

TRABAJO FINAL DE MASTER UNIVERSITARIO EN TELEMEDICINA

Estudio y evaluación del sistema recomendador basado en la captura y síntesis de la historia clínica electrónica a través del uso de lenguaje natural y ontologías sanitarias perteneciente al proyecto HOPE.

Alumno: Nesib Nicolás Achkar Tuglaman

**Tutor: Dr. Carlos Luis Sánchez Bocanegra
Dr. Luis Fernández Luque**

Junio 2018

Agradecimiento:

Este trabajo no habría sido posible si no fuera por unas personas que han sido parte de ello de forma directa e indirecta.

Para Cristina, por su apoyo y su paciencia, por los “juepintxos” que perdimos.

Para mis padres, quienes me han inculcado no parar de aprender nunca.

Para mis socios en ITH; Luis, Pedro y Gonzalo, que sin ellos, ni este master ni mi camino profesional habría sido igual.

Y como no, por mis compañeros de HOPE, especialmente Karla y por mi tutor, Carlos, por todo lo que me han enseñado en los últimos meses y espero que siga aprendiendo con ellos.

Contenido

Contenido	2
1. Introducción	3
a. Situación del cáncer en la evidencia científica:.....	3
b. Gran cantidad de información accesible.....	4
c. Fuentes de información científica fiable en medicina:.....	4
d. Ontología sanitaria:.....	5
e. Procesamiento del lenguaje natural (PLN).....	6
f. Términos semánticos:.....	6
g. Necesidad del sistema HOPE.....	7
h. Descripción del sistema:.....	7
i. Primer fase: comunicación inicial (Hope).....	7
ii. Segunda fase: Procesamiento Natural del Lenguaje (NLP).....	8
iii. Tercera fase: Ontologías sanitarias.....	8
iv. Cuarta fase: Conexión con fuentes de información sanitaria fiable.....	8
v. Quinta fase: Formateado.....	8
vi. Sexta fase: Incorporación en la Historia Clínica Electrónica.....	8
i. Grupo de investigación:.....	9
2. Objetivos	10
i. Objetivos primarios.....	10
ii. Objetivos secundarios.....	10
3. Metodología	11
a. Selección e introducción de los datos de casos clínicos.....	11
i. Descripción de la interfaz:.....	11
ii. Selección de casos:.....	11
iii. Criterios de exclusión:.....	12
b. Evaluación externa de precisión del sistema en PubMed.....	13
4. Resultados	15
a. Descripción de resultados:.....	15
b. Evaluación interna del sistema:.....	16
c. Evaluación externa de concordancia en Pubmed:.....	17
5. Discusión:	20
a. Limitaciones identificadas.....	20
b. Sugerencias de mejora.....	21
6. Conclusión	22
7. Bibliografía	23
8. Anexos:	0

Estudio y evaluación del sistema recomendador basado en la captura y síntesis de la historia clínica electrónica a través del uso de lenguaje natural y ontologías sanitarias perteneciente al proyecto HOPE.

1. Introducción

La práctica médica actual tiene dos características que le distinguen de la medicina del siglo XX, estar centrada en el paciente y estar basada en la evidencia (MBE).

La MBE empezó como un movimiento de mejora de la calidad clínica. Para que la MBE pueda ser llevada a la práctica se requiere integrar la experiencia clínica individual con la de otros profesionales de la salud. cuando se toma una decisión terapéutica, tomando en cuenta los valores y las preferencias de los pacientes a lo que se conoce como la tríada de la medicina basada en la evidencia(1).

El acceso a los datos y la información más actualizados a nivel local e internacional permite ampliar el conocimiento médico y con ello proporciona al profesional de la salud más opciones para la para la correcta toma de decisiones. La MBE se aplica desde el diagnóstico de un padecimiento hasta el tratamiento del paciente. La MBE tiene el potencial de cambiar exponencialmente el pronóstico y la calidad de vida de los pacientes.

a. Situación del cáncer en la evidencia científica:

El cáncer es un padecimiento con una incidencia en crecimiento exponencial a nivel global. En la cual es fundamental el realizar un diagnóstico oportuno para poder proporcionar un manejo en estadios tempranos. Estos factores son fundamentales para el pronóstico. .Constantemente a nivel mundial se están realizando investigaciones para encontrar alternativas de manejo a las diversas neoplasias. .La eficiencia de los tratamiento depende en el estadio clínico en el que se encuentra en el momento del inicio y al presentar un estadio más avanzado se reduce exponencialmente las opciones terapéuticas llegado el caso a la limitación del esfuerzo terapéutico y adoptar estrategias paliativas.

b. Gran cantidad de información accesible.

Todos los días se publican numerosos artículos científicos a nivel internacional, (pubmed acumula a junio 2018: 28.567.794 de artículos)(2), siendo temas relacionados con el cáncer un 16% hasta 2016 (3) y están activos gran número de estudios sobre el cáncer que van creando conocimiento que modifica protocolos médicos que pueden ser muy beneficiosos para nuestros pacientes, pero la capacidad individual para poder estar continuamente actualizado, es muy limitada.

En la práctica diaria, la buena comunicación con los pacientes es primordial, y darle acceso a información médica verificada, permite un mejor empoderamiento que retorna a ellos con mejor salud y más confianza en el sistema sanitario(4).

Los nuevos avances tecnológicos al alcance de los médicos, permiten una mejor recogida de datos, especialmente en la redacción de la historia clínica electrónica, pero esta información se redacta en texto plano que no permite su correcta explotación como por ejemplo la utilización para su indexación o mapeo.

c. Fuentes de información científica fiable en medicina:

- i. **Pubmed:** Es una de las bases de datos bibliográfica y biomédica, más grandes y con más autoridad en el mundo(5). Además existe evidencia que los usuarios experimentados de Pubmed acceden más rápido y a documentos más relevantes y más rápidos que usuarios con menos experiencia.(6)
- ii. **ClinicalTrials (7)** por otro lado, es un registro de ensayos clínicos activos y accesible en línea desde su portal, y se cuenta con un sistema de información con base en la web en USA, en el cual se encuentran 247,748 estudios en 50 Estados de 202 Países.

- iii. **MedlinePlus:** el sitio web del Instituto Nacional de la Salud de EEUU para pacientes y familiares. Producida por la Biblioteca Nacional de Medicina de los Estados Unidos, MedlinePlus da información dirigida a pacientes enfermedades, afecciones y bienestar en un lenguaje fácil de leer. Es fuente de información confiable y actualizada en todo momento.(8)

Nuestro objetivo es desarrollar un sistema que identifique y extraiga de la historia clínica electrónica y de forma automatizada, los datos clínicos de los pacientes en base a ontologías sanitarias.

Esta información se vincule con Pubmed, con Medline Plus y con clinicalTrails para proporcionar al médico una selección de información relativa a los últimos avances científicos personalizados a su paciente, proporcionar información al médico acerca de posibles tratamientos en base a artículos científicos. junto a los posible ensayo clínicos en los que pueda participar y la información mas relevante respecto a su enfermedad de una fuente fiable.

Nuestro reto consiste en permitir que un lenguaje natural (texto plano) con el que los médicos escriben en la historia clínica electrónica, se pueda transformar en texto comprensible informáticamente y que permite su búsqueda en bases de datos. Esto se consigue a través de un procesamiento automatizado que permite identificar los términos de interés en el texto libre, convertirlo en términos semánticos y después en ontología sanitaria específica de cada base de datos, Hecho que permite realizar una búsqueda con este tipo de descriptores de forma más precisa.

d. **Ontología sanitaria:**

Una ontología es un recurso constituido por un conjunto de conceptos organizados por sus relaciones(9). Los conceptos y las relaciones incluidas deben describir el conocimiento consensuado y aceptado de un área o dominio de interés para ser utilizado tanto por los humanos como por los sistemas de información (SI) (9). Este recurso permite que los SI trabajen automáticamente a partir del significado de los signos lingüísticos (términos) y sus combinaciones.

Las ontologías son un recurso que permite trabajar informáticamente con la conceptualización del significado y evitar la limitación impuesta por los términos normalizados.(10)

Uno de los repositorios de ontologías médicas bioontology(14)

Como ejemplos de Ontologías están:

- La Clasificación Internacional de Enfermedades en su decima revisión, Modificación Clínica (CIE-10-MC)(11)
- Los Medical Subject Headings (MeSH)(12) Éste es el vocabulario controlado de la National Library Medicine (NLM) para indexar/referenciar. Se realiza así porque se realiza una búsqueda con este tipo de descriptores de forma más precisa.
- Systematized Nomenclature of Medicine, Clinical Terms (SNOMED CT) (13)

e. Procesamiento del lenguaje natural (PLN)

Es el campo que combina las tecnologías de la ciencia computacional (como la inteligencia artificial, el aprendizaje automático o la inferencia estadística) con la lingüística aplicada, con el objetivo de hacer posible la comprensión y el procesamiento automatizado de información expresada en lenguaje humano para determinadas tareas, como la traducción automática, los sistemas de diálogo interactivos, el análisis de opiniones, etc.(15)

Uno de los componentes de PLN es el análisis semántico; La extracción del significado de la frase, y la resolución de ambigüedades léxicas y estructurales, que es lo que utilizaremos en HOPE.

f. Términos semánticos:

En el campo de aplicación de las tecnologías de la información, la “semántica” tiene un significado más concreto y se refiere al conjunto de tecnologías que proporcionan una manera de representar el conocimiento de un dominio en términos entendibles tanto por humanos como por máquinas(16).

g. Necesidad del sistema HOPE

La búsqueda de información desde fuentes bibliográficas científicas debe llevarse siguiendo una metodología, que permita establecer un proceso de búsqueda y recuperación de información válido, actual, personalizado, eficaz, eficiente y en tiempo aceptable.

Esta metodología se fundamenta en varios pilares siendo uno de ellos la información clínica del paciente la cual es tomada de la historia clínica electrónica. Se identifica y se toma un conjunto de variables como sexo, edad, diagnóstico (contexto definido). Esta información se analiza y en base a esta se obtienen términos semánticos. Los cuales son utilizados como un vínculo para realizar la búsqueda de recomendaciones en PubMed, MedlinePlus y ClinicalTrials.

Por ejemplo, en la HCE se describe un "cáncer de colon", si se realiza una búsqueda en PubMed a través de las palabras claves "cáncer de colon". El resultado obtenido son más de 124000 posibles artículos. Ahora bien si seleccionamos el término MeSH "Colonic Neoplasms" [Mesh], el número de resultados baja a casi la mitad. Se ha desarrollado un sistema en el que cata de forma automatizada, texto en español, convierte este texto en términos semánticos y lo busca automáticamente en diferentes bases de datos.

La Implementación de este sistema se podría utilizar como una herramienta de actualización constante del profesional médico en base a medicina basada en evidencia y en base a artículos científicos y ensayos clínicos.

h. Descripción del sistema:

Técnicamente el sistema HOPE consiste en 6 fases:

i. Primer fase: comunicación inicial (Hope)

La primer fase del método, establece la comunicación inicial entre la historia clínica electrónica y HOPE. En esta primera comunicación, desde la historia clínica se provee a HOPE de los datos vertidos por el profesional médico en el historial del paciente.

ii. Segunda fase: Procesamiento Natural del Lenguaje (NLP).

Este procesamiento hace posible la comprensión y el procesamiento asistido por algoritmo de información expresada en lenguaje humano combinando la ciencia computacional aplicada a la lingüística aplicada.

iii. Tercera fase: Ontologías sanitarias.

En esta misma fase, los términos estructurados que provienen de la historia clínica electrónica, también son procesados por bioontology para encontrar su término MESH.

iv. Cuarta fase: Conexión con fuentes de información sanitaria fiable.

En esta fase se utilizan diferentes herramientas externas que permiten la comunicación con las múltiples bibliotecas virtuales. Esta instancia del proceso, es la que define qué estrategia de búsqueda se va a emplear sobre las fuentes de información como por ejemplo PubMed, ClinicalTrials y MedlinePlus.

v. Quinta fase: Formateado.

Tras enviar la consulta a las diferentes fuentes de información válidas que se encuentran en la Red, se extraen los resultados correspondientes. Dichos resultados son desglosados y procesados y devueltos a la historia clínica electrónica como parte de la consulta que realiza el profesional médico

vi. Sexta fase: Incorporación en la Historia Clínica Electrónica

La última fase de HOPE es el envío de los enlaces obtenidos a artículos de PubMed y casos de ClinicalTrials a la historia clínica electrónica con su posterior valoración realizada por el profesional médico, además de incluir recomendaciones sobre MedlinePlus que pueda valer al paciente

i. Grupo de investigación:

El grupo de investigación del proyecto HOPE es el resultado de juntar los conocimientos y habilidades de varios médicos y varios ingenieros en un grupo multidisciplinario bajo el liderazgo del Dr. Sánchez Bocanegra, para desarrollar una herramienta con criterios médicos usando las tecnologías emergente en inteligencia artificial para apoyar a los profesional médicos oncólogos en su labor diaria para la búsqueda automatizada de evidencias científicas centralizadas en el paciente para la toma de decisiones médicas.

El resultado de ello ha sido el desarrollo de un sistema recomendador en salud capaz de extraer datos estructurados de documentos de texto libre utilizando el procesamiento del lenguaje natural (NLP).

Es una plataforma tecnológica que utiliza una implementación de NLP para procesar datos no estructurados. Tras la validación técnica realizada en dos estudios previos (17,18).

En este estudio presentamos los resultados de la fase de prueba de concepto sujeto a una evaluación inicial previa a un proyecto Piloto.

En el presente estudio se evalúa la coincidencia de las recomendaciones con el diagnóstico del paciente con las búsquedas de un médico experto en la base de datos de Pubmed, Clínica Trial y MedlinePlus para prescribir recomendaciones de diagnóstico y tratamiento en patología oncológica basada en MBE que cuenta con evidencia, junto a la medicina centrada en el paciente en la mejora de la calidad y resultados de la atención médica(19).”

2. Objetivos

i. Objetivos primarios

- Evaluar la precisión interna de las recomendaciones proporcionadas por HOPE en base a historia clínica de pacientes según los primeros 10 primeros artículos recomendados y en las 3 bases de datos, Pubmed, MedlinePlus y ClinicalTrials.
- Evaluar la precisión de los primeros 5 artículos recomendados y la concordancia por parte de 2 evaluadores externos.

ii. Objetivos secundarios.

- Evaluar el procesamiento de datos de la HCE, sus resultados de extracción de ontologías sanitarias y la calidad de resultados de búsqueda .
- Crear metodología de evaluación y sugerencias para la realización del piloto inicial

3. Metodología

a. Selección e introducción de los datos de casos clínicos.

La selección de casos y la evaluación interna se ha realizado por parte del equipo médico del grupo HOPE, realizándose conexión con una interfaz intermedia del sistema que se ha programado para la evaluación.

i. Descripción de la interfaz:

La interfaz de prueba de la herramienta implica 4 pestañas principales,

La primera corresponde al formulario de carga de datos, donde se rellenan los siguientes campos:

- Nombre del caso
- Edad y sexo del paciente
- Diagnóstico principal Caso clínico e informe de pruebas complementarias extraídos de la historia clínica.
- El idioma de introducción puede que sea en Español o en inglés, teniendo en cuenta que el resultado para las búsquedas es en inglés siendo el idioma principal de las bases de datos utilizadas. (anexo 1)
- Campos automáticos con la carga de los términos Mesh extraídos de los datos de los pacientes y su historia clínica.

Los siguientes 3 pestañas corresponden a los resultados de búsqueda en Pubmed, MedlinePlus y ClinicalTrials.

ii. Selección de casos:

Se realiza una búsqueda manual en publicación en internet:

- Búsqueda en Pubmed de 5 casos que correspondan a las 5 neoplasias más frecuentes(20): Pulmonar, Hepático, Colorrectal, Gástrico y Mamario.
- Se buscan los siguientes 5 casos con las mismas patologías con artículos en inglés.

- Se realiza una búsqueda de casos clínicos de patología neoplásicas diferentes procurando no repetir patologías.
- Se agregan en el idioma original de la publicación.

iii. Criterios de exclusión:

- Casos con idiomas diferentes al español e inglés.
- Casos sin definición del tumor principal
- Casos clínicos dentro de un grupo de casos en el artículo.

iv. Las métricas analizadas fueron las siguientes:

- La correcta identificación y codificación del lenguaje natural del profesional de la salud de HOPE desde el texto de la historia clínica hasta los términos MeSH, SnomedCT.
- Coincidencia de diagnósticos principales con las recomendaciones proporcionadas por el sistema.
- Términos Mesh opcionales
- El nivel de acierto en búsqueda en las 3 bases datos y el número de artículos validos entre los 10 primeros artículos de la búsqueda.
- Concordancia de estos datos entre dos evaluadores médicos independientes.

Se ha buscado el número de artículos de cada caso que concuerda con el diagnostico suyo en las Pubmed, ClinicalTrials y MedlinePlus.

Se realiza un control del número de casos que resultado correcto en Mesh.

- Detección de palabras claves en la búsqueda e identificar los campos obligatorios y opcionales.
- Se realiza una verificación manual de la carga de datos de cada patología en la que se verifica:
 - Correcta carga de la historia y la detección de términos Mesh, calidad de la detección del diccionario.

- Correcto funcionamiento del código de Pubmed según “main diagnosis” : número de artículos que presentan en título y/o en abstracto el diagnóstico principal.
- Correcto funcionamiento de los términos detectados para ClinicalTrials: número de artículos que presentan en título y/o en abstracto el diagnóstico principal.
- Correcto funcionamiento de términos para MedlinePlus NCI: número de artículos que presentan en título y/o en abstracto el diagnóstico principal.

Se realiza una matriz de evaluación para realizar la base de datos con datos de concordancia, y detección de errores. (Gráfica 1).

				Pubmed	Pubmed	Medline plus		Clinical Trials	
	idioma	patologia	Termino principal MesH	casos con patologia correcta	Acierto	articulos correctos	Acierto	articulos	
paciente 1	español	Adenocarcinoma colon	adenocarcinoma colon	8 de 10	80,00%	1,00	100,00%	0 de 0	0

Gráfica 1 Matriz de evaluación interna.

Se codifica cada resultado en cada una de las bases de datos con una puntuación de 1-10 según el número de artículos en las que los evaluadores consideran que corresponden al caso clínico analizado.

Se evalúa:

- Número de artículos con mención a la patología principal.
- Número de artículos que el médico evaluador considera significativo para el caso.

b. Evaluación externa de precisión del sistema en PubMed.

Se realiza una encuesta en GoogleForms® en que consistía en 150 preguntas correspondientes a 30 casos clínicos.

La pregunta que se plantea a los evaluadores es: *“Buscamos valorar el acierto de un sistema automático de generación de resultados en Pubmed basado en una historia clínica básica.”*

Inicialmente se describe el caso clínico con la información más relevante y como respuestas se expone solo el título de los 5 primeros artículos en Pubmed.

Se solicita una respuesta con SI/NO a la siguiente pregunta: “Acepto este resultado como útil para este caso” (anexo 2).

Se solicita a 2 evaluadores médicos externos acceder y responder a la encuesta realizada y posteriormente se exporta la encuesta para analizar los resultados con el programa estadístico STATA®, realizando estudio de la concordancia según **Kappa de Cohen**, que es un índice que relaciona el acuerdo que exhiben los observadores, más allá del debido al azar, con el acuerdo potencial también más allá del azar. Se calcula basado en la siguiente fórmula(21):

$$k = \frac{Pr(a) - Pr(e)}{1 - Pr(e)},$$

donde Pr(a) es concordancias observadas y Pr(e) es la concordancia atribuida al azar. Si los evaluadores tienen máximo acuerdo, el valor k sería de 1, y si tuviesen desacuerdo total, el valor sería de 0.

Para la valoración de los valores intermedios, seguiremos la siguiente tabla (22)

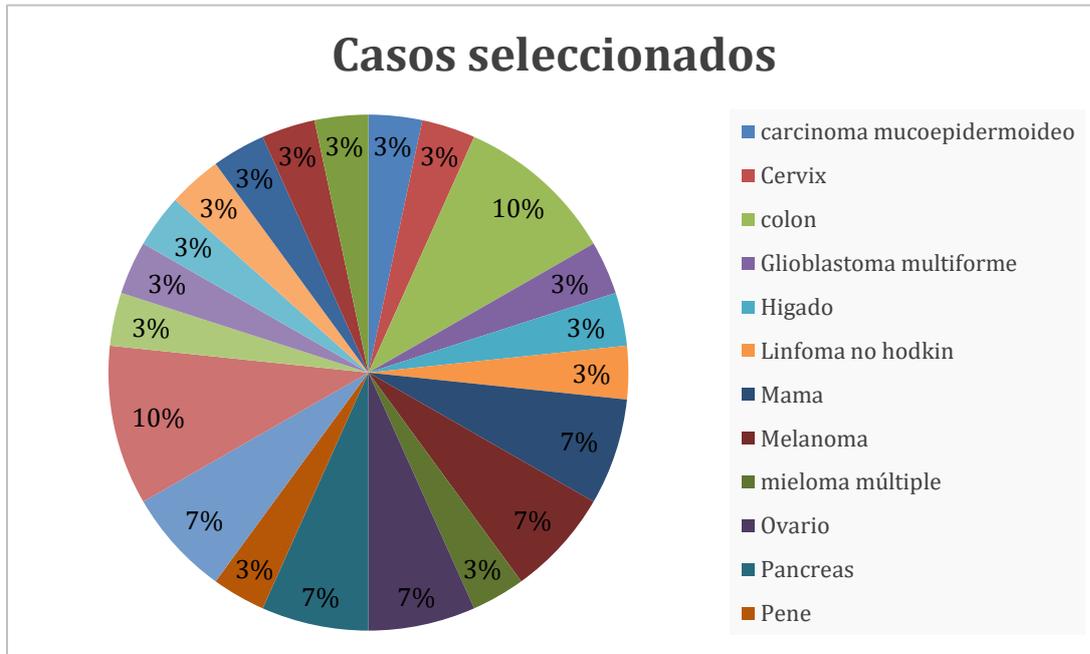
Coefficiente kappa	Fuerza de la concordancia
0,00	Pobre (<i>Poor</i>)
0,01 - 0,20	Leve (<i>Slight</i>)
0,21 - 0,40	Aceptable (<i>Fair</i>)
0,41 - 0,60	Moderada (<i>Moderate</i>)
0,61 - 0,80	Considerable (<i>Substantial</i>)
0,81 - 1,00	Casi perfecta (<i>Almost perfect</i>)

Gráfica 2 Valoración del índice Kappa

4. Resultados

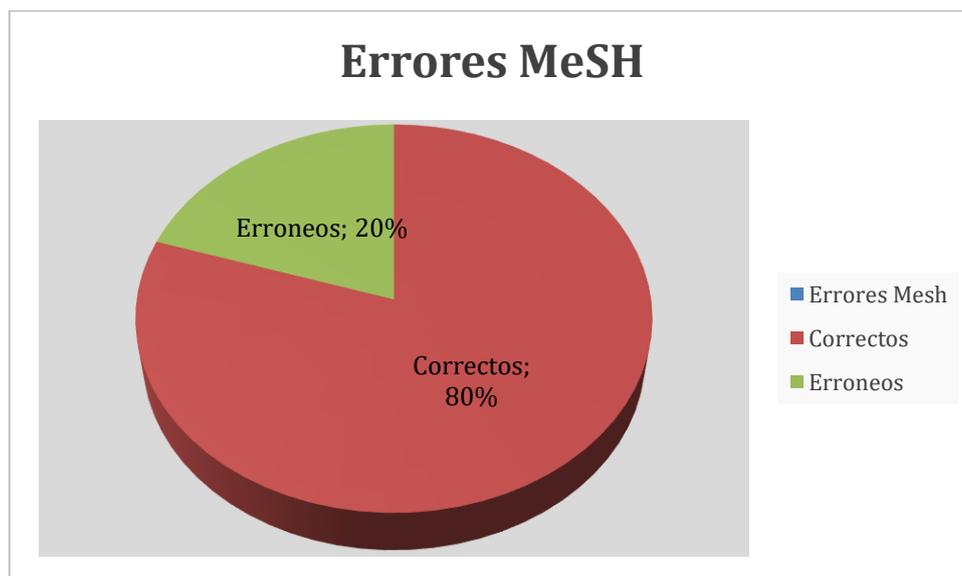
a. Descripción de resultados:

Se han recogido un total de **30 casos** clínicos que corresponden a los siguientes tipos de canceres. (gráfico 2)



Gráfica 3 Número de casos recogidos

De todos los casos, se han evaluado la adecuación de la detección de términos MeSH del diagnóstico principal, habiendo un 80% de acierto. (gráfico 2)

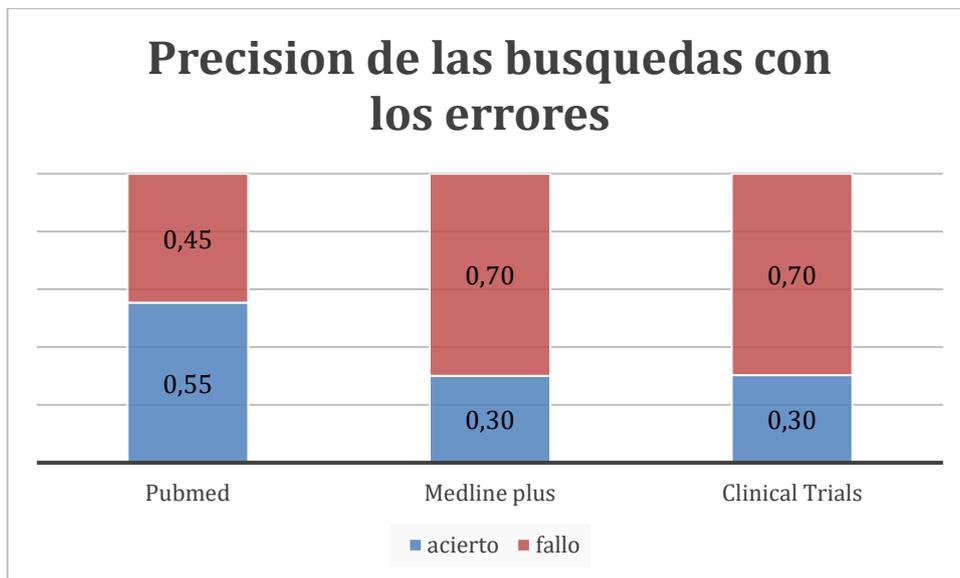


Gráfica 4 Errores detección de términos Mesh

Los principales errores en Mesh han sido en patologías correspondientes a: Pulmón, Próstata, Páncreas, Riñón, Recto.

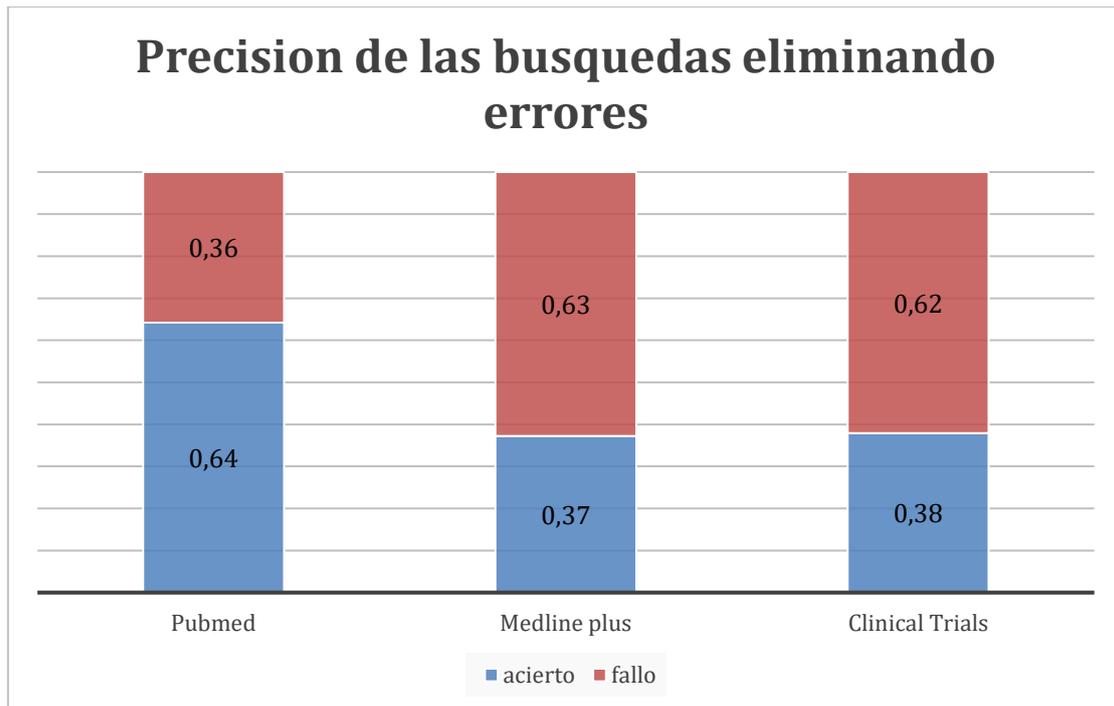
b. Evaluación interna del sistema:

En la evaluación inicial, por parte del equipo médico del grupo de investigación se han apreciado un nivel de precisión en los resultados de 55,33% en Pubmed, siendo la más precisa, mientras en MedlinePlus y ClinicalTrials ha llegado a 30% de aciertos.



Gráfica 5 Precisión de las búsquedas en las 3 bases de datos.

Teniendo en cuenta los errores de la detección en los términos MeSH, retirando estos casos con error, se mejora la precisión hasta 64% en Pubmed y 37% en las otras dos bases de datos.



Gráfica 6 Precisión de las búsquedas en las 3 bases de datos sin errores MeSH.

c. Evaluación externa de concordancia en Pubmed:

Se han evaluado los resultados del cuestionario dando resultado de 150 preguntas respuestas por cada uno de los 2 evaluadores. Tras su evaluación en STATA® y el resultado de concordancia entre ambos evaluadores haya sido medida vía Kappa de Cohen con un resultado inicial de Kappa 0.31 correspondiente a un índice de concordancia **discreto**.

Agreement	Expected Agreement	Kappa	Std. Err.	Z	Prob>Z
64.43%	48.44%	0.3101	0.0718	4.32	0.0000

Gráfica 7 Kappa de Cohen entre evaluador 1 y 2

d. Evaluación externa de precisión en Pubmed:

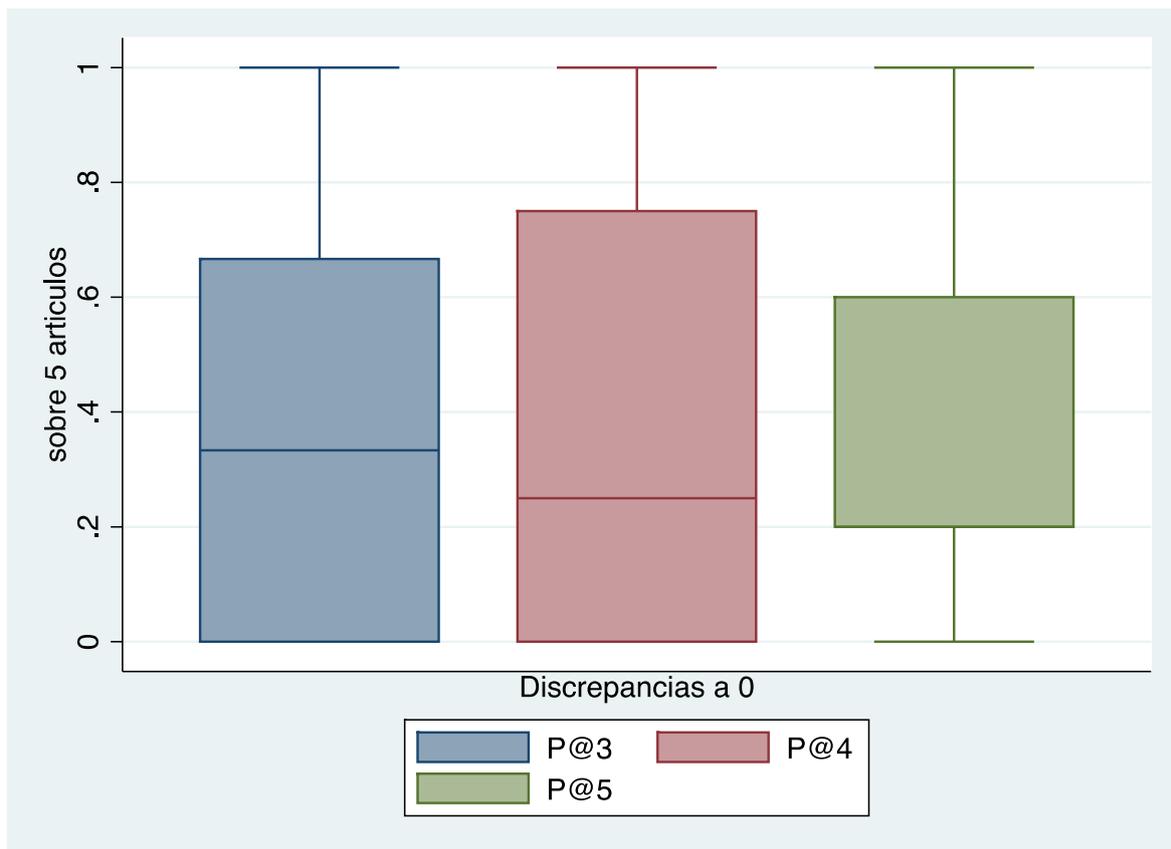
Se evalúa la precisión de las respuestas de los evaluadores, se ha buscado la precisión de las respuestas con los primeros 3 artículos, los primeros 4 y en todos los artículos (los primeros 5 que nos ha recomendado HOPE). Se han aplicado dos criterios, uno más restrictivo y otro más flexible.

En el primer caso más restrictivo, en el que se aplica la fórmula para que ambos evaluadores hayan aceptado el resultado de los artículos, en este apreciamos que la media está entre 0 y 1 es:

Precisión a 3 preguntas: 0,38 +- 0,38

Precisión a 4 preguntas: 0,41 +- 0,34

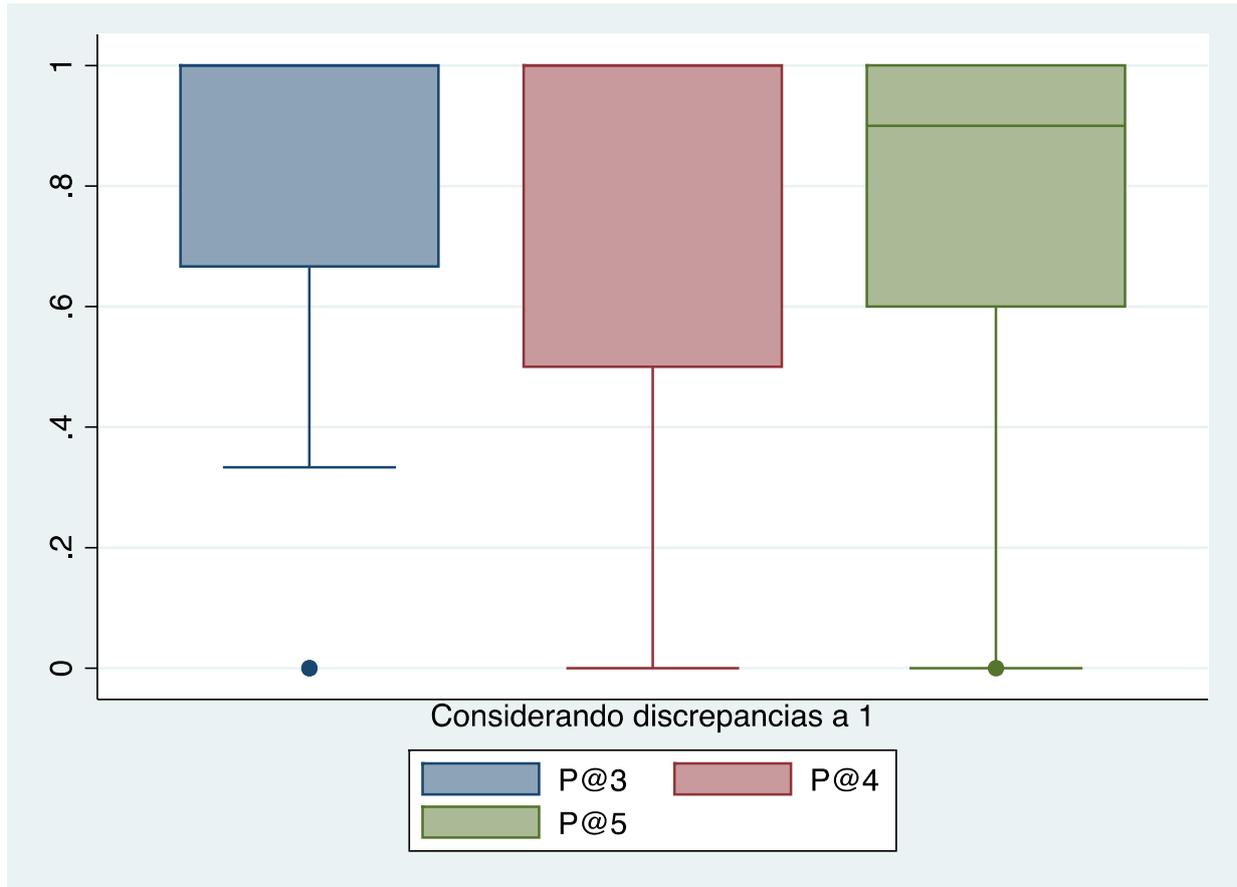
Precisión a 5 preguntas: 0,41 +- 0,33



Gráfica 8: Discrepancias si ambos evaluadores están conformes

Aplicando un criterio más flexible, en el que aceptamos que al menos uno de los dos evaluadores haya aceptado el resultado, detectamos los siguientes resultados:

- Precisión a 3 preguntas: 0,77 +- 0,34
- Precisión a 4 preguntas: 0,76 +- 0,33
- Precisión a 5 preguntas: 0,76 +- 0,32



Gráfica 9: Discrepancias si al menos un evaluador conforme

5. Discusión:

En esta fase del proyecto, nos encontramos en la fase de prueba de concepto del sistema. Es un proyecto que se encuentra análisis, evaluación y mejora continua de los procesos implicados en éste. Estos resultados y hallazgos identificados en ésta evaluación son fundamentales para, en base a ellos, realizar las siguientes versiones del sistema con las mejoras propuestas. Siendo uno de los principales asuntos por perfeccionar, la obtención de los términos y de los resultados proporcionados por el sistema.

Los principales errores que se han detectado han sido la ausencia de algunos términos en los diccionarios NLP, por lo que se han producido fallos en la detección de palabras claves, y teniendo en cuenta que al eliminar los casos con error de este tipo, hemos podido mejorar la precisión de la recomendación.

En cuanto a la concordancia de los artículos recomendados en Pubmed, el índice Kappa nos indica una concordancia discreta, aunque durante la realización de la encuesta por pare de los evaluadores, han expresado que algunas respuestas les ha causado dudas por falta de información que explique mejor el resultado.

Cuando hemos medido la precisión, detectamos que no hay diferencia significativa a la hora del punto de corte a 3, 4 ó 5 resultados, aunque puede que sea debido a que los resultados no están en orden de relevancia, y mejora sustancialmente la precisión el hecho de aplicar criterios más flexibles entre ambos evaluadores.

En próximas versiones de evaluación, sugerimos mejorar las descripciones de los resultados y aumentaríamos el número de evaluadores para conseguir una mayor significación estadística, mejorando la concordancia entre ellos y pudiendo acceder a evaluadores médicos de diferentes especialidades que pueda que los resultados les fueran valido y menos válidos a otros, por ejemplo, en el caso de cáncer de colon, puede que el resultado de paciente con hematoquecia sea de valor para un oncólogo clínico ó un médico de familia, pero no es válido para a un oncólogo radioterapéutico.

a. Limitaciones identificadas.

Los evaluadores fueron parte del equipo de análisis, aunque se ha mantenido un esfuerzo para responder de forma objetiva, teniendo en cuenta la fase de la evolución inicial y que en próximas fases, se recurrirá a evaluadores externos.

Los evaluadores voluntarios solo han tenido acceso al título del artículo en pubmed, y en algunos casos, los títulos han sido poco descriptivos o escasos en información. En esta fase, aunque el sistema HOPE sí que lo proporciona, no hemos añadido los resúmenes ni enlaces a los artículos en el formulario, para no extender en el tiempo necesario para responder a toda la encuesta, primando el número de casos clínicos evaluados (anexos 3).

b. Sugerencias de mejora.

Las principales sugerencia al equipo de desarrollo están en relación a los diccionarios NLP y a la mejora de la eficiencia de las búsquedas:

- Se evalúan diferente opciones de campos de detección automática para limitar los resultados y hacerlos más eficiente:
- Añadir un glosario de términos genéricos que su identificación resta precisión a la búsqueda, para que no se tengan en cuenta a la hora de la recomendación.
- Añadir funcionalidad de limitar búsquedas por diagnóstico y/o tratamientos.
- Agregar la estatificación de las neoplasias.
- Se ha creado una matriz para evaluación clínica que sirve para futuras evaluación internas.
- Se ha creado una encuesta que permite reproducir los resultados de precisión y de concordancia en las siguientes versiones del sistema, siendo una base para prueba piloto con criterios de evaluación clínica.

6. Conclusión

- 1) La medicina basada en evidencias tiene como pilar fundamental la búsqueda de información desde fuentes bibliográficas científicas. Para realizar este proceso se requiere una gran inversión de tiempo lo cual limita su uso en la práctica clínica habitual.
- 2) Se ha desarrollado un sistema que permite transformar un lenguaje estructurado para la obtención de términos semánticos de ontologías sanitarias que nos permite poder extraer recomendaciones de fuentes fiables de salud, pudiendo utilizar el idioma español para acceder a la literatura médica actualizada.
- 3) En esta fase de prueba de concepto, el sistema HOPE ha demostrado poder recomendar artículos en 3 bases de datos científicos con una precisión aceptable.
- 4) La precisión de las recomendaciones, aunque no son muy concluyentes a nivel de precisión, demuestran que están el desarrollo de la aplicación está yendo en una buena dirección, y con unos ajustes en la capa de programación, se puede mejorar el rendimiento y la precisión del sistema HOPE.
- 5) En las siguientes etapas se han planteado realizar una evaluación a mayor escala, con historia clínica de pacientes reales y un número de evaluadores superior e independiente, para ir escalando la calidad de la recomendación y su adecuación a las necesidades de los médicos oncólogos a la hora de la toma de decisiones clínicas.

7. Bibliografía.

1. Gérvas J PFM. Uso apropiado de la medicina basada en pruebas, revisión de diez artículos recientes. *Amf.* 2005;1(1):46–56.
2. Para ver el tamaño de la base de datos de pubmed, buscar “1800:2100[dp]” en <https://www.ncbi.nlm.nih.gov/pubmed/>.
3. Reyes-Aldasoro CC. The proportion of cancer-related entries in PubMed has increased considerably; is cancer truly “The Emperor of All Maladies”? Novelli G, editor. *PLoS One* [Internet]. 2017 Mar 10 [cited 2018 Jun 21];12(3):e0173671. Available from: <http://dx.plos.org/10.1371/journal.pone.0173671>
4. March Cerdà JC. Pacientes empoderados para una mayor confianza en el sistema sanitario. *Rev Calid Asist* [Internet]. 2015 Jan [cited 2018 Jun 21];30(1):1–3. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S1134282X15000056>
5. Yoo I MM. Recent research for MEDLINE/PubMed: short review. *Proc ACM fourth Int Work Datatext Min Biomed informatics.* 2010;October 26.
6. Yoo I, Mosa ASM. Analysis of PubMed User Sessions Using a Full-Day PubMed Query Log: A Comparison of Experienced and Nonexperienced PubMed Users. *JMIR Med informatics* [Internet]. 2015 Jul 2 [cited 2018 Jun 12];3(3):e25. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26139516>
7. *Clinical Trials* [Internet]. [cited 2018 Jun 1]. Available from: <https://www.clinicaltrials.gov/ct2/about-site/background>
8. *MedlinePlus* [Internet]. Available from: <https://medlineplus.gov/spanish/aboutmedlineplus.html>
9. Gruber TR. Toward principles for the design of ontologies used for knowledge sharing? *Int J Hum Comput Stud* [Internet]. 1995 Nov 1 [cited 2018 Jun 21];43(5–6):907–28. Available from: <https://www.sciencedirect.com/science/article/pii/S1071581985710816>
10. Teresa Romá-Ferri M, Palomar M. Análisis de terminologías de salud para su utilización como ontologías computacionales en los sistemas de información clínicos. *Gac Sanit* [Internet]. 2008 Sep 1 [cited 2018 Jun 21];22(5):421–33. Available from: <https://www.sciencedirect.com/science/article/pii/S0213911108724178>
11. CIE-10-MC [Internet]. Available from: https://eciemaps.msssi.gob.es/ecieMaps/browser/index_10_mc_old.html
12. *Medical Subject Headings (MeSH), Introduction 2004.* Bethesda: US National Library of Medicine (NLM). Available from: <https://www.nlm.nih.gov/mesh/introduction.html>
13. *SnomedST* [Internet]. [cited 2006 Aug 20]. Available from: <https://www.snomed.org/snomed-ct/what-is-snomed-ct>
14. *Bioontology* <http://www.biontology.com/>.
15. *PLN* [Internet]. Available from: <http://www.vicomtech.org/t4/e11/procesamiento-del-lenguaje-natural>
16. Shadbolt N, Berners-Lee T, Hall W. The Semantic Web Revisited. *IEEE Intell Syst* [Internet]. 2006 May [cited 2018 Jun 21];21(3):96–101. Available from: <http://ieeexplore.ieee.org/document/1637364/>
17. Chacón Vargas KA. HOPE Desarrollo de una metodología piloto para vinculación a Clinical Trials como herramienta de medicina basada en evidencias, para proporcionar medicina centrada en el paciente. 2017 Jul 1 [cited 2018 Apr 25]; Available from: <http://hdl.handle.net/10609/66645>
18. Passadore NA. Recomendaciones de artículos a través de PUBMED. 2017; Available from: <http://hdl.handle.net/10609/66786>
19. Sacristán JA. Evidence based medicine and patient centered medicine: some thoughts on thier integration. *Rev Clin Esp* [Internet]. 2013 Dec [cited 2018 Apr 25];213(9):460–4. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24409523>
20. Plummer M, de Martel C, Vignat J, Ferlay J, Bray F, Franceschi S. Global burden of cancers attributable to infections in 2012: a synthetic analysis. *Lancet Glob Heal* [Internet]. 2016 Sep [cited 2018 Apr 25];4(9):e609–16. Available from: <http://linkinghub.elsevier.com/retrieve/pii/S2214109X16301437>
21. Sim J WC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther.* 2005;85(2):257–68.
22. Landis JR KG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159–74.

8. Anexos:

Paciente prueba 1: neo colon (70cc9061-3f13-4dc9-8b40-f02d0d4c06cb) Load EMR

ID: **Name:**

Changed: 1525974974 / Created: 1522955393

Age: **Gender:** **Affected organ:**

Metastasis: **Region:**

Main diagnostic: **Main diagnostic (mesh terms):** Adenocarcinoma Colon

Medical History:

Procedencia: Localidad rural
 APP: Madre fallecida por cáncer de pulmón
 Padre fallecido por cáncer de cólon
 Abuela fallecida por cáncer de mama
 Motivo de Consulta: Dolor abdominal, distensión y vómitos.
 AEA: Paciente que comienza 2 días antes de la consulta con dolor en FI, tipo cólico, no irradiado, no se alivia ni se exacerba con ninguna conducta.
 También refiere cambios en el hábito intestinal, y en el momento de la

Medical History (mesh terms): Pain Eructation Recurrence Adenocarcinoma Colon
 Abdomen Rectum Constipation Vomiting
 Colonoscopy Dilatation, Pathologic Analgesics
 Tomography, X-Ray Computed Signs and Symptoms Touch
 Diagnosis Breast Neoplasms Hysterectomy
 Digital Rectal Examination Breast Intestines Biopsy
 Symptom Flare Up Substance-Related Disorders
 Neoplasms Colic Herpes Zoster Enema
 Abdominal Pain Catharsis

Anexo 1 Ejemplo caso clínico en formulario principal HOPE

Form **PubMed** MedlinePlus Clinical Trials

PubMed

Query

```
[*db*: "pubmed", *term*: "("Adenocarcinoma"[mesh] AND ("Colon"[mesh] AND ("middle aged"[mesh] AND ("female"[mesh] AND ("Pain"[mesh] OR ("Erectation"[mesh] OR ("Recurrence"[mesh] OR ("Abdomen"[mesh] OR ("Rectum"[mesh] OR ("Constipation"[mesh] OR ("Vomiting"[mesh] OR ("Colonoscopy"[mesh] OR ("Dilatation, Pathologic"[mesh] OR ("Analgescics"[mesh] OR ("Tomography, X-Ray Computed"[mesh] OR ("Signs and Symptoms"[mesh] OR ("Touch"[mesh] OR ("Diagnosis"[mesh] OR ("Breast Neoplasms"[mesh] OR ("Hysterectomy"[mesh] OR ("Digital Rectal Examination"[mesh] OR ("Breast"[mesh] OR ("Intestines"[mesh] OR ("Biopsy"[mesh] OR ("Symptom Flare Up"[mesh] OR ("Substance-Related Disorders"[mesh] OR ("Neoplasms"[mesh] OR ("Colic"[mesh] OR ("Herpes Zoster"[mesh] OR ("Enema"[mesh] OR ("Abdominal Pain"[mesh] OR ("Catharsis "[mesh]))", *datatype*: "eedit", *retmax*: 10, *sort*: "relevance" ]
```

Data

#1 - Approach to the Patient With Hematochezia.

Cotter Thomas G Buckley Niamh S Loftus Conor G

The evaluation of the patient with hematochezia can be complex because of the broad differential diagnosis and the number of management strategies available. In this article, a simplified approach to the history and physical examination is presented, with management illustrated in a case-oriented manner.

pubmed: 28473039 pii: S0025-6196(17)30071-X doi: 10.1016/j.mayocp.2016.12.021

#2 - Incidental focal uptake in colorectal location on oncologic ¹⁸F-DG PET and PET/CT studies: histopathological findings and clinical significances.

Fuertes Jordi Montagut Clara Bullich Santi Coma Mar Iglesias Mestre-Fusco Antoni Suárez-Piñera Marina Trampal Carlos Belmont Joaquim

Unexpected focal colonic or rectal radiotracer activity is an usual finding in patients subjected to a PET study. The aim of this work has been to evaluate the clinical significance of this finding in the prediction of an existing colorectal malignancy.

During the last three years, all patients studied with (18)F-FDG PET/CT and PET for oncologic work-up purposes were prospectively surveyed for focal colorectal radiotracer activity. Colonoscopy was performed in all patients with this incidental finding in order to exclude colonic malignancy. CEA level, maximum standardized uptake value (SUVmax), CT findings, colonoscopy findings and histopathological results were prospectively analyzed in all patients.

A total of 2290 patients were evaluated, 158 of whom were studied with PET and the remainder with a hybrid PET/CT. Focal FDG colorectal activity was incidentally detected in 27 patients with no previous history of colorectal cancer. Colorectal adenocarcinoma was diagnosed in seven (25.9%) patients. A pre-cancerous lesion was found in eleven patients (40.7%). Eight patients (29.6%) had no macroscopic lesions. One patient was diagnosed with a benign lesion. Any focal activity found in the colon by (18)F-FDG PET/CT examination predicts a probability greater than 50% of an underlying malignant or premalignant lesion in the histopathological analysis (logistic regression, p=0.01), independently of the calculated SUVmax.

According to the results of the present study, we recommend the performance of a colonoscopy and biopsy of any suspicious lesions, in all patients with unexpected focal FDG activity found in colon or rectum during a (18)F-FDG PET/CT examination.

pubmed: 25263718 pii: S2253-654X(14)00149-8 doi: 10.1016/j.rem.2014.07.008

#3 - Primary Adenocarcinoma with Focal Choriocarcinomatous Differentiation in the Sigmoid Colon.

Oh Soek Kyoung Kim Hyung Wook Kang Dae Hwan Choi Cheol Woong Choi Yu Yi Lim Hong Kyu Goo Ja Jun Choi Sung Yeol

Primary colorectal choriocarcinoma is a rare neoplasm. Only 19 cases have been reported worldwide, most of which involved adenocarcinomas. The prognosis is usually poor, and the standard therapy for this tumor has not been established. A 61-year-old woman presented with constipation and lower abdominal discomfort. She was diagnosed with primary adenocarcinoma with focal choriocarcinomatous differentiation in the sigmoid colon and liver metastasis. Because the serum beta-human chorionic gonadotropin level was not significantly elevated, and because only focal choriocarcinomatous differentiation was diagnosed, we selected the chemotherapy regimen that is used for the treatment of metastatic colorectal adenocarcinoma. The patient survived for 13 months after the initial diagnosis. This is the first case in Korea to assess the suppressive effects of the standard chemotherapy for colorectal adenocarcinoma against coexisting colorectal choriocarcinoma and adenocarcinoma.

pubmed: 26588354 pii: 10.4166/kjg.2015.66.5.291 doi: 10.4166/kjg.2015.66.5.291

#4 - In Vitro Adenosine Triphosphate-Based Chemotherapy Response Assay as a Predictor of Clinical Response to Fluorouracil-Based Adjuvant Chemotherapy in Stage II Colorectal Cancer.

Anexo 2 Ejemplo resultados completos de PubMed

Evaluacion HOPE 2.0

Buscamos valorar el acierto de un sistema automatico de generacion de resultados en Pubmed basado en una historia clinica básica.

Paciente 1: Mujer 51 años. Diagnostico principal: Adenocarcinoma colon estadio 4. DUKES 4.

Short-answer text
.....

#1 - Approach to the Patient With Hematochezia.

- Acepto este resultado como útil para este caso
- NO acepto este resultado.

#2 - Incidental focal uptake in colorectal location on oncologic ¹⁸F-FDG PET and PET/CT studies: histopathological findings and clinical significances.

- Acepto este resultado como útil para este caso
- NO acepto este resultado.

#3 - Primary Adenocarcinoma with Focal Choriocarcinomatous Differentiation in the Sigmoid Colon.

- Acepto este resultado como útil para este caso
- NO acepto este resultado.

#4 - In Vitro Adenosine Triphosphate-Based Chemotherapy Response Assay as a Predictor of Clinical Response to Fluorouracil-Based Adjuvant Chemotherapy in Stage II Colorectal Cancer.

- Acepto este resultado como útil para este caso
- NO acepto este resultado.

#5 - En-bloc pelvic resection with concomitant rectosigmoid colectomy and immediate anastomosis as part of primary cytoreductive surgery for patients with advanced ovarian cancer.

Anexo 3 Ejemplo de pregunta de la encuesta