

# Leave No Patient Behind:

## Enhancing Medication Recommendation for Rare Disease Patients

Zihao Zhao, Yi Jing, Fuli Feng, Jiancan Wu, Chongming Gao, Xiangnan He  
*University of Science and Technology of China*

SIGIR 2024 · Washington, DC

Grupo 7

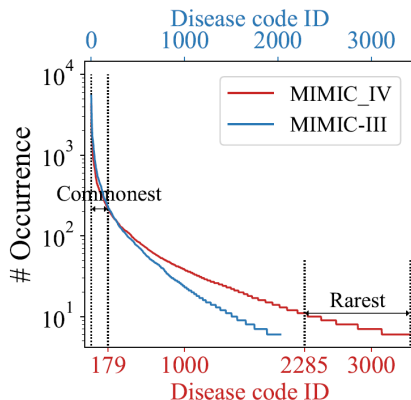
*Joaquín De Ferrari, Nicolás Schiaffino, Gabriel Venegas*

Se estima que al menos **1.500.000** de personas son afectadas por errores de medicación *prevenibles* **al año**.

Los sistemas de recomendación de medicamentos son prometedores, pero sufren de problemas de **equidad (fairness)**.

# Causa principal: El Sesgo en los Datos

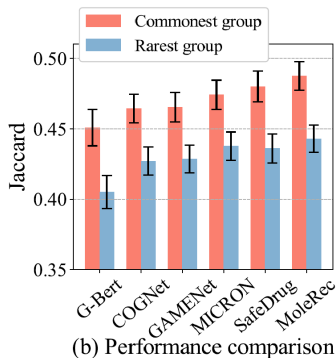
- Pocas enfermedades son comunes.
- El resto son consideradas **raras**.
- Datos de entrenamiento insuficientes para los casos raros.



(a) Clinical code distribution

## Recomendaciones Injustas

Las recomendaciones son más precisas para pacientes con enfermedades comunes que para aquellos con condiciones raras. Esta disparidad en la precisión socava la fiabilidad y utilidad de estos sistemas de recomendación.



Desarrollar un modelo que sea preciso en la recomendación de medicamentos, haciendo especial énfasis en la equidad. Es decir, obtener una alta precisión en el grupo de pacientes con condiciones **raras**.

Modelar y recomendar para enfermedades raras es una tarea difícil. En base a lo anterior el paper resume 3 desafíos centrales:

- **Escasez de datos de calidad:** La baja prevalencia de enfermedades raras impide discernir con precisión los patrones de recomendación de medicamentos.
- **Asociaciones intrincadas:** Las enfermedades raras a menudo presentan combinaciones complejas de síntomas y eventos clínicos, lo que confunde la selección de medicamentos.
- **Necesidades de medicación a medida:** La fisiopatología atípica de las enfermedades raras a menudo necesita intervenciones farmacológicas dirigidas con mayor precisión.

3 Áreas clave:

- 1 **Recomendación de Fármacos**
- 2 **Fairness en Recomendación**
- 3 **Pre-entrenamiento en Recomendación**



# 1: Recomendación de Fármacos

## a) Métodos Longitudinales

- Utilizan el historial médico de los pacientes.
- Menos adecuados para pacientes con una sola visita.
- *Ejemplos: GAMENet, SafeDrug, RETAIN, COGNet, MICRON y MoleREC.*

## b) Métodos basados en instancias

- Se centran en la condición médica actual del paciente.
- No capturan adecuadamente información específica del paciente, lo que puede reducir la precisión.
- *Ejemplo: LEAP (Zhang et al.).*

## 2: Fairness en Recomendación

### a) Métodos de preprocesamiento

- Modifican los datos de entrenamiento para eliminar sesgos.
- *Ejemplo: Re-muestreo para balancear datos (Ekstrand et al.).*

### b) Métodos en procesamiento

- Añaden regularización o aprendizaje adversario durante el entrenamiento.
- *Ejemplo: Métricas de fairness como regularización (Yao et al.).*

### c) Métodos de postprocesamiento

- Re-ordenan la lista de resultados para mejorar la equidad.
- *Ejemplo: Ranking con restricciones de intervalo (Geyik et al.).*

#### Mejorar representaciones latentes

- Se usan para enriquecer las representaciones de ítems o usuarios.
- Aprovechan datos a gran escala (no necesariamente de la tarea final) para aprender patrones generales.
- *Ejemplos: U-BERT (Qiu et al.), G-BERT (Shang et al.).*

# Formulación del problema:

- Electronic Health Record (EHR) denoted as  $\mathcal{R} = \{\mathcal{V}^{(1)}, \dots, \mathcal{V}^{(j)}, \dots, \mathcal{V}^{(N)}\}$
- Set of Diseases denoted as  $\mathcal{D} = \{d_1, d_2, \dots, d_{|\mathcal{D}|}\}$
- Set of Procedures denoted as  $\mathcal{P} = \{p_1, p_2, \dots, p_{|\mathcal{P}|}\}$
- Set of Medications denoted as  $\mathcal{M} = \{m_1, m_2, \dots, m_{|\mathcal{M}|}\}$

## Input

$$\mathcal{V}^{(j)} = \{d^{(j)}, p^{(j)}, m^{(j)}\}$$

donde:

- $d^{(j)}$  : Subset de enfermedades ordenadas por prioridad o importancia según EHR dataset.
- $p^{(j)}$  : Subset de procedures ordenadas por prioridad igualmente.
- $m^{(j)}$  : Multi-hot Encoded vector de medicamentos. El 1 indica que el medicamento fue prescrito y 0 indica ausencia.

# Formulación del problema:

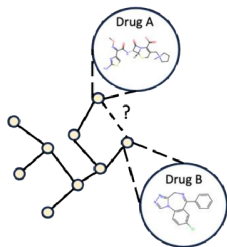


Figura: (Xu et al., 2024)

## Drug-Drug Interaction (DDI) graph

Una matriz simétrica binaria  $A \in \{0, 1\}^{|\mathcal{M}| \times |\mathcal{M}|}$  en donde cada elemento  $A_{uv}$  es 1 sólo si existe una interacción dañina entre la  $u$ -ésima y la  $v$ -ésima droga.

## Problema de Recomendación de medicamentos

Dada una secuencia de diagnóstico de enfermedades  $d^{(k)}$  y una secuencia de tratamientos  $p^{(k)}$  del paciente  $k$ , junto con con el DDI graph, el objetivo es:

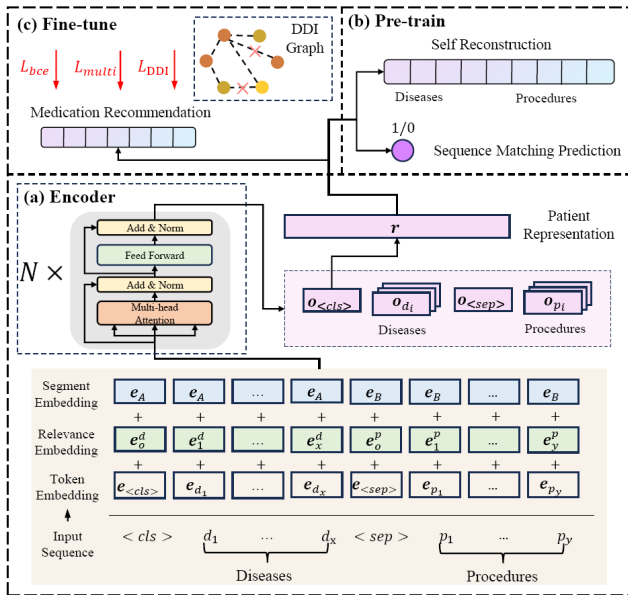
- Generar una apropiada combinación de medicamentos  $\hat{y}$  para el paciente que **maximice** la precisión de la recomendación *mientras* se **minimiza** el riesgo de interacciones adversas en los medicamentos.

## RAREMed

(Robust and Accurate REcommendations for Medication)

Un modelo diseñado para la **Precisión** y la **Equidad**.

# Arquitectura General de RAREMed

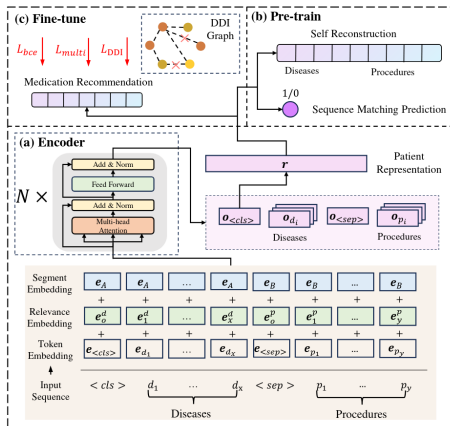




# Paso 1: El Encoder (1/3) - Entrada Unificada

Se combinan diagnósticos  $d^{(r)}$  y procedimientos  $p^{(r)}$  en una sola secuencia para capturar interacciones complejas.

$$\text{input}_r = [\text{CLS}] \oplus d^{(r)} \oplus [\text{SEP}] \oplus p^{(r)}$$

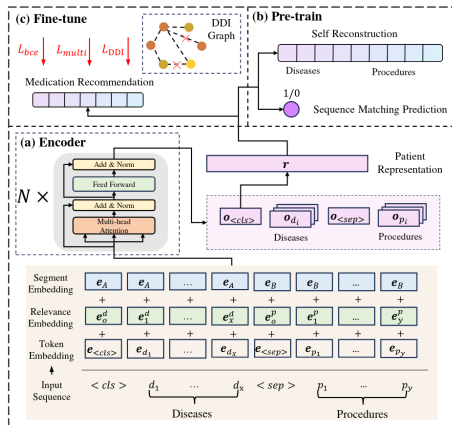


# Paso 1: El Encoder (2/3) - Embeddings Aumentados

Se enriquece la entrada con 3 capas de embeddings:

- **Token** ( $E_{tok}$ ): ¿Qué código es?
- **Relevance** ( $E_{rel}$ ): ¿Qué tan importante es?
- **Segment** ( $E_{seg}$ ): ¿Es diagnóstico o procedimiento?

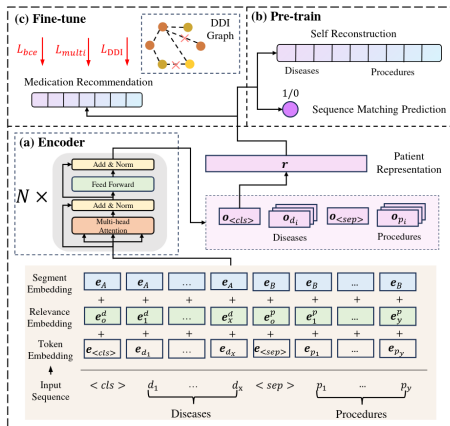
**Resultado:** Una representación mucho más completa del paciente



# Paso 1: El Encoder (3/3) - Representación Final

La secuencia de embeddings pasa por un **Transformer Encoder**. La representación final del paciente ( $r$ ) se toma del token '[CLS]'.

$$r = \text{Encoder}(E_{tok} + E_{seg} + E_{rel})[0]$$



## Idea Principal

Se entrena al modelo en tareas auxiliares para que **entienda la estructura** de los datos clínicos antes de la tarea principal de recomendación.

Se utilizan 2 tareas:

- 1 Sequence Matching Prediction (SMP)
- 2 Self Reconstruction (SR)

# Sequence Matching Prediction (SMP)

- **Objetivo:** Discernir si las secuencias de enfermedades y procedimientos pertenecen al mismo paciente.
- **Proceso:** Para cada par de secuencias de un paciente, se crea un ejemplo negativo reemplazando aleatoriamente una de las secuencias con la de otro paciente.

## Ecuación de Pérdida (BCE Loss)

$$L_{SMP} = -\log(\hat{y}_i) - \log(1 - \hat{y}_j)$$

Donde  $\hat{y}_i$  es la probabilidad predicha para el par correcto y  $\hat{y}_j$  para el par incorrecto.

# Self Reconstruction (SR)

- **Objetivo:** Reconstruir la secuencia de códigos clínicos de entrada a partir de la representación del paciente  $r$ .
- **Proceso:** Clasificación multietiqueta, prediciendo enfermedades y procedimientos para cada paciente.

## Ecuación de Reconstrucción

$$L_{SR} = - \sum_{j=1}^{|\mathcal{D}|+|\mathcal{P}|} [c_j \log(\hat{c}_j) + (1 - c_j) \log(1 - \hat{c}_j)]$$

Donde  $c$  es el vector *ground truth* de los códigos de entrada y  $\hat{c}$  es el vector de probabilidades reconstruidas.

- Se adapta el modelo preentrenado para la tarea de **recomendar medicamentos**.
- Esto se plantea como un problema de **clasificación multietiqueta**, donde se predice la probabilidad  $\hat{o}$  para cada medicamento:

## Capa de Clasificación Multilabel

$$\hat{o} = \sigma(W_3 r + b_3)$$

$\hat{o}$  es el vector de probabilidades de cada medicamento y  $r$  es la representación del paciente obtenida por el encoder.

## Loss para Accuracy (1/2)

Se trata cada medicamento como una tarea de clasificación binaria independiente.

### Binary Cross-Entropy Loss

$$L_{bce} = - \sum_{i=1}^{|\mathcal{M}|} [m_i \log(\hat{o}_i) + (1 - m_i) \log(1 - \hat{o}_i)]$$



Asegura que los medicamentos correctos tengan una puntuación mayor que los incorrectos.

### Ecuación (7): Multi-label Margin Loss

$$L_{multi} = \sum_{\substack{i:m_i=1 \\ j:m_j=0}} \frac{\text{máx}(0, 1 - (\hat{o}_i - \hat{o}_j))}{|\mathcal{M}|}$$

Penaliza si el modelo recomienda dos medicamentos con una interacción adversa conocida.

Ecuación (8): Loss de Interacción Fármaco-Fármaco

$$L_{ddi} = \sum_{i=1}^{|\mathcal{M}|} \sum_{j=1}^{|\mathcal{M}|} A_{ij} \cdot \hat{o}_i \cdot \hat{o}_j$$

- $A_{ij}$  es 1 si los fármacos  $i$  y  $j$  tienen una interacción peligrosa.
- La pérdida es alta si las probabilidades  $\hat{o}_i$  y  $\hat{o}_j$  de dos fármacos que interactúan son altas.

# Pérdida Total: El Balance Final

La función de pérdida final combina precisión y seguridad.

Ecuación (9):  $L$

$$L = (1 - \beta) \underbrace{((1 - \alpha)L_{bce} + \alpha L_{multi})}_{\text{Accuracy}} + \beta \underbrace{L_{ddi}}_{\text{Safety}}$$

- $\alpha$  y  $\beta$  son hiperparámetros que controlan la importancia de cada componente de la pérdida.

La recomendación final se realiza imponiendo un umbral sobre cada la probabilidad de cada medicamento.

## Predicción Final

$$\hat{y} = \{i \mid \hat{o}_i > 0,5, 1 \leq i \leq |\mathcal{M}|\}$$

- $\hat{y}$  es el **conjunto final** de medicamentos que se recomienda al paciente.

- 1 ¿Cómo se compara el **rendimiento** del RAREMed propuesto con el de los métodos de recomendación de medicamentos existentes?
- 2 ¿Mitiga RAREMed eficazmente la **injusticia** (unfairness) en las recomendaciones de medicamentos?

# Métrica de Precisión: Coeficiente de Jaccard

Mide la similitud entre los medicamentos recetados y los recomendados.

## Fórmula de Jaccard

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

- $A$ : Es el conjunto de medicamentos **reales** (recetados al paciente).
- $B$ : Es el conjunto de medicamentos **predichos** por el modelo.

## Precision-Recall Area Under Curve.

Resume el rendimiento del modelo, especialmente en datos desbalanceados.

### Fórmula de PRAUC

$$\text{PRAUC} = \sum_{k=1}^{|\mathcal{M}|} P_k (R_k - R_{k-1})$$

- $P_k$ : Es la **Precisión** en el punto de corte  $k$ . (De los recomendados, ¿cuántos son correctos?)
- $R_k$ : Es el **Recall** (exhaustividad) en el punto de corte  $k$ .

# Resultados: Rendimiento en Visita Única (MIMIC-III)

Método	Jaccard ↑	F1 ↑	DDI ↓	#MED
LR	0.4933	0.6519	0.0786	16.85
LEAP	0.4526	0.6154	0.0722	19.04
GAMENet	0.5210	0.6762	0.0781	19.78
SafeDrug	0.5255	0.6804	0.0688	20.84
COGNet	0.5109	0.6676	0.0737	25.05
MICRON	0.5119	0.6676	0.0610	20.94
MoleRec	0.5303	0.6844	0.0692	21.30
<b>RAREMed</b>	<b>0.5414</b>	<b>0.6942</b>	<b>0.0529</b>	19.66



# Resultados: Rendimiento en Visita Única (MIMIC-IV)

Método	Jaccard ↑	PRAUC ↑	F1 ↑	DDI ↓	#MED
LR	0.4150	0.6782	0.5650	0.0736	9.85
LEAP	0.3907	0.5540	0.5437	0.0548	12.66
GAMENet	0.4401	0.6833	0.5933	0.0718	12.89
SafeDrug	0.4560	0.6858	0.6098	0.0689	14.07
COGNet	0.4313	0.6112	0.5850	0.0866	15.46
MICRON	0.4495	0.4353	0.6033	0.0502	14.50
MoleRec	0.4502	0.6867	0.6040	0.0699	14.05
<b>RAREMed</b>	<b>0.4625</b>	<b>0.7008</b>	<b>0.6158</b>	<b>0.0508</b>	12.51

# Resultados: Rendimiento en Multi-Visita (MIMIC-III)

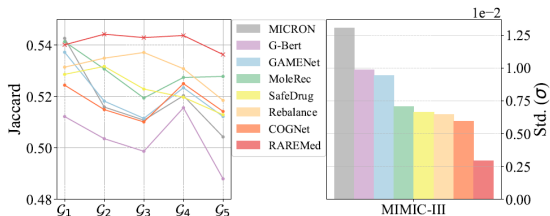
Método	Jaccard $\uparrow$	F1 $\uparrow$	DDI $\downarrow$	#MED
LR	0.4933	0.6519	0.0786	16.85
LEAP	0.4526	0.6154	0.0722	19.04
RETAIN	0.4922	0.6517	0.0792	24.76
G-Bert	0.5037	0.6617	0.0832	24.46
GAMENet	0.5205	0.6754	0.0774	19.52
SafeDrug	0.5232	0.6781	0.0698	20.67
COGNet	0.5177	0.6715	0.0757	24.95
MICRON	0.5188	0.6735	0.0658	19.51
MoleRec	0.5293	0.6838	0.0698	21.27
<b>RAREMed</b>	<b>0.5414</b>	<b>0.6942</b>	<b>0.0529</b>	19.66

# Resultados: Rendimiento en Multi-Visita (MIMIC-IV)

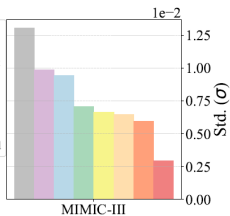
Método	Jaccard $\uparrow$	PRAUC $\uparrow$	F1 $\uparrow$	DDI $\downarrow$	#MED
LR	0.4150	0.6782	0.5650	0.0736	9.85
LEAP	0.3907	0.5540	0.5437	0.0548	12.66
RETAIN	0.4113	0.6543	0.5674	0.0839	16.95
G-Bert	0.4292	0.6746	0.5822	0.0794	16.19
GAMENet	0.4462	0.6923	0.5979	0.0807	12.60
SafeDrug	0.4566	0.6867	0.6080	0.0644	14.00
COGNet	0.4518	0.6359	0.6026	0.0846	15.70
MICRON	0.4555	0.5350	0.6081	0.0528	13.24
MoleRec	0.4576	0.6933	0.6097	0.0683	14.19
<b>RAREMed</b>	<b>0.4625</b>	<b>0.7008</b>	<b>0.6158</b>	<b>0.0508</b>	12.51

# Resultados: Equidad (Fairness)

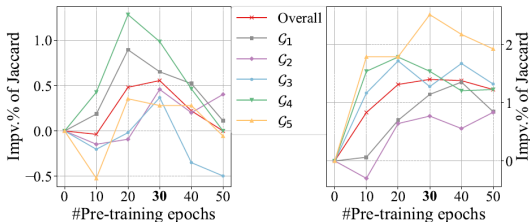
**RAREMed es significativamente más justo.**



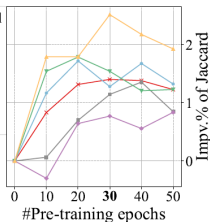
(a) MIMIC-III



(b) Std. of Jaccard



(c) MIMIC-III



(d) MIMIC-IV

- El rendimiento de otros modelos **cae drásticamente** en grupos raros ( $G_4$ ,  $G_5$ ).
- **RAREMed** mantiene un rendimiento **estable y alto**.

- El sesgo contra enfermedades raras es un problema **crítico** y real.
- **RAREMed** lo mitiga con un **encoder expresivo** y un **pre-entrenamiento** inteligente.
- **Resultado:** Recomendaciones más **precisas, seguras** y, fundamentalmente, más **JUSTAS**.

- ① Edward Choi, Mohammad Taha Bahadori, Jimeng Sun, Joshua Kulas, Andy Schuetz, and Walter Stewart. 2016. Retain: An interpretable predictive model for healthcare using reverse time attention mechanism. *Advances in neural information processing systems* 29 (2016).
- ② Michael D Ekstrand, Mucun Tian, Ion Madrazo Azpiazu, Jennifer D Ekstrand, Ogheemaro Anuyah, David McNeill, and Maria Soledad Pera. 2018. All the cool kids, how do they fit in?: Popularity and demographic biases in recommender evaluation and effectiveness. In *Conference on fairness, accountability and transparency*. PMLR, 172-186.
- ③ Sahin Cem Geyik, Stuart Ambler, and Krishnaram Kenthapadi. 2019. Fairness-aware ranking in search & recommendation systems with application to linkedin. talent search. In *Proceedings of the 25th acm sigkdd international conference on knowledge discovery & data mining*. 2221-2231.
- ④ Junyuan Shang, Tengfei Ma, Cao Xiao, and Jimeng Sun. 2019. Pre-training of graph augmented transformers for medication recommendation. In *28th International Joint Conference on Artificial Intelligence, IJCAI 2019*. International Joint Conferences on Artificial Intelligence, 5953-5959.

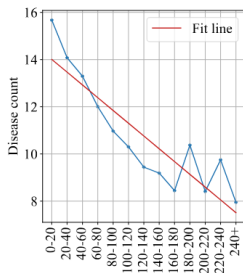
- 5 Junyuan Shang, Cao Xiao, Tengfei Ma, Hongyan Li, and Jimeng Sun. 2019. Gamenet: Graph augmented memory networks for recommending medication combination. In *proceedings of the AAAI Conference on Artificial Intelligence, Vol. 33*. 1126-1133.
- 6 Rui Wu, Zhaopeng Qiu, Jiacheng Jiang, Guilin Qi, and Xian Wu. 2022. Conditional generation net for medication recommendation. In *Proceedings of the ACM Web Conference 2022*. 935-945.
- 7 Chaoqi Yang, Cao Xiao, Lucas Glass, and Jimeng Sun. 2021. Change Matters: Medication Change Prediction with Recurrent Residual Networks. In *30th International Joint Conference on Artificial Intelligence, IJCAI 2021*. International Joint Conferences on Artificial Intelligence, 3728-3734.
- 8 Chaoqi Yang, Cao Xiao, Fenglong Ma, Lucas Glass, and Jimeng Sun. 2021. SafeDrug: Dual Molecular Graph Encoders for Recommending Effective and Safe Drug Combinations. In *30th International Joint Conference on Artificial Intelligence, IJCAI 2021*. International Joint Conferences on Artificial Intelligence, 3735-3741.

- 9 Nianzu Yang, Kaipeng Zeng, Qitian Wu, and Junchi Yan. 2023. Molerec: Combinatorial drug recommendation with substructure-aware molecular representation learning. In *Proceedings of the ACM Web Conference 2023*. 4075-4085.
- 10 Sirui Yao and Bert Huang. 2017. Beyond parity: Fairness objectives for collaborative filtering. *Advances in neural information processing systems* 30 (2017).
- 11 Yutao Zhang, Robert Chen, Jie Tang, Walter F Stewart, and Jimeng Sun. 2017. LEAP: learning to prescribe effective and safe treatment combinations for multimorbidity. In *proceedings of the 23rd ACM SIGKDD international conference on knowledge Discovery and data Mining*. 1315-1324.

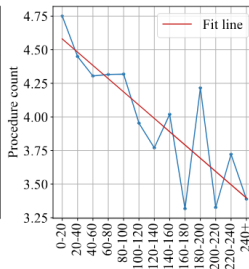


# ¿Preguntas?

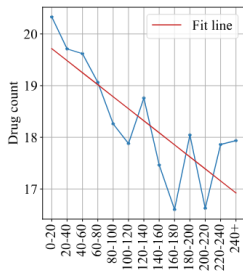
# Anexo 1 - Perfil de Pacientes con Enfermedades Raras



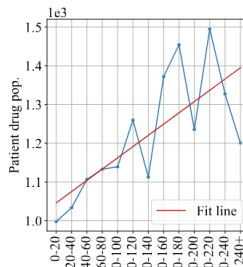
(a) Disease count w.r.t. disease pop.



(b) Procedure count w.r.t. disease pop.



(c) Drug count w.r.t. disease pop.



(d) Drug popularity w.r.t. disease pop.

Para abordar estos desafíos, se propone un nuevo modelo llamado **RAREMed** (Robust and Accurate Recommendations for Medication).

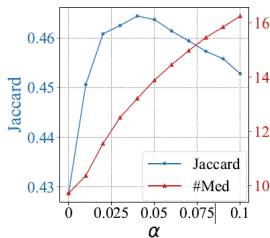
- **Estrategia Clave:** Utiliza un paradigma de **pre-entrenamiento y ajuste fino** para mejorar la precisión en enfermedades raras.
- **Encoder Unificado:** Emplea un Transformer que procesa los códigos de enfermedades y procedimientos en una secuencia unificada para capturar sus complejas relaciones.
- **Tareas de Pre-entrenamiento:** Introduce dos tareas auto-supervisadas (Sequence Matching y Self Reconstruction) para aprender las necesidades de medicación especializadas y las interrelaciones entre los códigos clínicos.

## Anexo 3 - Métodos a Comparar

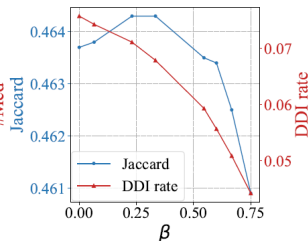
Método	Descripción
RETAIN	Utiliza atención neural de dos niveles para predecir la condición futura.
GAMENet	Usa redes de memoria y GCN para codificar EHR y el grafo de interacciones DDI.
SafeDrug	Aprovecha el grafo molecular de fármacos y el grafo DDI para garantizar la seguridad.
COGNet	Utiliza un novedoso mecanismo de "copiar o predecir".
MICRON	Modelo de aprendizaje residual recurrente enfocado en el cambio de medicación.
MoleRec	Modela la dependencia entre la salud del paciente y las subestructuras moleculares.
LEAP	Usa un modelo LSTM para generar secuencias de medicación.
G-Bert	Integra representaciones de GNN en encoders Transformer.

- 1 ¿Cómo influyen los **componentes individuales** de RAREMed en su rendimiento con respecto a la precisión y la equidad?
- 2 ¿Cuáles son los **factores influyentes**, como el gráfico de interacciones fármaco-fármaco (DDI), que impactan notablemente el rendimiento de RAREMed?

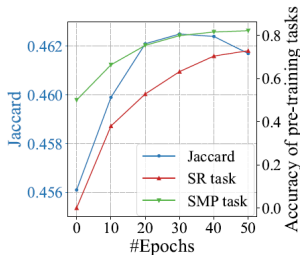
# Anexo 4 - Gráficos RQ4



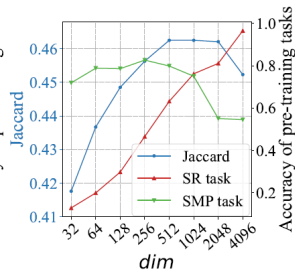
(a) Weight of  $L_{multi}$



(b) Weight of  $L_{ddi}$



(c) Number of pre-training epochs



(d) Embedding dimension