

# Aplicaciones en ciencia de datos

---

Modelo predictivo basado en síntomas para el diagnóstico de  
cáncer colorrectal

**Profesores:**

Jean Paul Maidana González, PhD  
David Araya Gálvez, PhD

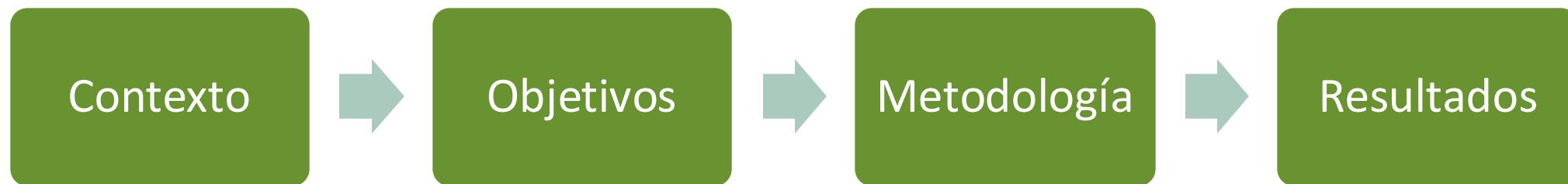
**Estudiantes**

Brian Guzmán  
Eric Silva

Noviembre - 2025



Estudio analizado: Modelo predictivo basado en síntomas para CCR  
(Revista Médica de Chile, 2025)

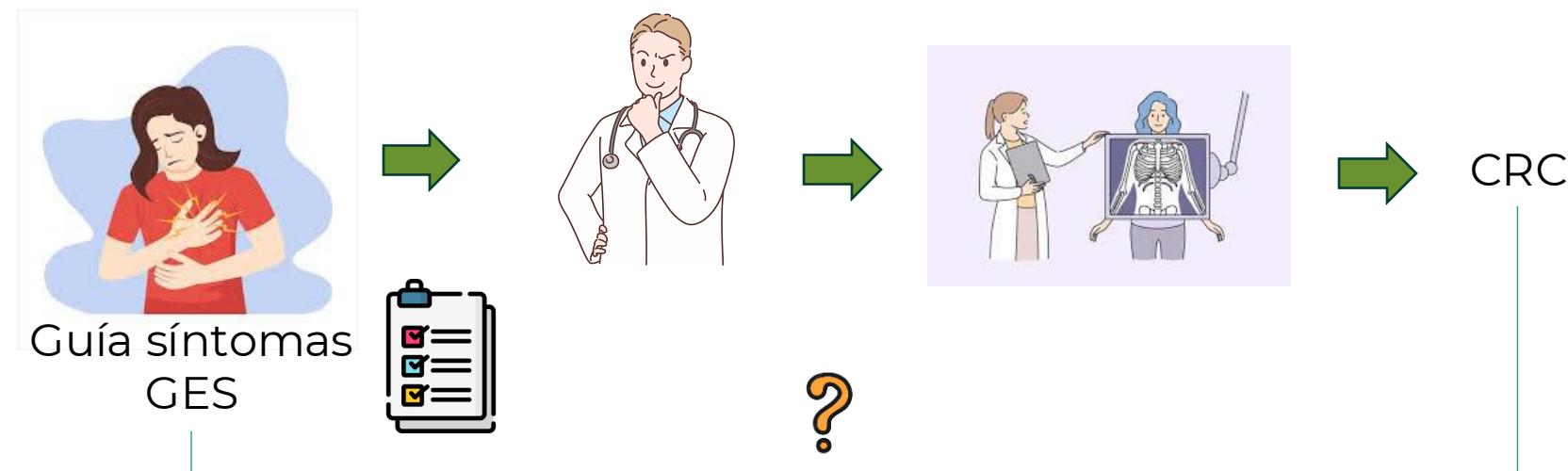


## ¿POR QUÉ ES RELEVANTE?

- ✓ Aplicación real en Chile
- ✓ Técnicas de ML en salud pública
- ✓ Impacto en políticas públicas (GES)
- ✓ Optimización de recursos limitados



- Sistema público de Salud (Hospital Guillermo Grant Benavente – Concepción)
- Marco garantías explicitas en salud (GES)
- Sintomatología
- Diagnóstico CCR



## Objetivo General :

Desarrollar un modelo predictivo basado en síntomas clínicos para el diagnóstico de cáncer colorrectal, optimizando la priorización de pacientes derivados a colonoscopia según las directrices del GES.

## Objetivos Específicos :



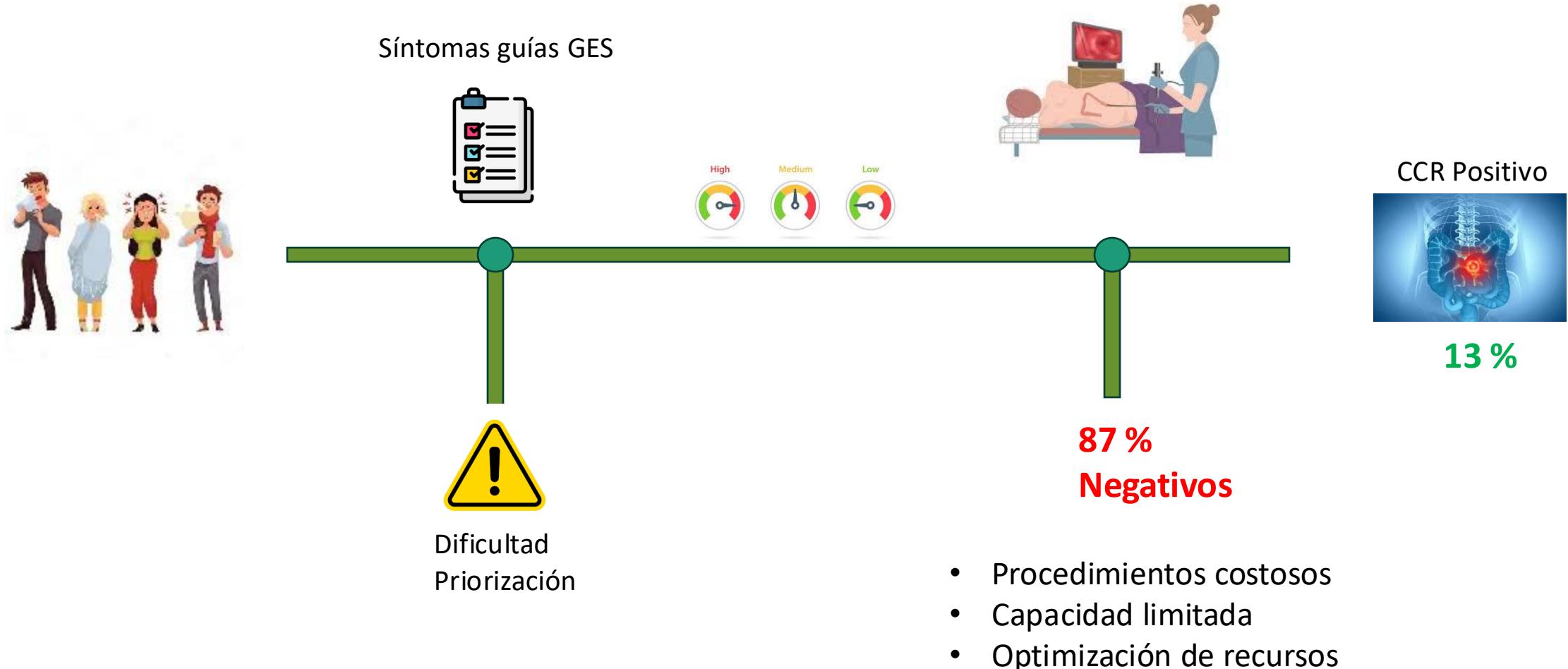
**1. Identificar las variables clínicas y sintomáticas** más relevantes asociadas al diagnóstico de CCR en pacientes derivados según protocolo GES.



**2. Desarrollar un modelo de regresión logística** que estime el riesgo de CCR basándose en síntomas y hallazgos clínicos.



**3. Validar el rendimiento del modelo** mediante métricas estadísticas (sensibilidad, especificidad, AUC-ROC).



## Metodología : Estudio retrospectivo analítico

Universo de pacientes: 1099

Variables evaluadas: 16

Periodo del estudio:  
Julio 2016 – diciembre 2021

Lugar :  
Hospital Guillermo Grant Benavente - Concepción

## Técnica principal: regresión logística binaria

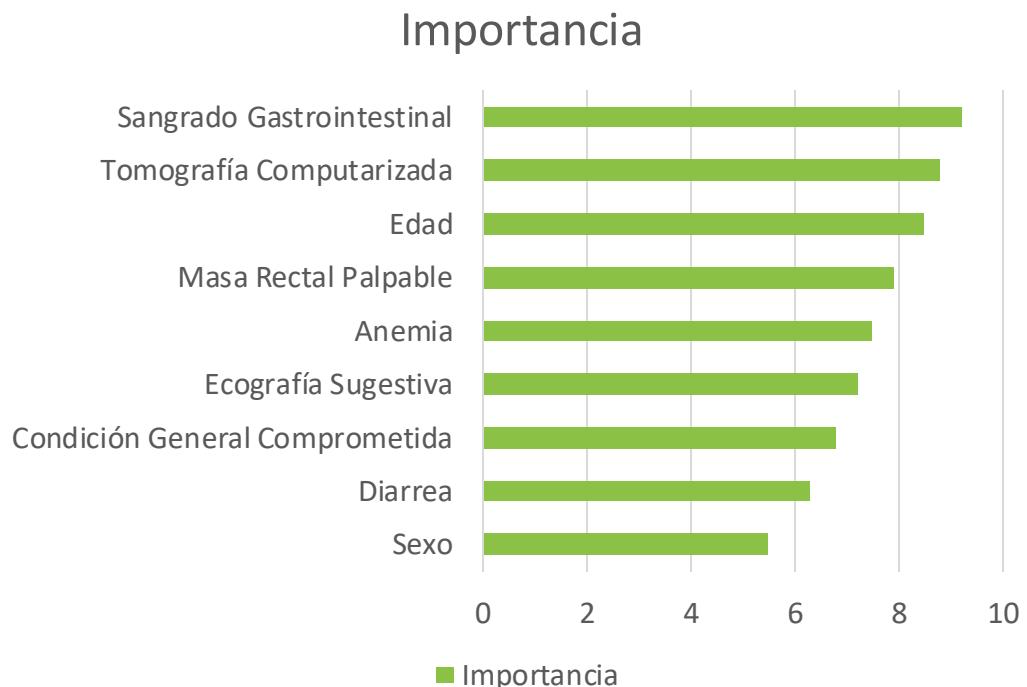
Técnica apropiada para variable dependiente binaria (tiene / no tiene CCR).

Ampliamente aceptada en investigación médica.

Permite incorporar múltiples predictores simultáneamente.

# Resultados Obtenidos

## Importancia de variables en el estudio



### Variables GES (ya en protocolo)

- ✓ Sangrado GI      ( $p<0,001$ )
- ✓ Edad                ( $p=0,002$ )
- ✓ Diarrea             ( $p=0,008$ )
- ✓ Masa rectal        ( $p=0,003$ )
- ✓ Anemia             ( $p=0,011$ )

### Variables NO Significativas

- Pérdida de peso      ( $p=0,12$ )
- Dolor abdominal      ( $p=0,28$ )
- Estreñimiento        ( $p=0,45$ )
- Cambio hábito intestinal      ( $p=0,33$ )
- Antecedentes familiares CCR ( $p=0,19$ )
- Leucocitosis            ( $p=0,51$ )
- Colonoscopía previa    ( $p=0,38$ )

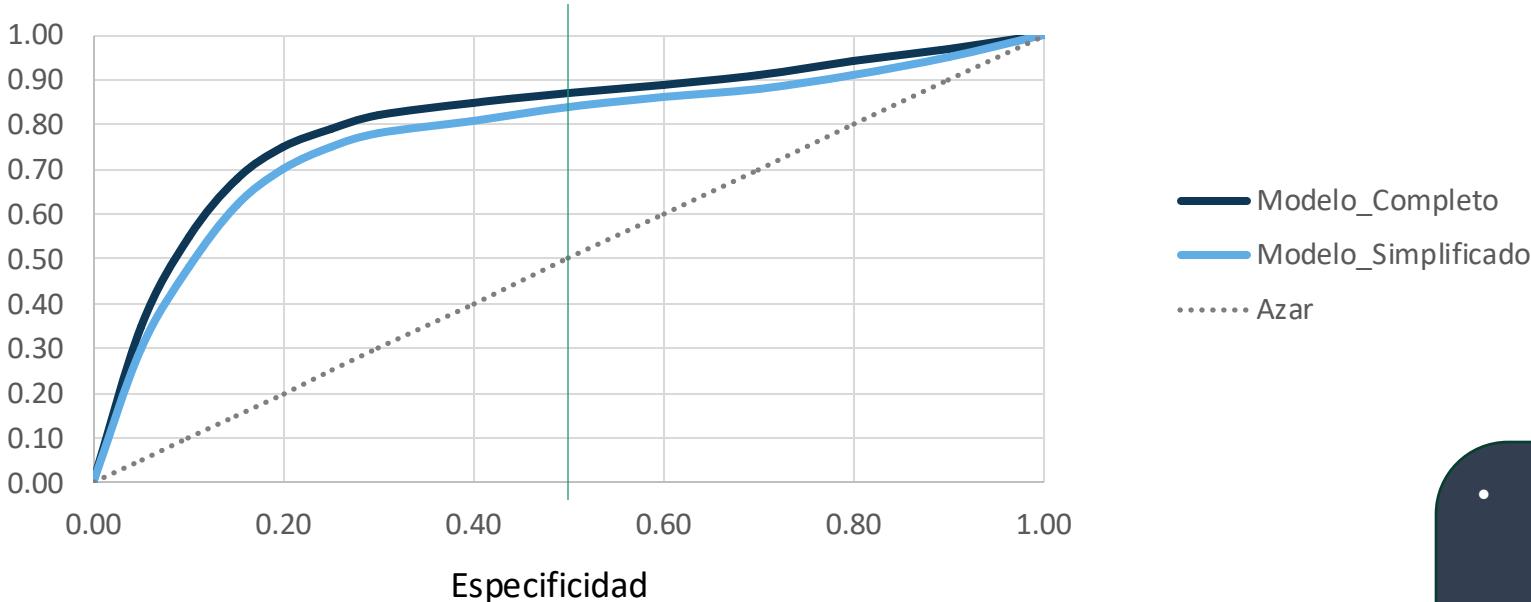
### Variables nuevas (no están en protocolo)

- ✓ Tomografía CT      ( $p<0,001$ )
- ✓ Ecografía            ( $p=0,004$ )
- ✓ Condición general    ( $p=0,015$ )
- ✓ Sexo                  ( $p=0,023$ )

Criterio:  $p<0,05$  para inclusión en modelo. (p) Valor de probabilidad.

## Rendimiento de los modelos predictivos

Sensibilidad



Interpretación AUC:

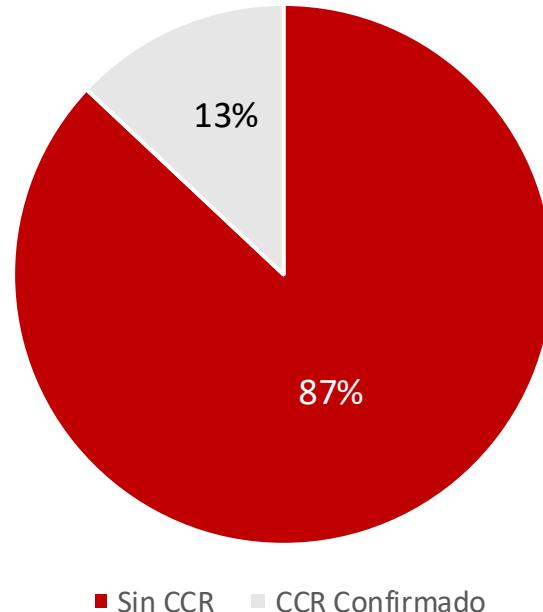
- 0,90-1,00 = Excelente
- 0,80-0,90 = Muy bueno
- 0,70-0,80 = Bueno
- 0,60-0,70 = Regular
- 0,50-0,60 = Malo

- Modelo Completo (9 variables): AUC = **0,86**  
→ Rendimiento MUY BUENO
- Modelo Simplificado (7 variables): AUC = **0,81**  
→ Rendimiento MUY BUENO
- Ambos modelos superan significativamente la línea de azar (AUC = 0,50)

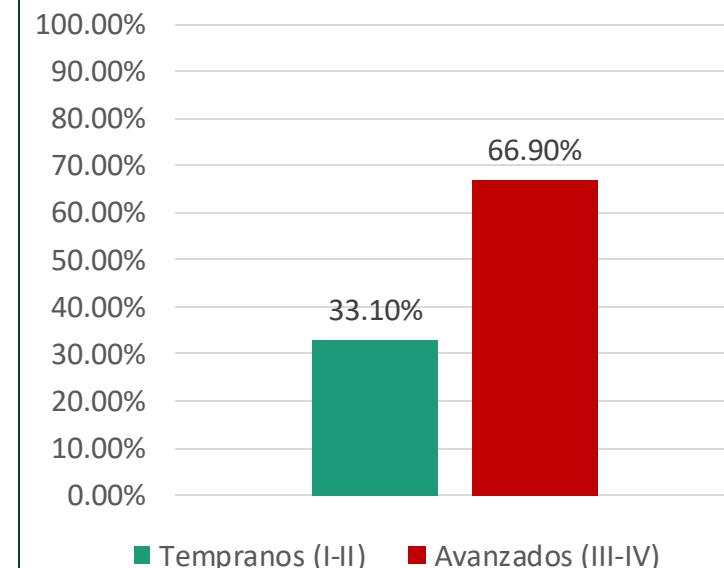
El modelo completo supera en 5 puntos al simplificado, demostrando el valor de incluir estudios de imagen.

## Impacto Clínico : El Problema de la Detección Tardía

Tasa de Detección



Estadificación al Diagnóstico  
(% de pacientes)



### IMPACTO EN SUPERVIVENCIA A 5 AÑOS:

Estadio I-II (temprano)	: 90% supervivencia
Estadio III (avanzado)	: 64% supervivencia
Estadio IV (metastásico)	: 14% supervivencia

→ La detección temprana multiplica por 6 las probabilidades de supervivencia

2 de cada 3 pacientes (67%) se diagnostican en estadios avanzados, reduciendo opciones terapéuticas y supervivencia.

## Optimización del protocolo GES

### VENTAJAS:

- Muy buen rendimiento (AUC = 0,86)  
→ Supera estándares clínicos
- Uso de variables accesibles  
→ Aplicable en sistema público
- Validación robusta ( $n = 1.099$ )  
→ 5 años de datos retrospectivos
- Compatible con GES actual  
→ No requiere cambio radical

### DESAFÍOS:

- Estudio retrospectivo  
→ Necesita validación prospectiva
- Un solo centro hospitalario  
→ Generalización limitada
- Falta análisis de costo-beneficio  
→ Implementación incierta
- No considera comorbilidades complejas  
→ Población real más heterogénea
- Dependencia de estudios imagenológicos  
→ Disponibilidad variable en regiones

### SI GES ADOPTARA LAS 4 VARIABLES NUEVAS:

Modelo actual GES: 5 variables validadas  
Modelo propuesto: 9 variables (+ 4 nuevas)

### IMPACTO ESTIMADO (por cada 1.000 pacientes):

- Mejora en detección temprana: +12-15%
  - Reducción diagnóstico tardío: -8-10%
  - Vidas salvadas potenciales: 8-12 pacientes
  - Mejora supervivencia a 5 años: +18-22%
- Costo adicional: Mínimo (variables clínicas)  
→ Beneficio: Alto impacto en salud pública

### CONCLUSIÓN CRÍTICA:

- La incorporación de 4 variables adicionales al protocolo GES podría mejorar significativamente la detección temprana de CCR sin aumentar sustancialmente los costos operacionales del sistema público.



Se analizaron 5 años de datos de más de mil pacientes



Se encontraron 4 variables que el protocolo actual no considera y 7 que no tienen relevancia a la hora de detectar CCR lo cual mejora el nivel predictivo del modelo.

## Problema de Negocio



## Datos y Preparación



## Modelamiento



## Evaluación



- Benavides, C., & Alvarado, J. (2025). Modelo predictivo basado en síntomas para el diagnóstico de cáncer colorrectal: Optimización según las directrices de la política pública de salud chilena. Revista Médica De Chile, 153 (03) [citado en Abr 4, 2023]. Recuperado a partir de <https://revistamedicadechile.cl/index.php/rmedica/article/view/10992>
- WHO. Colorectal cancer – IARC. [citado en May 14, 2023]. Recuperado a partir de: <https://www.iarc.who.int/cancer-type/colorectal-cancer/>
- Colorectal cancer trends in Chile: A Latin-American country with marked socioeconomic inequities PLOS ONE. [citado en May 14, 2023]. Recuperado a partir de: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0271929>
- Cancer today. [citado en August 9, 2022]. Recuperado a partir de: <http://gco.iarc.fr/today/home>
- Colorectal Cancer Awareness Month 2021–IARC. [citado en September 29, 2021]. Recuperado a partir de: <https://www.iarc.who.int/featured-news/ccam2021/>



Universidad  
Andrés Bello®