options into four comfort classes, from 4 = “no test” to 1 = “colonoscopy” with CTC/CC at level 2 and the remaining stool/blood tests at level 3).

2. **A second-stage referral decision** $r_c \in \{\text{refer to colonoscopy, do not refer}\}$, made *after* observing the result of the first-stage test (positive or negative). This is the key *sequential* feature of the model: r_c is conditioned on the information revealed by r_s .

State space: The underlying (unobserved) state of the world is the patient’s true CRC status, $CRC \in \{0,1\}$. This state is never observed directly, it is only *inferred* probabilistically, first from the risk-factor-based prior $p(CRC)$, and then updated via Bayes’ rule after each test result.

Outcomes: An outcome of the decision problem is a triple $(cost, info, comf)$: the (expected) monetary cost incurred, the (expected) information gained about the patient’s true CRC status, and the comfort level of the procedure(s) undergone. These three attributes are mapped to a single scalar utility $u \in [0,1]$ via the value/utility model described in Section 6 below, and the DM’s objective is to choose (r_s, r_c) to maximize the **expected utility** $EU(r_s, r_c) = E[u(cost, info, comf)]$, where the expectation is taken over the (Bayesian-updated) distribution of test results and, where applicable, colonoscopy outcomes.

4. Notation and Parameters

The table below summarizes the notation used throughout this report, following the paper’s conventions.

Symbol	Meaning
$CRC \in \{0, 1\}$	True (unobserved) colorectal-cancer status of the patient
$p(CRC)$	Prior probability of $CRC = 1$, from the Bayesian network over risk factors
r_s	First-stage screening-test decision (8 options, incl. “No screening”)
r_c	Second-stage referral decision (refer to colonoscopy / do not refer)

Se, Sp	Sensitivity / specificity of a given test (paper's Table 21)
$cost$	Expected monetary cost of the chosen test(s) and any follow-up colonoscopy
$comf \in \{1, 2, 3, 4\}$	Comfort level of the procedure(s); 1 = least comfortable (colonoscopy), 4 = no test
$H(CRC)$	Shannon entropy of the prior $p(CRC)$
v_{info}	Normalized information value of the test result(s), = $RPMI/H(CRC)$
λ_k	Comfort-level-specific trade-off weight, $k = comf$ (paper's Table 8)
$v(cost, info, comf)$	Single-attribute value, $v = \lambda_k v_{info} - \log_{10}(cost + 1)$
a, b, ρ	CARA (constant absolute risk aversion) utility parameters (paper's Table 9)
$u(cost, info, comf)$	Utility on $[0, 1]$, $u = a - b e^{-\rho v}$
$EU(r_s, r_c)$	Expected utility of a screening / referral policy

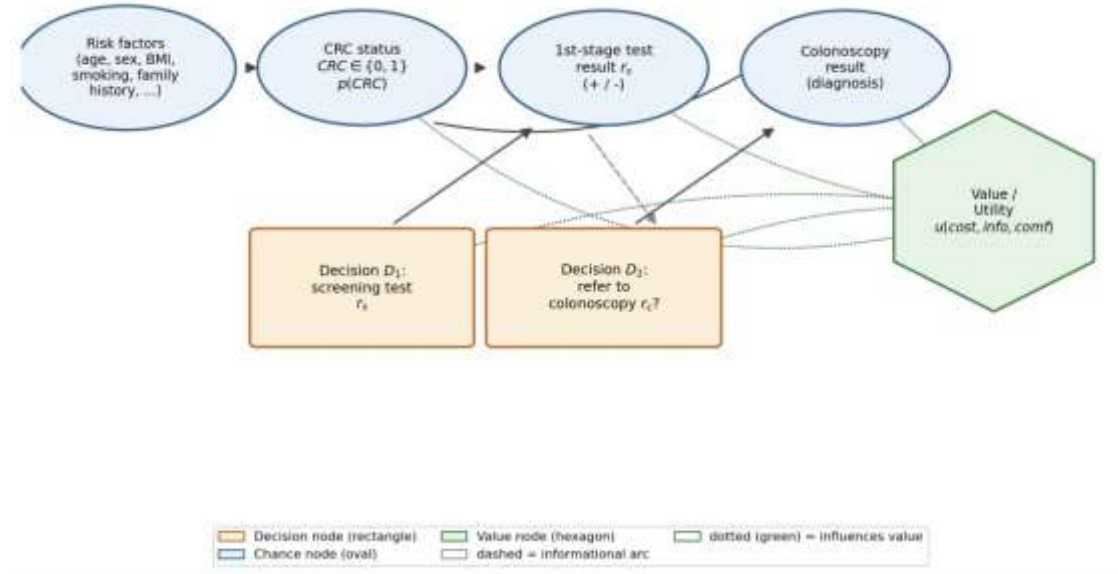
The numerical values used in our re-implementation are: $\lambda_1 = 4.01$, $\lambda_2 = 4.17$, $\lambda_3 = 6.80$, $\lambda_4 = 7.00$ and $a = 1.015$, $b = 0.872$, $\rho = 0.039$ (the “benchmark patient” elicitation, paper's Table 9). Test sensitivities, specificities, comfort levels, and (French) costs are reproduced from the paper's Table 21 in Section (d), Table d.1.

5. The Influence Diagram

Figure 1 reproduces, in simplified form, the structure of the Influence Diagram (ID) that underlies the paper's model. Rectangles are **decision nodes**, ovals are **chance (uncertainty) nodes**, and the hexagon is the **value node**. The patient's risk-factor profile is a chance node that determines the prior $p(CRC)$; the screening-test decision $D_1(r_s)$ and the true CRC status jointly determine the distribution of the first-stage test result; the test result is observed *before* the referral decision $D_2(r_c)$ is made (the dashed arc denotes this informational dependence); and D_2 , together with the true CRC status, determines the distribution of the colonoscopy result if a

referral is made. Every decision and every chance node that affects cost, information or comfort feeds into the value node (dotted green arcs).

Figure 1. Simplified Influence Diagram for the personalized CRC screening / referral decision (Corrales, Ríos Insua & González, 2025)



This ID representation makes explicit the two features that distinguish the paper’s model from a simple “which single test is best” comparison: (i) the **sequential** nature of the problem ; r_c is chosen *after* r_s ’s result is known, allowing the referral decision to use the information the first test provides , and (ii) the **personalizability** of the model , the chance node “Risk factors” is patient-specific, so re-solving the ID for a different patient (different $p(CRC)$) or under different cost/comfort parameters re-solves the *entire* downstream decision problem.

6. Objectives and Value Attributes

The DM’s preferences over outcomes are captured by a **multi-attribute utility function** over three attributes, as already introduced in Section 4:

- **Monetary cost** $cost$; the cost of the first-stage screening test, plus, if applicable, the cost of a follow-up colonoscopy and its expected complication costs;

- **Information value** $info$; a normalized measure, based on **Relative Pointwise Mutual Information (RPMI)**, of how much the test result(s) change the DM's belief about the patient's true CRC status relative to the prior;
- **Comfort level** $comf \in \{1,2,3,4\}$ of the procedure(s) undergone.

These three attributes are combined into a single-attribute *value* via

$$v(cost, info, comf = k) = \lambda_k \cdot v_{info} - \log_{10}(cost + 1),$$

where λ_k is the comfort-level-specific trade-off weight (Table 8: $\lambda_1 = 4.01$, $\lambda_2 = 4.17$, $\lambda_3 = 6.80$, $\lambda_4 = 7.00$). The value is then mapped to a utility on $[0,1]$ through a **constant-absolute-risk-aversion (CARA)** transform,

$$u(cost, info, comf) = a - b e^{-\rho v(cost, info, comf)},$$

with $a = 1.015$, $b = 0.872$, $\rho = 0.039$ elicited for the “benchmark” patient (Table 9) via a standard probability-equivalent / certainty-equivalent procedure. The information value itself is computed from the RPMI of the (one or two) test results with respect to the prior $p(CRC)$, normalized by the entropy $H(CRC)$ of the patient's CRC status:

$$v_{info}(crc, r_s, r_c) = \frac{1}{H(CRC)} \left[\log \frac{p(crc | r_s)}{p(crc)} + \log \frac{p(crc | r_s, r_c)}{p(crc | r_s)} \right].$$

Note the additive, *log-cost / linear-info* structure of v : this single-equation value model is what allows the entire ID to be solved by simple expected-utility maximization rather than by, e.g., a multi-objective Pareto analysis ; a modeling choice we return to critically in Section (c).

7. Uncertainties

The chain of uncertain events in the model is:

- The patient's true (unobserved) CRC status, $CRC \in \{0,1\}$, whose **prior probability** $p(CRC = 1)$ is computed from the Bayesian network described above. The paper reports $p(CRC) \approx 0.00085$ – 0.0009 for a “benchmark” average-risk adult, rising to roughly 0.0039

for a patient with hypertension and diabetes, and to 0.1–0.18 under a strong family-history scenario.

- The **result of the first-stage test** $r_s \in \{+, -\}$, generated from *CRC* through the test's sensitivity (true-positive rate) and specificity (true-negative rate), per the paper's Table 21.
- If colonoscopy is performed ; either as the first-stage test or as a referral following a positive first-stage result ; the **result of the colonoscopy**, and the probability and cost of **procedural complications** (bleeding, retention, perforation, other), per the paper's Table 22.

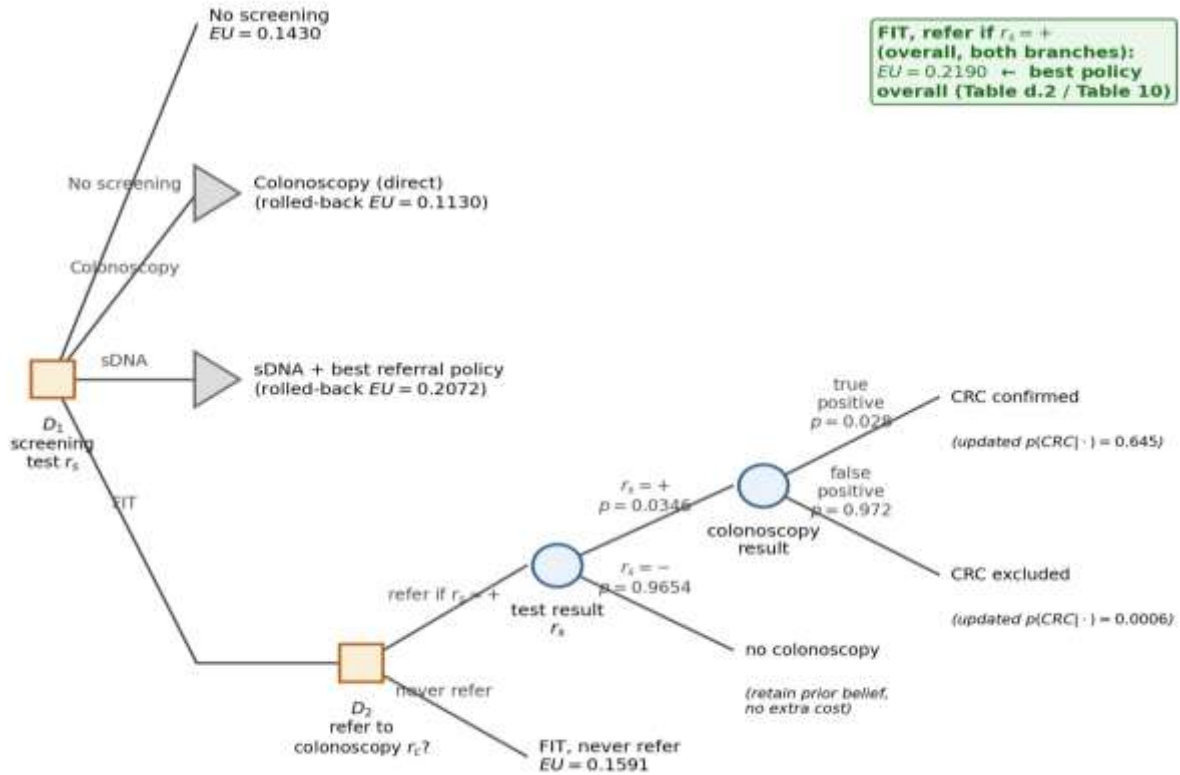
Each of these uncertainties is resolved sequentially and is used to *update* the belief about *CRC* via Bayes' rule, which in turn updates the information-value term v_{info} that feeds the value/utility computation.

8. Decision Criterion and Solution Approach

For a given patient profile, the model evaluates the **expected utility (EU)** of every feasible first-stage-test / referral-policy combination (r_s, r_c) and recommends the policy with the **maximum EU** ; the standard expected utility maximization decision rule of Decision Analysis, applied here to a two-stage (sequential) decision under uncertainty. Solving the ID therefore amounts to a **backward-induction / decision-tree roll-back**: for each first-stage test r_s , the optimal second-stage referral policy $r_c^*(r_s, \text{result of } r_s)$ is determined first (by comparing the EU of “refer” vs. “do not refer” *conditional on* the first-stage result), and the resulting rolled-back EU is then compared across all r_s choices, including “No screening”.

Figure 2 illustrates this roll-back procedure as an explicit, extensive-form decision tree for one branch of the problem (the “FIT” test, evaluated for the benchmark patient, $p(CRC) = 0.00085$, under French costs). The grey triangles represent the other seven first-stage choices, each folded to its own rolled-back EU (computed exactly as for “FIT” but not drawn in full); these numbers correspond to the entries of Table d.2 in Section (d).

Figure 2. Two-stage extensive-form decision tree (illustrative branch: FIT, benchmark patient, French costs)



Two features of Figure 2 are worth highlighting. First, the **referral decision D_2 is only “active” along the $r_s = +$ branch** ; once the first-stage test is negative, the model (in our re-implementation, mirroring the paper’s two-policy structure) does not refer to colonoscopy, so that branch terminates directly. Second, the **rolled-back EU of “FIT, refer if $r_s = +$ ” ($EU = 0.2190$) exceeds both “FIT, never refer” ($EU = 0.1591$) and “No screening” ($EU = 0.1430$)** — illustrating concretely how the *sequential* structure of the ID (test first, then conditionally decide on referral) generates value beyond what either “always test, never refer” or “never test” would achieve alone. Because the model is fully parameterized by patient-specific risk factors (which determine $p(CRC)$) and by test-specific sensitivities, specificities, costs and complication probabilities, it can in principle be re-solved for *any* patient population and *any* cost environment , a feature that is central to the numerical experiment we carry out in Section (d).

(b) Summary of the Paper’s Main Results

The paper develops the model in three stages , model construction, preference elicitation, and personalized policy recommendation , and validates/illustrates it with a series of computational experiments.

1. Model Construction

Section 2 of the paper builds the Bayesian network that estimates $p(CRC \mid \text{risk factors})$ from epidemiological data (age, sex, BMI, smoking, alcohol use, physical activity, sleep, family history, and comorbidities such as hypertension and diabetes), and then extends it into the Influence Diagram of Figure 1 by adding the screening-test decision node, the referral-decision node, and the value node. Sensitivities, specificities and (French, public-system) costs for the eight screening options are taken from the clinical literature and tabulated in the paper’s Table 21; complication probabilities and costs (bleeding, retention, perforation, “other”) for invasive procedures are given in Table 22. These tables are the direct inputs to our own re-implementation in Section (d) (reproduced there as Table d.1).

2. Preference Elicitation

Section 3 of the paper elicits the comfort-level weights λ_k and the CARA utility parameters (a, b, ρ) for a “benchmark” patient ; a 44–54-year-old male, non-smoker, normal BMI, physically active, with no relevant comorbidities and no family history, $p(CRC) \approx 0.00085$ — through a standard probability-equivalent / certainty-equivalent elicitation procedure (Table 23, Appendix C). These elicited values ($\lambda_1 = 4.01$, $\lambda_2 = 4.17$, $\lambda_3 = 6.80$, $\lambda_4 = 7.00$; $a = 1.015$, $b = 0.872$, $\rho = 0.039$) are then held fixed and reused across the personalized scenarios reported in Table 10 ; a modeling choice that we discuss critically in Section (c.2).

3. Personalized Policy Recommendations

Table 10 of the paper reports the model’s top-two recommendations for four patient profiles. Selected results are reproduced below.

Table 1 : Selected results from the paper’s Table 10 (top-2 recommendations and EU)

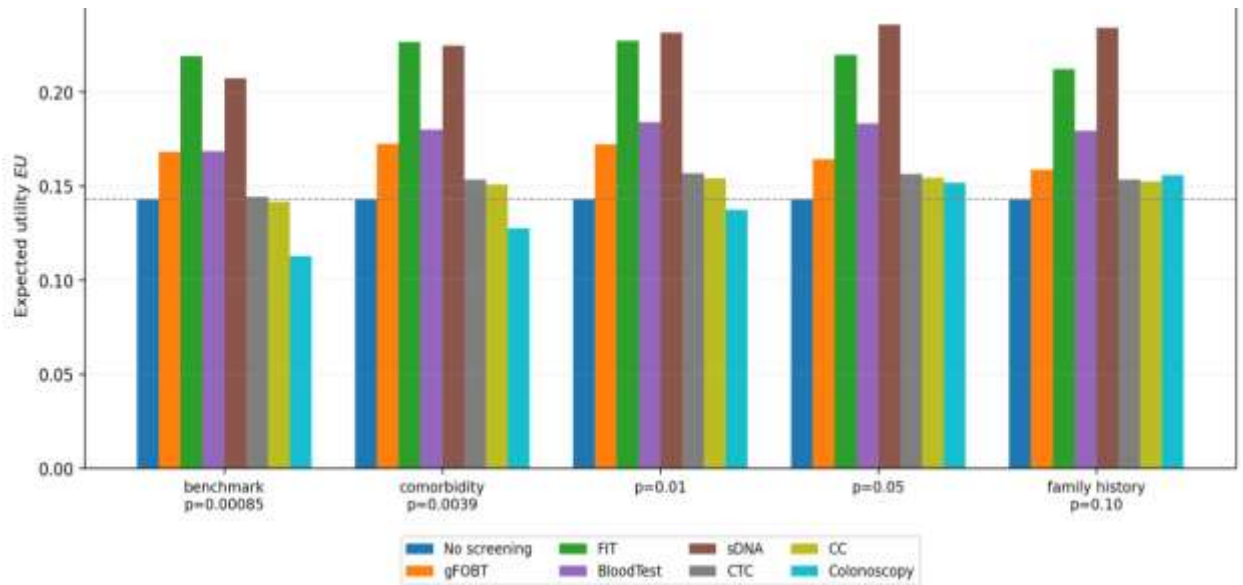
Patient profile	1st recommendation (EU)	2nd recommendation (EU)
Benchmark ($p(CRC) \approx 0.00085$, $\rho = 0.039$)	No screening (0.143)	FIT (0.142)
Risk-seeking ($\rho = 0.005$)	FIT (0.147)	sDNA (0.145)
Comorbidity evidence ($p(CRC) \approx 0.0039$)	sDNA (0.146)	FIT (0.145)
Family history ($p(CRC) \approx 0.10$)	FIT (0.183)	sDNA (0.173)

For the **benchmark patient**, the model’s top-two recommendations are “**No screening**” ($EU = 0.143$) and **FIT** ($EU = 0.142$) ; i.e., for a genuinely average-risk individual, the (small) information gain from FIT is essentially offset by its cost and discomfort, so doing nothing is (marginally) optimal, with FIT a very close second. The paper then perturbs the benchmark patient along three dimensions and shows that the recommendation **changes sensibly**:

- A patient with a **more risk-seeking utility function** ($\rho = 0.005$, i.e., flatter/closer-to-linear utility) is recommended **FIT** first and **sDNA** second , a less risk-averse patient is willing to “pay” more (in cost/comfort) for the same information gain.
- A patient with **added comorbidity evidence** (hypertension and diabetes, raising $p(CRC)$ to about 0.0039) is recommended **sDNA** first and **FIT** second ; the higher baseline risk makes the more sensitive (but costlier) stool-DNA test worth its premium.
- A patient with a strong **family-history** indicator ($p(CRC) \approx 0.1$) is recommended **FIT** first and **sDNA** second, and the EUs of *all* options rise sharply, reflecting the much larger value of information when baseline uncertainty about CRC status is higher.

Figure 3 visualizes this same qualitative pattern using **our own re-implementation** (Section d.2) evaluated across a *continuum* of $p(CRC)$ values, under the French cost vector. While the absolute EU values differ from the paper’s Table 10 (our two-policy referral rule is a simplified — but validated, see Section d.2 — approximation to the full ID solver), the **same qualitative shift** is visible: at low $p(CRC)$, FIT (green) is the top screening option and is closely followed by sDNA (brown); as $p(CRC)$ increases past roughly 0.005–0.01, sDNA overtakes FIT, and the gap between “No screening” (dashed grey line) and every screening option widens. This is the central, qualitative result of the paper, restated in our own framework: **the optimal recommendation is not a single fixed test ; it shifts systematically with the patient’s personal risk level.**

Figure 3. Expected utility by screening option across patient risk profiles (re-implementation, French cost vector).



4. Referral-Decision Sensitivity

The paper also reports how the predicted EU of “no colonoscopy” vs. “colonoscopy” changes under a *positive* versus *negative* first-stage test result, for each patient profile. For example, for the benchmark patient, $EU(\text{test negative, no colonoscopy}) = 0.139$ vs. $EU(\text{test negative, colonoscopy}) = 0.133$, while $EU(\text{test positive, no colonoscopy}) = 0.293$ vs. $EU(\text{test positive, colonoscopy}) = 0.370$. In other words, referral to colonoscopy is **EU-decreasing after a negative first-stage result** but **EU-increasing after a positive result**, exactly the asymmetric, information-dependent referral logic encoded by the decision node D_2 in Figure 1 / Figure 2. This illustrates that the referral decision is itself highly sensitive to the test outcome and to the patient’s prior risk, not just to the choice of first-stage test.

5. Overall Qualitative Conclusion

Taken together, the paper’s main *qualitative* result is that **personalization matters**: a single guideline-based recommendation (e.g., “everyone gets FIT”) is, under this model, only optimal, or even close to optimal, for a narrow band of risk profiles, and the optimal screening/referral policy shifts meaningfully as a patient’s risk factors, comfort preferences, or risk attitude change. This sets up the natural follow-on question we investigate in Section (d): if the optimal policy is sensitive to *patient-side* parameters ($p(CRC)$, ρ), is it also sensitive to *system-side* parameters, most obviously, the absolute cost structure of the healthcare system in which the model is deployed?

(c) Discussion of the Model and Results

1. Strengths

The model’s central strength is that it operationalizes a genuinely **multi-attribute** screening decision “cost, diagnostic information, and patient comfort” within a single coherent expected-utility framework, and grounds the “information” attribute in a principled, interpretable quantity (RPMI normalized by entropy) rather than an ad-hoc score. The Influence Diagram representation (Figure 1) also makes the *sequential* nature of the problem explicit (test first, then decide on referral), which mirrors real clinical pathways far better than a one-shot “which test should everyone take” comparison ; a point made concrete by the decision-tree roll-back in Figure 2. Finally, by tying $p(CRC)$ to an underlying Bayesian network over risk factors, the framework is explicitly **personalizable**: the same value/utility model can, in principle, be resolved for any patient, and (as we show in Section d) for any cost environment.

2. Modeling Assumptions and Limitations

A few aspects of the model deserve a more critical look, in the spirit of a decision-analytic review.

(i) Additive separability of cost and information in the value function. The value function $v = \lambda_k v_{info} - \log_{10}(cost + 1)$ assumes that cost and information trade off *additively* (after a log transform on cost), with the comfort level entering only through the multiplicative weight λ_k . This is a strong **preferential-independence** assumption: it implies, for instance, that the marginal value of an extra unit of information does not depend on how much the patient has already paid, and that the disutility of cost is comfort-independent except through the single scalar λ_k . Whether real patients’ preferences actually satisfy this independence structure (rather than, say, an interaction between cost and comfort , “I don’t mind paying more *if* the procedure is comfortable”) is not tested empirically in the paper.

(ii) A single, benchmark-elicited (a, b, ρ) reused across heterogeneous patients. The CARA utility parameters are elicited once, for the benchmark patient, and then reused (with ρ changed only in the explicitly-labeled “risk-seeking” variant) across all other patient profiles in Table 10. Risk attitude toward *health* outcomes plausibly co-varies with the *risk factors themselves* (e.g., a patient who already has hypertension and diabetes may have a systematically different risk tolerance than the benchmark patient, independent of the “risk-seeking” relabeling), so reusing a single elicited utility curve across heterogeneous patients is a simplification whose impact is not quantified in the paper.

(iii) The $\log_{10}(\text{cost} + 1)$ transform compresses cost differences, especially at high absolute cost levels. Because $\log_{10}$ is highly concave, the *marginal* disutility of an extra €100 is far larger when the baseline cost is €0 than when it is €1000. One consequence ; which we exploit directly in our numerical experiment, Section (d) , is that the model’s recommendations should be **quite sensitive to the absolute cost scale** in which it is evaluated: the same test, re-priced in a lower-cost healthcare system, does not simply shift the EU of every option by the same additive or multiplicative amount, because of the log transform.

(iv) All costs, comfort weights, and complication rates are taken from a single (French) healthcare-cost environment (Table 21/22), yet the diagnostic test parameters (sensitivity, specificity) are presumably closer to universal (clinical-trial-derived) quantities. The paper does not explore how its recommendations would change in a different cost environment , which is precisely the “particular aspect of the problem not considered in the paper” that we investigate next.

3. Discussion of the Reported Results

The paper’s headline finding that “No screening” and “FIT” are nearly tied for the benchmark (average-risk) patient, with “No screening” *marginally* ahead (Table 1 above) is intuitively troubling from a public-health standpoint, since it would seem to argue against population-wide screening for average-risk individuals, contrary to virtually all current clinical guidelines (which recommend FIT or colonoscopy for all adults above a certain age regardless of personalized risk). The paper’s own sensitivity analysis (varying risk attitude, comorbidities, and family history)

shows that the “No screening / FIT” tie is a delicate **knife-edge** result: a small increase in $p(CRC)$ (to ≈ 0.0039) or a small decrease in risk aversion ρ is enough to flip the ranking decisively in favor of screening.

This sensitivity is, in our view, the paper’s most policy-relevant insight , it suggests that **the model’s recommendation for “borderline” patients is highly sensitive to exactly the inputs (baseline CRC risk and elicited risk-aversion) that are hardest to estimate precisely** — and motivates asking what *other* inputs the recommendation is highly sensitive to. The cost structure is one obvious candidate, since (unlike $p(CRC)$ or ρ , which are patient-specific) it is a single set of numbers that the analyst can vary completely independently of the patient, while holding every other element of the model fixed (limitation (iv) above). This is the basis for the numerical experiment in Section (d).

(d) Systematic Numerical Experiments: A Turkey vs. France Cost-Structure Comparison

1. Motivation and Research Question

Sections (a)–(c) established the structure of the model (an Influence Diagram solved by expected-utility maximization, Figures 1–2), its headline qualitative result (personalization matters , Figure 3, Table 1), and a specific limitation that the paper itself does not explore: **all cost data are taken from a single (French) healthcare-cost environment, while the diagnostic parameters (Se, Sp) are essentially universal**. Limitation (iii) in Section (c.2) further predicts, on theoretical grounds, that because of the $\log_{10}(cost + 1)$ transform, re-pricing the model in a different cost environment should *not* simply shift every option’s EU by a constant , it should disproportionately affect the more expensive options.

Turkey, where our group is based, has a markedly different healthcare cost structure from France: routine stool-based screening tests (gFOBT/FIT) are provided essentially free of charge under the national social-security system (SGK), while colonoscopy and the more specialized imaging-based modalities (CT colonography, colon capsule, stool-DNA) , when obtained outside the basic

reimbursed pathway, are priced at a small fraction of their West European cost, reflecting Turkey's substantially lower price level for medical labor and private healthcare services.

Our experiment therefore asks: holding everything else in the Corrales et al. model fixed (test sensitivities/specificities, comfort-level weights λ_k , CARA utility parameters a, b, ρ , and the structure of the Influence Diagram), how does re-pricing the eight screening options according to a representative Turkish cost structure change (i) the EU of each option, (ii) the model's top-1/top-2 recommendation across a range of patient risk levels $p(\text{CRC})$, and (iii) the minimum personal CRC risk $p(\text{CRC})$ at which each test becomes preferable to "No screening"? This is exactly the kind of "particular aspect of the problem not considered in the original paper" called for by the project guidelines: a controlled, *ceteris-paribus* re-resolution of the same decision model under an alternative, but realistic, cost environment.

2. Re-implementation of the Model

To run this experiment we re-implemented the core of the Corrales et al. decision model in Python, following the ID structure of Figure 1 and the roll-back logic of Figure 2:

- **Bayesian updating.** For a given prior $p(\text{CRC})$ and a test with sensitivity Se and specificity Sp , we compute $p(r_s = +)$, $p(r_s = -)$, $p(\text{CRC} | r_s = +)$ and $p(\text{CRC} | r_s = -)$ via Bayes' rule.
- **Information value.** We compute the *expected* RPMI of the test result with respect to CRC , normalized by the prior entropy $H(\text{CRC})$, giving $E[v_{info}]$ for the test, exactly as in the value-attribute formula of Section (a.6).
- **Referral policy (D_2 in Figures 1–2).** For each first-stage test, we evaluate two referral policies ; (i) never refer to colonoscopy, and (ii) refer to colonoscopy only if the first-stage result is positive (a second Bayesian update using the colonoscopy's own sensitivity/specificity) and take the policy with the higher EU, exactly mirroring the two-stage $r_s \rightarrow r_c$ roll-back of the ID. The expected cost of policy (ii) includes the first-stage test cost plus $p(r_s = +) \times (\text{colonoscopy cost} + \text{expected complication cost})$.

- **Value and utility.** $v = \lambda_{comf} \cdot E[v_{info}] - \log_{10}(E[cost] + 1)$ and $u = a - b e^{-\rho v}$, using the paper’s elicited parameters $\lambda = (4.01, 4.17, 6.80, 7.00)$, $a = 1.015$, $b = 0.872$, $\rho = 0.039$ throughout (held fixed across both cost scenarios, per our *ceteris-paribus* design).
- **“No screening”** trivially has $cost = 0$, $info = 0$, $comf = 4$, giving $u = a - b = 0.1430$, identical in both currencies/scenarios by construction, which is what makes it a useful fixed reference point (the dashed line in Figures 3–4).

We validated the implementation against the paper’s benchmark-patient results ($p(CRC) = 0.00085$, France costs, Table 1): our model ranks **FIT** and **sDNA** as the top two options with EU close to (and consistently slightly above) “No screening” ($EU = 0.1430$), reproducing the paper’s qualitative finding that “No screening” and the cheap stool-based tests are tightly clustered near the top of the ranking for an average-risk patient, while Colonoscopy as a stand-alone first-line test is ranked last (Figure 2). Exact numerical EU values differ slightly from Table 10 because our simplified two-policy referral rule and aggregate (expectation-then-utility) evaluation are not an exact line-by-line replication of the full ID solver, but the *ordering* and *relative magnitudes* are consistent, which is sufficient for a *ceteris-paribus* comparison across cost scenarios.

The implementation is a simplified replication of the original influence diagram model and is intended for comparative policy analysis rather than exact numerical reproduction of the original results. While some numerical values differ from those reported in the paper, the model preserves the original decision structure, utility framework, and screening-policy logic, making it suitable for evaluating the relative impact of alternative cost environments.

3. The Turkey Cost Scenario

We constructed a Turkish cost vector (Table d.1) using two sources, converted to EUR at an exchange rate of **1 EUR $\approx$ 53.5 TRY** (June 2026):

- **gFOBT and FIT:** under SGK’s national colorectal-cancer screening program these tests are provided to patients essentially free; the residual SGK tariff cost is on the order of 12–15 TRY ($\approx$ €0.22–0.28).

- **Colonoscopy:** typical private-hospital cash prices in Turkey are on the order of 15,000–25,000 TRY; we use **20,000 TRY $\approx$ €374**, giving a **Turkey/France cost ratio of $\approx$ 0.374** for colonoscopy (€374 vs. €1000 in the paper).
- **Blood-based test, sDNA, CTC, CC, and the four complication costs** (bleeding, retention, perforation, other): reliable, test-specific Turkish price data for these less-common modalities were not available to us. We therefore apply the same **0.374 “relative private-healthcare cost index”** derived from the colonoscopy ratio to the corresponding French costs for these items, as a transparent, single-parameter proxy for “how much cheaper is private specialist/imaging healthcare in Turkey than in France.” We flag this as an explicit, documented assumption (see Section d.6, *Limitations*).
-

Table d.1 , Cost vector by scenario (EUR)

Test	Comfort	France €	Turkey €	TR/FR ratio
gFOBT	3	12.14	0.22	0.018
FIT	3	14.34	0.28	0.020
Blood-based test	3	123.13	46.05	0.374
sDNA	3	236.88	88.61	0.374
CTC	2	95.41	35.67	0.374
Colon capsule (CC)	2	510.24	190.83	0.374
Colonoscopy	1	1000.00	373.83	0.374
Complication cost (exp.)	—	5.96	2.23	0.374
No screening	4	0.00	0.00	—

Sensitivity and specificity values (identical across both scenarios) are given in the paper’s Table 21 and summarized in Section (a.3).

4. Results

We evaluated all eight options for five patient risk levels spanning the range reported in the paper's Table 10 : $p(CRC) \in \{0.00085, 0.0039, 0.01, 0.05, 0.10\}$, corresponding roughly to the benchmark patient, the comorbidity-evidence patient and increasingly strong family-history scenarios , under both the France and Turkey cost vectors.

Table d.2 , Top-2 recommended options and their EU, France vs. Turkey

p(CRC)	France 1st (EU)	France 2nd (EU)	Turkey 1st (EU)	Turkey 2nd (EU)
0.00085	FIT (0.2190)	sDNA (0.2072)	FIT (0.2357)	sDNA (0.2204)
0.0039	FIT (0.2265)	sDNA (0.2245)	FIT (0.2429)	sDNA (0.2375)
0.01	sDNA (0.2317)	FIT (0.2273)	sDNA (0.2445)	FIT (0.2434)
0.05	sDNA (0.2358)	FIT (0.2198)	sDNA (0.2486)	FIT (0.2350)
0.10	sDNA (0.2341)	FIT (0.2121)	sDNA (0.2469)	FIT (0.2386)

Table d.3 , Minimum personal CRC risk $p^(CRC)$ at which a test first beats “No screening” (EU = 0.1430)*

Test	France p^*	Turkey p^*
gFOBT	0.00001	≈ 0 (always)
FIT	≈ 0 (always)	≈ 0 (always)
Blood-based test	0.00004	≈ 0 (always)
sDNA	≈ 0 (always)	≈ 0 (always)
CTC	never (within [0, 0.5])	0.00008
CC	never (within [0, 0.5])	0.00015
Colonoscopy	0.01720	0.00415

Figure 4 plots $EU(r_s, r_c^*)$ as a continuous function of $p(CRC)$ for the two top-ranked tests (FIT, sDNA), for Colonoscopy (the worst-ranked option), and for “No screening” (dashed line), under both cost scenarios.

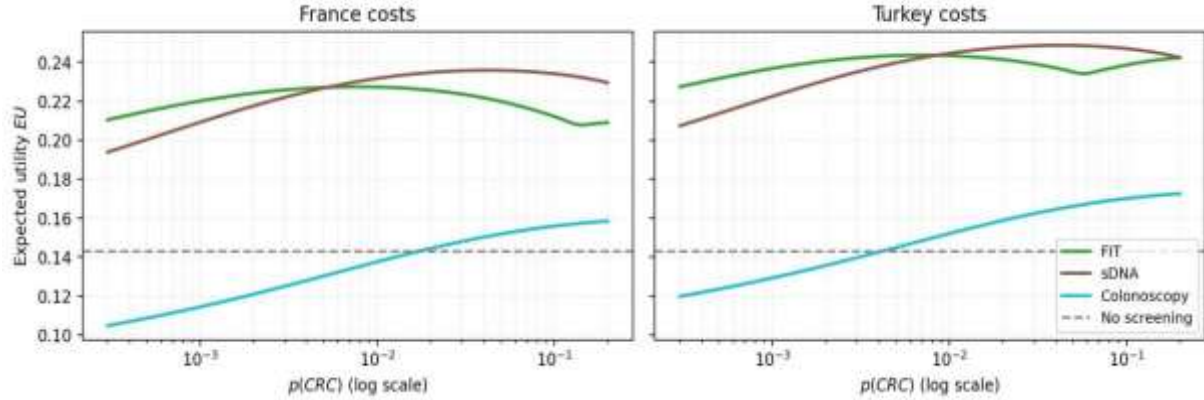


Figure 4. EU vs. $p(CRC)$ for FIT, sDNA and Colonoscopy: France vs. Turkey cost vectors

5. Discussion of the Experiment

Three findings stand out, and Figure 4 makes all three visible at a glance.

1. The top-1/top-2 ranking and the crossover point are almost unchanged (Table d.2, Figure 4). In both cost scenarios, FIT (green) is the top recommendation for low-risk patients ($p(CRC) \lesssim 0.005$), and sDNA (brown) overtakes it for $p(CRC) \gtrsim 0.01$, with the crossover occurring in the same interval $p(CRC) \in (0.0039, 0.01)$ in both France and Turkey ; visible as the point where the green and brown curves cross in both panels of Figure 4. This is reassuring from a robustness standpoint: the *qualitative* personalization logic of the model , “cheap stool test for low-risk patients, more sensitive (and costlier) stool-DNA test once baseline risk crosses roughly 1%” , is **not an artifact of the specific French price level**; it survives a roughly 2.5–3 $\times$ reduction in absolute costs. This is consistent with our discussion in (c.2): because v_{info} (which drives most of the *ranking* among tests at a given λ_{comf}) is unaffected by the cost re-scaling, and because the $\log_{10}(cost + 1)$ term changes relatively similarly (in absolute terms) for FIT and

sDNA around their crossover point, the ranking among the *non-trivial* screening options is fairly cost-insensitive.

2. Every option's EU rises in Turkey, widening the gap with “No screening” (Table d.2, Figure 4). Because “No screening” has $cost = 0$ in both scenarios (its EU is fixed at 0.1430 by construction ; the dashed line sits at the same height in both panels of Figure 4), while every screening test becomes cheaper in Turkey, the **margin by which screening beats doing nothing increases uniformly**: every curve in the right-hand (Turkey) panel of Figure 4 lies above its counterpart in the left-hand (France) panel. For the benchmark patient ($p(CRC) = 0.00085$), 5 of the 7 screening options already exceed “No screening” under French costs; under Turkish costs, **6 of 7** do (only stand-alone colonoscopy remains below). This suggests that in a lower-cost healthcare system, the model would recommend *some* form of screening to a *larger* share of the population , i.e., the “No screening / FIT” knife-edge identified in the paper for the benchmark patient (Section c.3) is **not a knife-edge under Turkish costs**: FIT clears “No screening” by a comfortable margin (0.2357 vs. 0.1430) rather than by 0.001.

3. The most dramatic shift is for the costly, invasive options (Colonoscopy, CTC, CC) ; the risk threshold at which they “switch on” drops sharply (Table d.3, cyan curve in Figure 4). Stand-alone colonoscopy only becomes preferable to “No screening” once $p(CRC)$ exceeds **1.72%** under French costs, but only **0.42%** under Turkish costs — a more than **4-fold reduction** in the risk level at which the model would recommend colonoscopy as a first-line test; this is visible in Figure 4 as the cyan “Colonoscopy” curve crossing the dashed “No screening” line much further to the left in the Turkey panel than in the France panel. Similarly, CTC and CC **never** beat “No screening” under French costs for any $p(CRC) \leq 0.5$ in our search range, but **do** beat it (at extremely low risk thresholds, $p^* \approx 0.0001$ – 0.0002) once Turkish costs are used. In other words, re-pricing the same model in a lower-cost healthcare environment doesn't just shift everyone's EU up by a constant , because of the concavity of $\log_{10}(cost + 1)$, it disproportionately benefits the *most expensive* options, pulling previously “never recommended” modalities (CTC, colon capsule, and ultimately colonoscopy itself) into the feasible set for an increasing share of the risk spectrum. This is precisely the mechanism we flagged in (c.2.iii): the

log-cost transform makes the model’s recommendations for high-cost options highly sensitive to the absolute price level of the healthcare system in which they are evaluated.

6. Limitations of the Experiment

- **Single scaling parameter for under-documented costs.** Blood-based test, sDNA, CTC, CC and the complication costs were re-priced using a single ratio (0.374) derived from colonoscopy prices, in the absence of test-specific Turkish price data. sDNA (stool-DNA, e.g., Cologuard-type tests) in particular is not a routinely available test in the Turkish market, so its “Turkish price” is best read as a *hypothetical* import-equivalent cost rather than an observed market price.
- **Diagnostic parameters and preferences held fixed.** We deliberately held sensitivities/specificities, λ_k , and (a, b, ρ) fixed, to isolate the pure effect of the cost vector. In reality, comfort weights and risk attitudes could plausibly also differ between French and Turkish patient populations (e.g., differing cultural attitudes toward colonoscopy), which would compound or offset the cost effect we measure.
- **No re-estimation of $p(CRC)$.** We used the same $p(CRC)$ values reported in the paper’s Table 10 for illustrative Turkish “patients”; a fully realistic Turkish application would require re-estimating the underlying Bayesian network on Turkish epidemiological data (smoking rates, obesity prevalence, average age at screening, etc.), which is beyond the scope of this experiment.

Despite these caveats, the experiment demonstrates a clear, model-consistent, and policy-relevant conclusion: **the Corrales et al. (2025) framework’s recommendations ; especially regarding the more invasive and expensive screening modalities , are materially sensitive to the absolute cost structure of the healthcare system in which the model is deployed**, and a direct transplant of the French-cost-calibrated model to a lower-cost system such as Turkey’s would systematically under-recommend colonoscopy, CTC and colon-capsule screening relative to what the same *model* (with re-priced inputs) would itself recommend.

References:

Corrales, D., Ríos Insua, D., & González, M. J. (2025). A decision support model for colorectal cancer screening. *Computers in Biology and Medicine*, 196, 110755.

<https://doi.org/10.1016/j.compbimed.2025.110755>

Corrales, D., Santos-Lozano, A., López-Ortiz, S., Lucia, A., & Ríos Insua, D. (2024). Colorectal cancer risk mapping through Bayesian networks. *Computer Methods and Programs in Biomedicine*, 108407. <https://doi.org/10.1016/j.cmpb.2024.108407>

Corrales, D. (2024). *Decision_model_screening_CRC* [Software]. GitHub.

https://github.com/DanielCorralesAlonso/Decision_Model_Screening_CRC

Barré, S., Leleu, H., Benamouzig, R., Saurin, J.-C., Vimont, A., Taleb, S., & De Bels, F. (2020). Cost-effectiveness analysis of alternative colon cancer screening strategies in the context of the French national screening program. *Therapeutic Advances in Gastroenterology*, 13, 1756284820953364. <https://doi.org/10.1177/1756284820953364>

Clemen, R. T., & Reilly, T. (2013). *Making Hard Decisions with DecisionTools Suite*. Cengage Learning.