

¹Hospital Público de Monforte, Monforte de Lemos, España. ²Hospital Universitario Lucus Augusti, Lugo, España. ³Hospital DA Costa, Burela, España.

Resumen

Métodos: Se extrajeron al azar 10 informes de alta de pacientes en formato pdf. Se editaron los archivos para eliminar todos los datos que pudieran identificar al paciente (nombre, teléfono, dirección, CIP y número de la seguridad social). Se elaboró una pregunta (fig.) que se realizó a un LLM (*large language model*, modelo grande de lenguaje) comercial no entrenado previamente para la tarea (ChatGPT, v4.0) y se le administraron los informes en pdf para la extracción del número de historia clínica, sexo, edad, presencia/ausencia de ocho antecedentes médicos preespecificados, tratamiento antibiótico recibido durante el ingreso, y destino al alta. Los informes se introdujeron primero juntos en un fichero y después por separado. Un digitador de datos humano realizó la misma tarea y se cuantificó la proporción de acuerdo observado y el coeficiente kappa de Cohen.

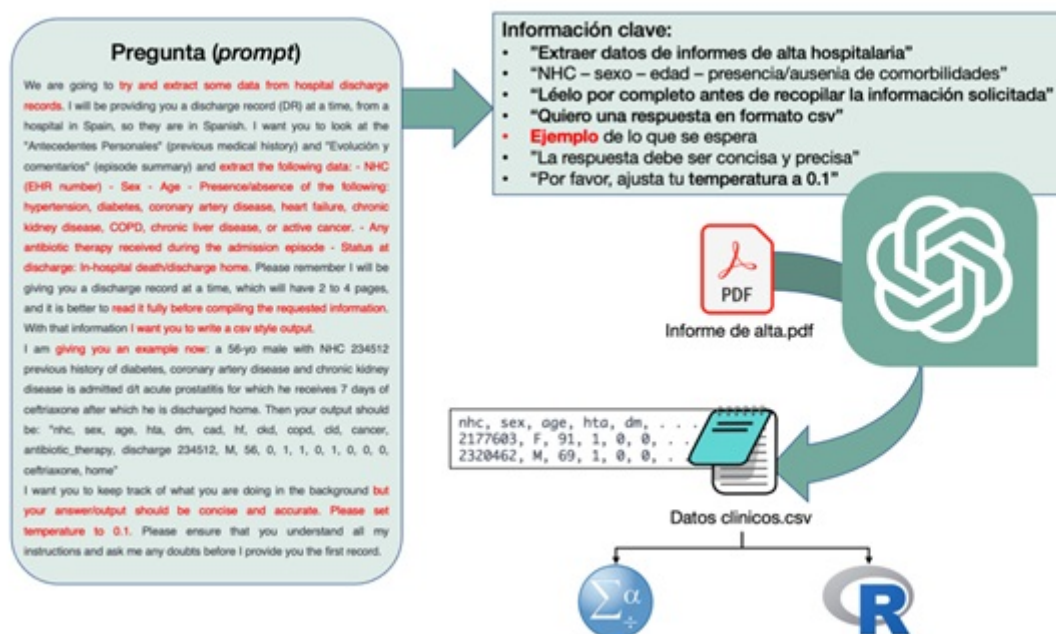
Resultados: En total se solicitó la extracción de 13 datos de cada uno de los 10 informes de alta (tabla). Mediante el fichero con los 10 informes conjuntos el LLM fue incapaz de realizar la tarea. Al introducir los informes uno a uno, el acuerdo total observado fue del 96,9% (IC95 95,4-98,4%), con un coeficiente kappa global de 0,8574. El acuerdo fue perfecto para número de historia, edad, y sexo, así como para los antibióticos recibidos y el destino al alta. Para los antecedentes el acuerdo total observado fue del 95,0% (IC95 92,6-97,4%), el acuerdo positivo observado del 87,5%, el acuerdo negativo del 96,9% y el coeficiente kappa de 0,8438. Las discordancias encontradas (4 datos) se muestran en la tabla.

Variable	NHC	Sexo	Edad	HTA	DM	CISQ	ICC	ERC	EPOC	EHC	Cáncer	Antibiótico	Alta
Acuerdo humano:IA	100%	100%	100%	100%	100%	100%	100%	90%	100%	80%	90%	100%	100%

Desacuerdos observados

1. ERC Antecedente de fracaso renal por aminoglucósidos, la IA interpretó ERC = Sí
2. EHC Antecedente de colelitiasis con colangitis de repetición, la IA interpretó EHC = Sí
3. EHC Antecedente de esteatosis hepática, la IA interpretó EHC = No
4. Cáncer Neoplasia vesical pendiente de resección transuretral, la IA interpretó Cáncer = No

NHC, número de historia clínica; HTA, hipertensión arterial; DM, diabetes mellitus; CISQ, cardiopatía isquémica; ICC, insuficiencia cardíaca; ERC, enfermedad renal crónica; EPOC, enfermedad pulmonar obstructiva crónica; EHC: enfermedad hepática crónica.



Discusión: La historia clínica electrónica a menudo no está diseñada para la extracción de datos para la investigación, como tampoco los informes que esta produce. Si bien los parámetros numéricos (constantes, resultados analíticos) pueden obtenerse desde los sistemas informáticos de laboratorios, antecedentes, evolución y recomendaciones al alta deben recogerse manualmente, a menudo desde campos tipo texto libre. Presentamos un ejemplo de extracción mediante Inteligencia Artificial de esta información a partir de informes de alta corrientes con un acuerdo muy bueno o excelente entre el LLM y un digitador humano. El experimento produjo fallos al introducir varios informes en un mismo fichero, pero no era un modelo específico ni había sido preentrenado para la tarea.

Conclusiones: En nuestro estudio, la concordancia de un LLM con un digitador humano fue excelente. Los LLM son una opción prometedora dentro de la inteligencia artificial para la extracción de datos clínicos a partir de documentos no específicos para ello.