Leveraging MCDA for Colorectal Cancer Screening Strategies

Daniel Corrales[1, 2] David Ríos-Insua[1]

[1]Instituto de Ciencias Matemáticas (ICMAT - CSIC), España
[2]Universidad Autónoma de Madrid (UAM), España

May 9, 2025

Table of contents

Context I

- Colorectal cancer (CRC) is the third most common cancer worldwide. [WHO, 2023]
- Only around 14% of susceptible EU citizens participate in screening programs.
- Early-onset incidence is rising.
- In 2020, 1.9M new cases and 930,000 associated deaths, worldwide. [Morgan et al., 2023]
- In 2015, its estimated annual costs were approximately 19 billion €. [Henderson et al., 2021]

Goal: Create predictive models and decision support tools for personalised CRC screening

Characterizing CRC Risk [Corrales et al., 2024] I

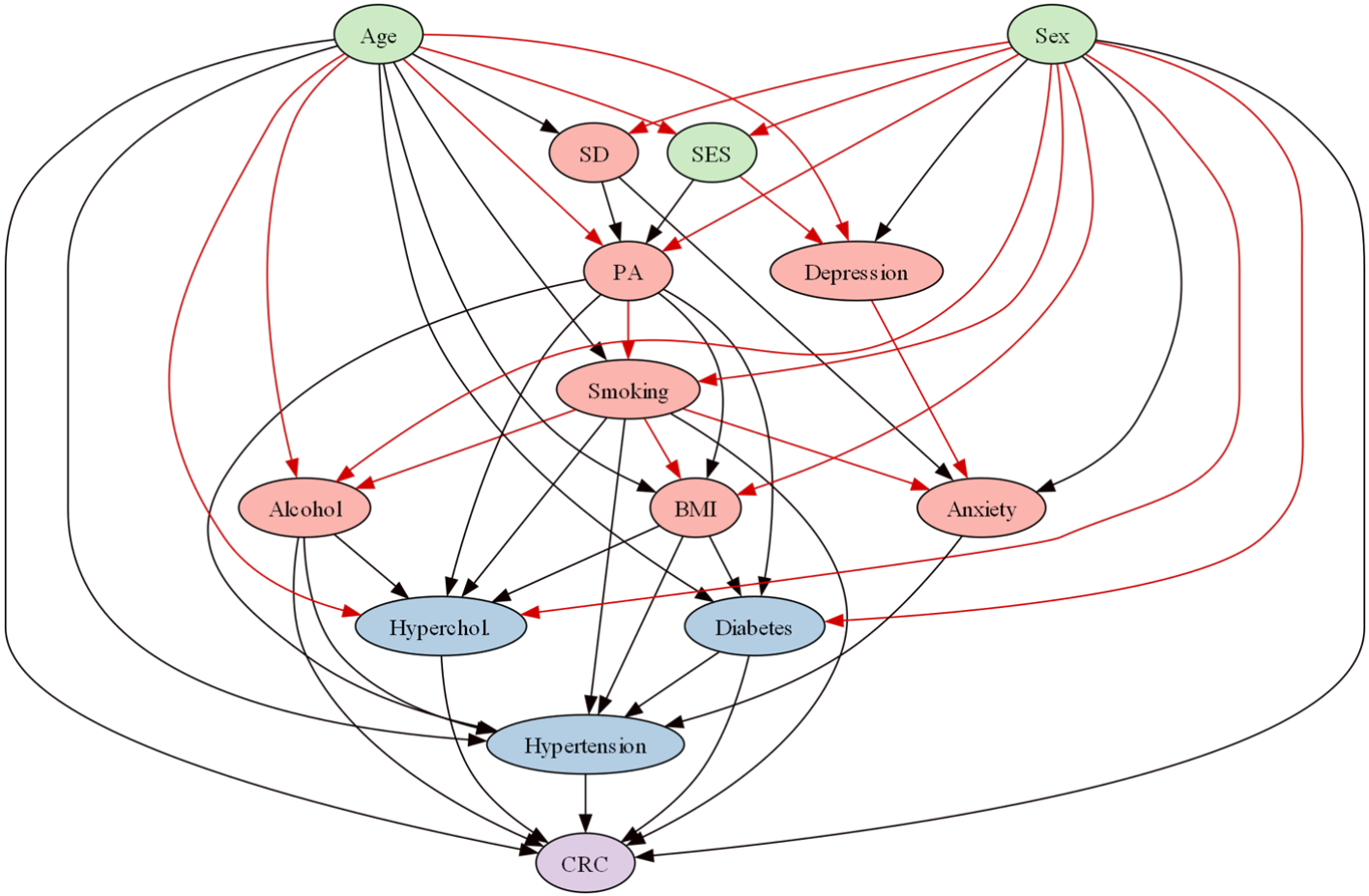
- Bayesian Networks (BNs) represent a natural framework to analyse dependence across CRC risk factors
- Data from annual health reviews database enriched through INE datalake. Total of 2M records and 66 variables.
- Kept non-modifiable (e.g. Age) and modifiable (e.g. BMI) risk factors, together with medical conditions (e.g. Hypertension).
- Literature review validated by experts.
- Variables discretized and an intense data cleaning work performed.
- Learn BN structure from expert judgement and data.

$$p(\theta_{X|\boldsymbol{u}}) \sim Dir(\alpha_{x^1|\boldsymbol{u}}, ..., \alpha_{x^K|\boldsymbol{u}}) \quad \text{(Empirical Bayes prior)}$$

$$p(\theta_{X|\boldsymbol{u}}|D) \sim Dir(\alpha_{x^1|\boldsymbol{u}} + m[\boldsymbol{u}, x^1], ..., \alpha_{x^K|\boldsymbol{u}} + m[\boldsymbol{u}, x^K]) \quad \text{(Posterior based on data)}$$

$$\hat{\theta}_{X=x^i|\mathbf{u}} = \frac{\alpha_{x^i|\mathbf{u}} + m[\boldsymbol{u}, x^i]}{\sum_j \alpha_{x^j|\mathbf{u}} + m[\mathbf{u}, \mathbf{x^j}]} \quad \text{(MAP estimator)}$$

Resulting Bayesian Network



Use cases

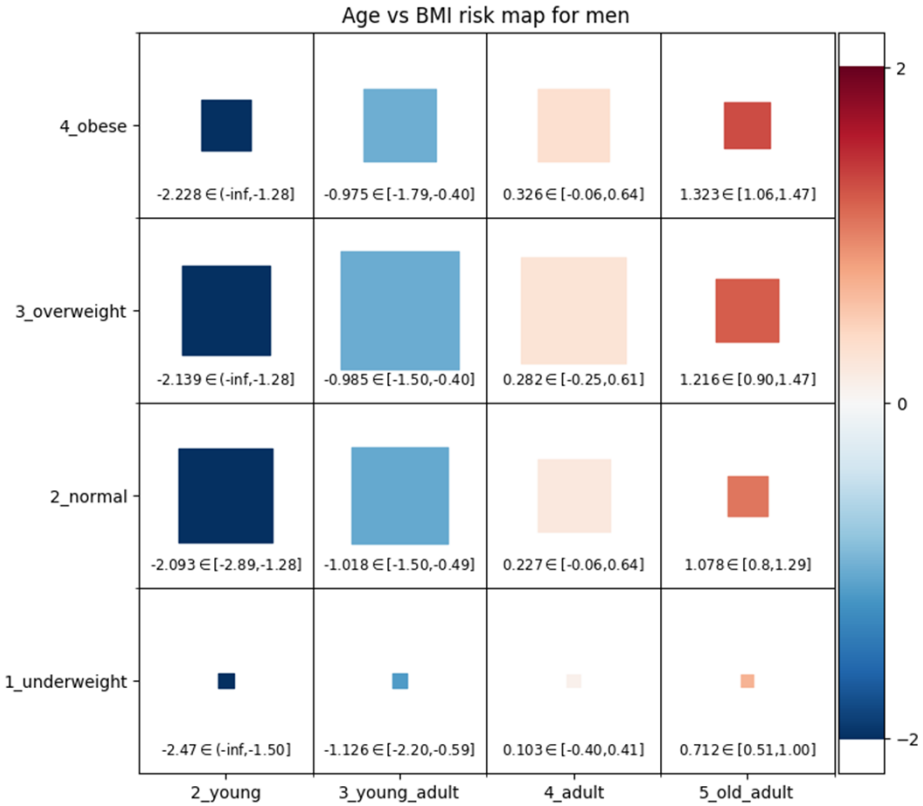


Figure: Risk map

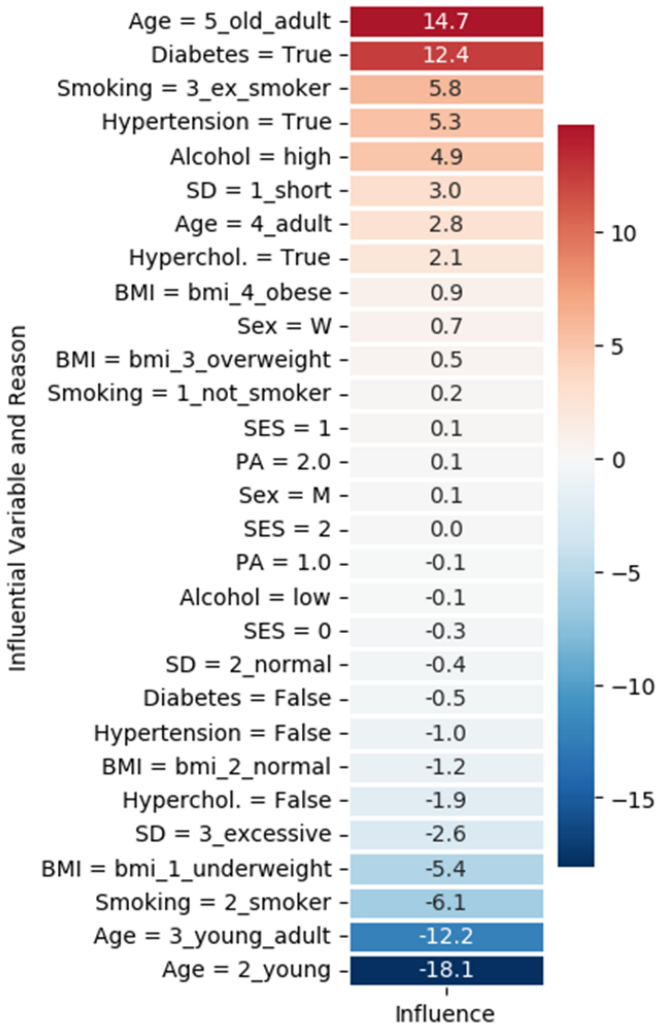


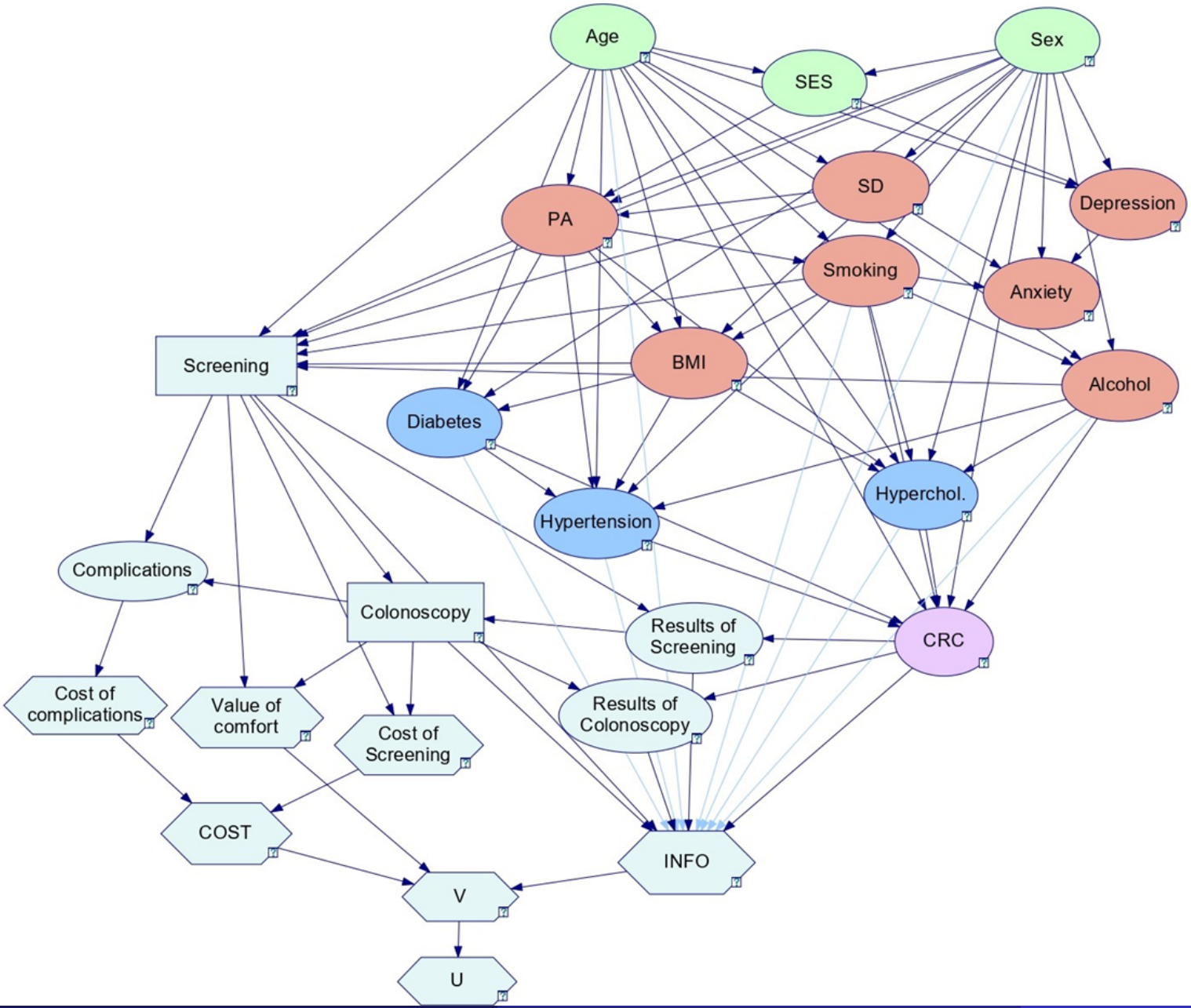
Figure: Influential variables

Decision model for CRC screening [Corrales et al., 2025] I

Aims:

1. Provide personalised advice for screening
2. Screening followed by colonoscopy if positive screening test
3. Focus on short-term information outcomes, costs and comfort.
4. Decision based on maximum expected utility.

Decision model for CRC screening [Corrales et al., 2025] II



Quantifying the decision model I

1. Probability models at chance nodes
   Distribution of results of screening and colonoscopy based on sensitivity and specificity of screening tools and patient features $p(R_S|crc, x)$, $p(R_C|crc, x)$. Additionally, probability of complications.

	gFOBT	FIT	BldBsd	sDNA	CTC	CC	Colons.
Sensitivity	0.45	0.75	0.66	0.923	0.8	0.87	0.97
Specificity	0.978	0.966	0.91	0.866	0.89	0.92	0.99

Quantifying the decision model II

2. Single criterion preferences
   - Costs (€) of interventions and complications
   - Comfort. Use a constructed attribute [Keeney and Gregory, 2005] based on level of discomfort

com	Description	Interventions
4	The patient does not experience any discomfort	No screening
3	The patient experiences a minor discomfort or the test implies a small inconvenience: time lost, emotional difficulty, or slight physical pain.	FIT, gFOBT, sDNA, Blood-test
2	The discomfort experienced by the patient is noticeable. There is a noteworthy emotional aversion and a few moments of physical discomfort.	CTC, CC
1	The discomfort is very significant. The test causes some periods of pain resulting in remarkable distress.	Colonoscopy

Figure: Comfort constructed attribute

Quantifying the decision model III

- Information v_{info} provided by the screening strategy. Measured in terms of *relative pointwise mutual information*. Intuition: proportion of uncertainty resolved by the screening strategy.

$$v_{info}(crc, R_S, R_C) = \frac{\log\left(\frac{p(crc|R_S)}{p(crc)}\right) + \log\left(\frac{p(crc|R_S,R_C)}{p(crc|R_S)}\right)}{-\sum p(crc)\log p(crc)} \quad (1)$$

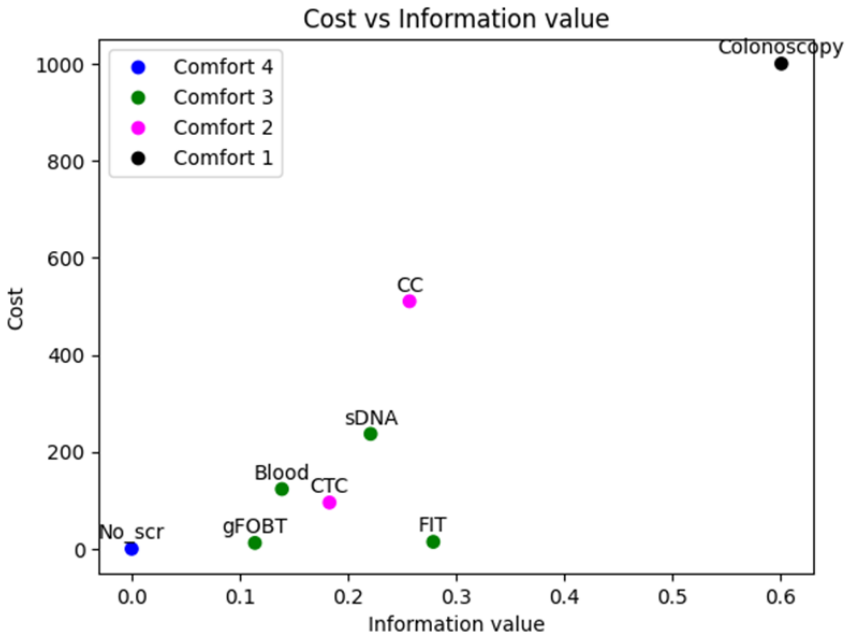


Figure: Cost vs Info for a fixed $p(crc)$

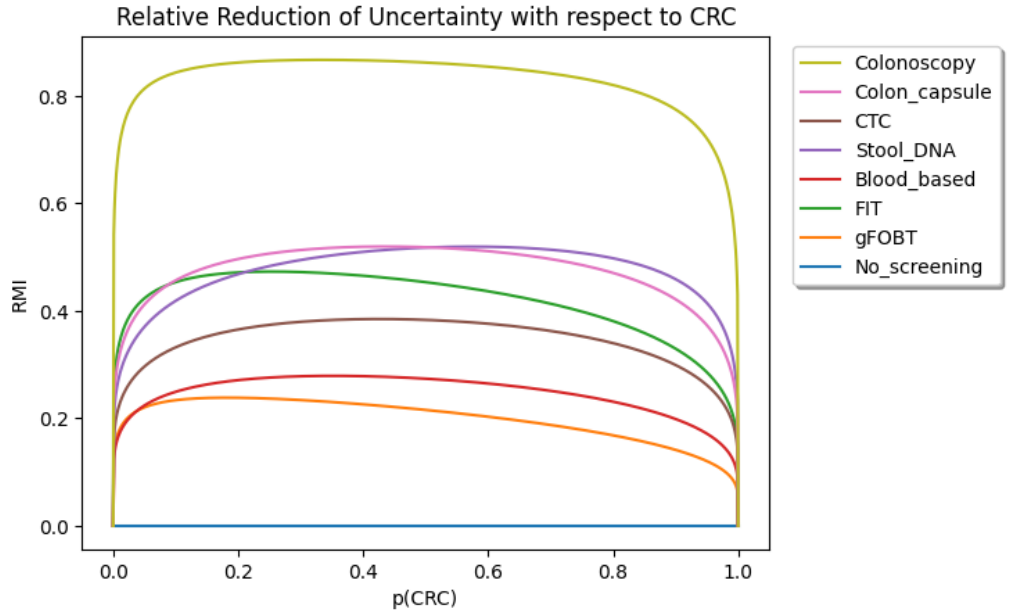


Figure: v_{info} for various screening alternatives

Quantifying the decision model IV

3. Multiple criteria preference aggregation
   A multicriteria value function aggregation followed by a risk aversion transformation used, where λ_k is a weighting factor that depends on comfort. Assume constant absolute risk-aversion (CARA) and elicit parameters a, b, ρ using the probability equivalent method (PE)

$$v(cost, value, comf) = \lambda_k \times value - \log_{10}(cost)$$
$$u(cost, value, comf) = a - b \times \exp(-\rho \times v(cost, value, comf))$$

Quantifying the decision model V

Comfort	Scr. method	Cost	Info	Preference	Indiff. cost	$\hat{\lambda}_k$	λ_k
1	Colonos	1000	0.530		–	$\lambda_1 = 4.01$	$\lambda_1 = 4.01$
	Synth.	–	0.4	×	300€		
2	CTC	95.41	0.159	×	–	$\lambda_2 = 4.17$	$\lambda_2 = 4.17$
	CC	510.24	0.225		180€		
3	gFOBT	12.14	0.129		3€	$\lambda_3 = 5.04$	$\bar{\lambda}_3 = 6.80$
	FIT	14.34	0.245	×	–		
3	gFOBT	12.14	0.128	×	–	$\lambda_3 = 10.57$	
	Blood test	125.13	0.121		10€		
3	gFOBT	12.14	0.128	×	–	$\lambda_3 = 16.28$	
	sDNA	236.88	0.197		170€		
3	FIT	14.34	0.244	×	–	$\lambda_3 = 6.40$	
	Blood test	125.13	0.121		1.5€		
3	FIT	14.34	0.244	×	–	$\lambda_3 = 7.2$	
	sDNA	236.88	0.197		6€		
3	Blood test	125.13	0.121		80€	$\lambda_3 = 6.17$	
	sDNA	236.88	0.197	×	–		
4	No scr.	0	0	–	–	–	$\lambda_4 = 7$

Figure: Elicitation of parameters through the probability equivalent method

Use cases I

- Individual recommendations. Personalised screening strategy depending on risk.

	No_screening	0.13376699
	gFOBT	0.11842714
▶	FIT	**0.13416683**
	Blood_based	0.11203061
	Stool_DNA	0.13116387
	CTC	0.12432938
	Colon_capsule	0.10352935

Figure: Best strategy for a male adult, age 44-54, with normal sleep duration, physically active, normal weight, non-smoker, and with low alcohol consumption.

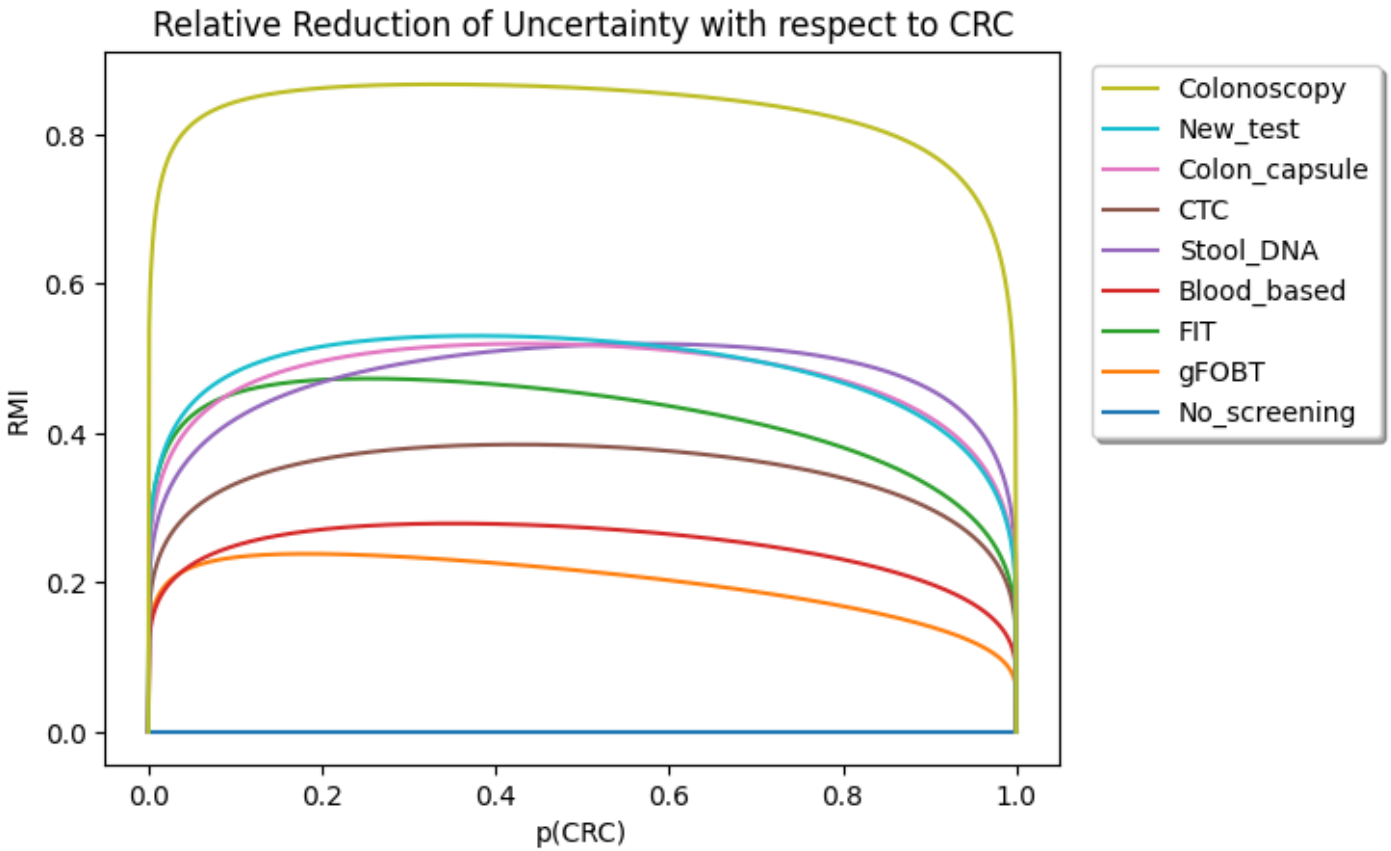
Use cases II

- Assessing the Spanish strategy on screening and designing a national strategy with budget and device constraints

Current Spanish Strategy	**Proposed Strategy**
• If patient >50 years old: Send FIT invitation regardless their features. • Else: No screening.	For patient with features X: • if $p(CRC\|X) < \theta_1 \rightarrow$ No screening • if $\theta_1 \leq p(CRC\|X) < \theta_2 \rightarrow$ FIT • if $p(CRC\|X) \geq \theta_2 \rightarrow$ sDNA
Example: Man under 50 with high risk (due to e.g. high alcohol consumption, overweight, exsmoking) would be detected by the model as a higher risk patient than a healthy man above 60.	
Results: Extrapolating results in Spain's target population, this strategy would detect 134 patients more.	

Use cases III

- Benchmarking for new screening devices (relevant for EU ONCOSCREEN project)



Use cases IV

	Predicted No CRC	Predicted CRC
No CRC	339472.1 ± 4.0	16.9 ± 4.0
CRC	139.8 ± 4.4	78.3 ± 4.4

Figure: Old strategy. Cost per patient 7.14€. F1 classification score 0.50

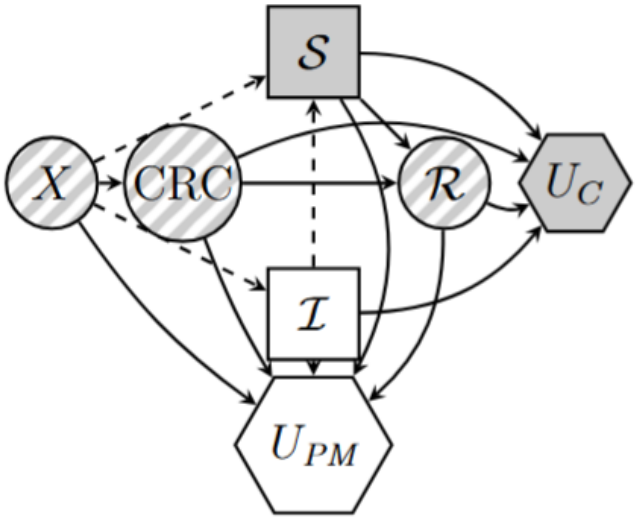
	Predicted No CRC	Predicted CRC
No CRC	339458.5 ± 5.4	30.5 ± 5.4
CRC	126.8 ± 3.8	91.2 ± 3.8

Figure: New strategy. Cost per patient 9.85€. F1 classification score 0.54

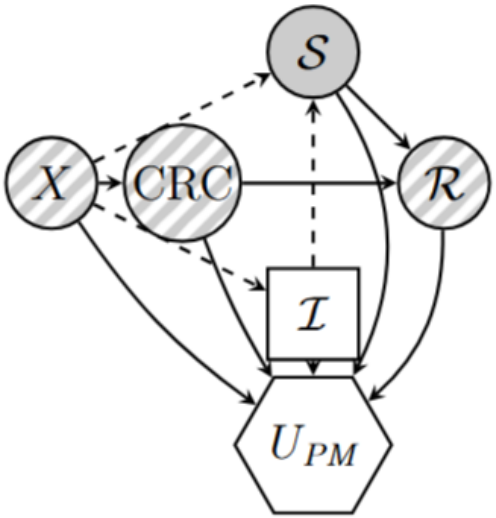
Ongoing work: Designing screening incentives through adversarial risk analysis I

- Suppose a policy-maker (PM) has chosen a screening strategy e.g. through the previous decision model. This does not mean that patients will be willing to participate in the screening program.
- Context: principal-agent problem. Reframe the usual framework through the Adversarial Risk Analysis framework.
- Modeled as a bi-agent influence diagram, which can be solved as in [González-Ortega et al., 2019].

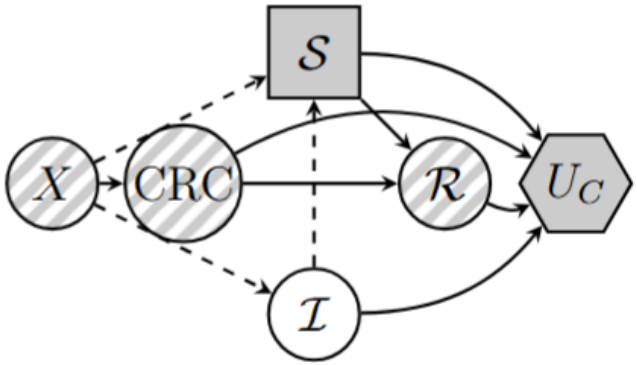
Ongoing work: Designing screening incentives through adversarial risk analysis II



(a) CRC screening incentive problem as principal-agent case from an ARA perspective

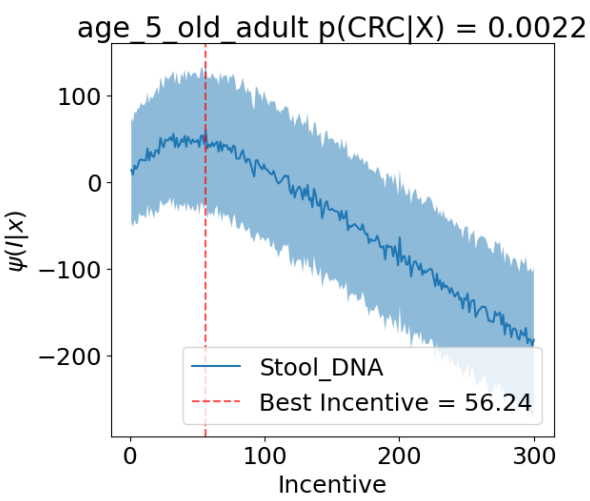


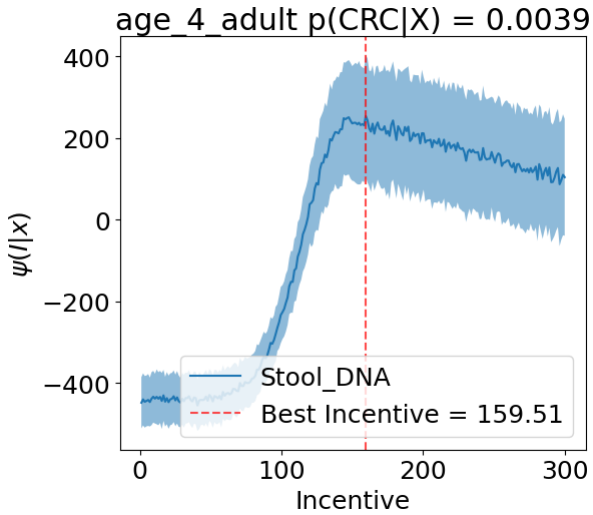
(b) Incentive problem from the PM perspective.

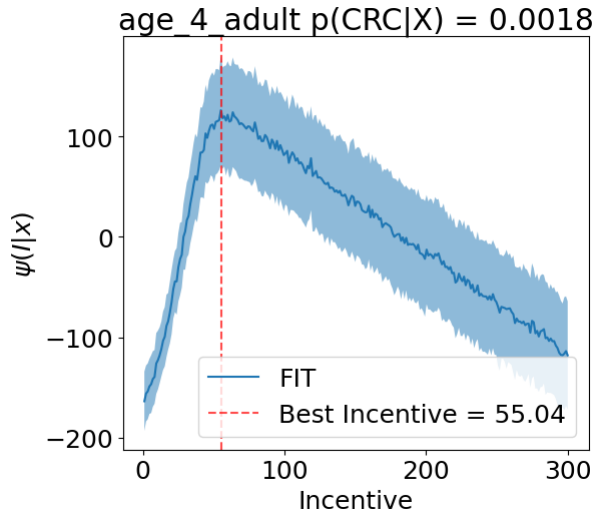


(c) Incentive problem from C's perspective.

Ongoing work: Designing screening incentives through adversarial risk analysis III







Take home messages

- We have developed a model for predicting CRC risk.
- This is embedded in a decision model with multiple criteria (cost, comfort, information provided)
- After the decision model is characterized, it can be used for individual recommendations, designing national screening strategies and benchmarking new devices.
- Decision models based on personalised risk approaches can be very relevant for early cancer detection.

Bibliography I

D. Corrales, A. Santos-Lozano, S. López-Ortiz, A. Lucia, and D. R. Insua. Colorectal cancer risk mapping through bayesian networks. *Computer Methods and Programs in Biomedicine*, 257:108407, 2024.

D. Corrales, D. R. Insua, and M. J. González. A decision analysis model for colorectal cancer screening. *arXiv preprint arXiv:2502.21210*, 2025.

J. González-Ortega, D. R. Insua, and J. Cano. Adversarial risk analysis for bi-agent influence diagrams: An algorithmic approach. *European Journal of Operational Research*, 273(3):1085–1096, 2019.

R. H. Henderson, D. French, T. Maughan, R. Adams, C. Allemani, P. Minicozzi, M. P. Coleman, E. McFerran, R. Sullivan, and M. Lawler. The economic burden of colorectal cancer across europe: a population-based cost-of-illness study. *The lancet Gastroenterology & hepatology*, 6(9):709–722, 2021.

R. L. Keeney and R. S. Gregory. Selecting attributes to measure the achievement of objectives. *Operations Research*, 53(1):1–11, 2005.

Bibliography II

E. Morgan, M. Arnold, A. Gini, V. Lorenzoni, C. Cabasag, M. Laversanne, J. Vignat, J. Ferlay, N. Murphy, and F. Bray. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from globocan. *Gut*, 72(2):338–344, 2023.

WHO. Colorectal cancer, 2023. `https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer`.

Thank You!

Feel free to reach out:

daniel.corrales@icmat.es

ICMAT
INSTITUTO DE CIENCIAS MATEMÁTICAS

CSIC
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

ONCOSCREEN