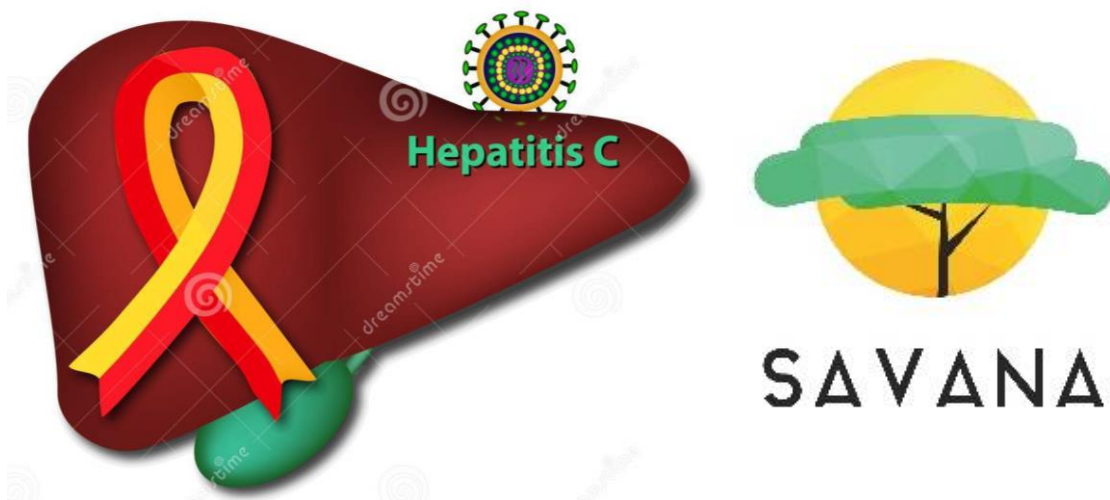


Trabajo de fin de grado 2019/2020

*Utilidad de un sistema de análisis
masivo de datos (Big Data)
insertado en la historia clínica
electrónica, en la búsqueda activa
de pacientes con hepatitis C*



Autor: Miguel Bardají Carrillo

Tutora: Dra. Gloria Sánchez Antolín





“Fija tu rumbo a una estrella,
y podrás navegar a través de cualquier tormenta”.

Leonardo da Vinci

Mi más sincero agradecimiento a la Dra. Sánchez Antolín, sin la cual este TFG no
habría sido posible.

Sin olvidarme de mis padres y mi hermana, de mi pareja, ni de Chavulería, por su
apoyo incondicional.



Índice

Resumen	4
1. Introducción	6
1.1. Historia de la hepatitis C	6
1.2. Epidemiología de la hepatitis C	6
1.3. Clínica y diagnóstico de la hepatitis C	8
1.4. Tratamiento de la hepatitis C	8
1.5. Eliminación de la hepatitis C	9
1.6. Inteligencia artificial y Big Data	11
2. Objetivos	12
2.1. Objetivos principales del estudio	12
2.2. Objetivos secundarios	12
3. Material y métodos	12
3.1. Selección de candidatos	12
3.2. Recogida de datos clínicos	13
3.3. Variables a estudio	13
3.4. Búsqueda bibliográfica	13
3.5. Aspectos éticos	14
3.6. Análisis estadístico	14
4. Resultados	15
4.1. Resultados generales	15
4.2. Resultados por servicios	18
4.3. Resultados en pacientes sin seguimiento por U. especializada	21
4.4. Resultados en pacientes seguidos previamente por U. especializada	22
5. Discusión	25
6. Conclusiones	28
7. Bibliografía	29
8. Póster	30
9. Resolución comité ético	31



Resumen

Introducción: La hepatitis C (VHC) es causa de cirrosis, hepatocarcinoma y la indicación más frecuente en el mundo de trasplante hepático. Es un objetivo de la OMS conseguir su eliminación para el 2030. Según el estudio de seroprevalencia del Ministerio de Sanidad, en España hay aproximadamente una prevalencia de Ac anti-VHC del 0,85%, y de RNA VHC + del 0,22%, la mayor parte asintomáticos. La eliminación del VHC pasa por la búsqueda de pacientes no diagnosticados o no controlados y, por tanto, sin acceso a tratamientos actualmente muy eficaces.

Objetivos: Detectar, mediante la herramienta informática Savana© (inteligencia artificial y Big Data), pacientes con hepatitis C que no reciben seguimiento por una Unidad especializada (U. de Hepatología o U. de Medicina Interna) para que puedan recibir tratamiento. Valorar a su vez la fiabilidad de esta herramienta.

Material y métodos: Estudio observacional retrospectivo, en el que se seleccionaron todos los pacientes diagnosticados de hepatitis C localizados por Savana entre el 01/01/2013 y el 31/12/2018. Se seleccionó a todos los pacientes que, con diagnóstico de hepatitis C, estaban controlados por servicios hospitalarios diferentes de Hepatología o Medicina Interna (MI), y, por tanto, sin acceso a tratamiento antiviral. Posteriormente se revisó su historia clínica electrónica para valorar la necesidad de confirmar el diagnóstico, y valorar la funcionalidad de Savana. Posteriormente se pretendía, además, contactar con los pacientes detectados por carta para confirmar su infección y ofertar tratamiento, pero la pandemia SARS-CoV-2 lo ha impedido.

Resultados: Se detectaron 398 pacientes “hepatitis C” en Savana, sin seguimiento por una Unidad especializada. Tras la revisión de las historias clínicas, se observó que 105 (26,4%) pacientes eran seguidos en Hepatología o MI y, por tanto, estaban inadecuadamente clasificados por Savana. Además, se localizaron 89 pacientes que no tenían seguimiento por una Unidad especializada. Hasta 16 servicios tenían pacientes con hepatitis C sin seguimiento por una Unidad especializada. A su vez, el diagnóstico más frecuente fue Enfermedad Renal Crónica (ERC).

Conclusiones: La Inteligencia artificial es una herramienta eficaz para detectar pacientes con hepatitis C sin control clínico adecuado. Aunque Savana necesite aprendizaje, hemos detectado 89 pacientes que podrían ser subsidiarios de tratamiento. Podría ser útil la creación de un protocolo o alerta para derivar a los pacientes a Unidades especializadas.

Palabras clave: hepatitis C, Savana, inteligencia artificial, Big Data, seguimiento.



Abstract

Background: Hepatitis C (HCV) is a main cause of cirrhosis, hepatocellular carcinoma, and the most frequent cause of hepatic transplant in the world. WHO aims to eliminate HCV by 2030. According to the Spanish seroprevalence study, in Spain there is, approximately, a 0,85% prevalence of antibodies anti-HCV and a 0,22% prevalence of RNA VHC+, the majority being asymptomatic. HCV elimination depends on the search of undiagnosed patients, or diagnosed patients not controlled in liver Units and, hence, without access to the highly effective treatments.

Objectives: To detect, by using the informatic tool Savana© (artificial intelligence and Big Data), hepatitis C patients who are not being monitored by and specialized Unit (Hepatology or Internal Medicine) so that they can be treated. Assess the reliability of this tool.

Material and methods: this is a descriptive observational study, in which we selected all the hepatitis C diagnosed patients detected by Savana between 01/01/2013 and 31/12/2018. We selected patients diagnosed with hepatitis C being supervised by medical controlled by hospital services different to Hepatology or Internal Medicine, and, hence, without access to the highly effective treatment. Afterwards, we reviewed their electronic medical records to confirm our findings and assess Savana's functionality. Afterwards, we intended to contact those patients by letter to confirm their infection and offer treatment, yet this could not be accomplished due to the SARS-CoV-2 pandemic.

Results: We found 398 "hepatitis C" patients in Savana uncontrolled by a liver Unit. After the medical histories' review, we detected 105 (26,4%) patients that were being controlled previously by a liver Unit and, consequently, inadequately classified by Savana. Simultaneously, we detected that 89 patients were not controlled in a liver Unit. There were patients diagnosed with hepatitis C without control by a specialized Unit in 16 medical departments. The most frequent diagnosis in this group of patients was the Chronic Kidney Disease (CKD).

Conclusions: Artificial Intelligence is an effective tool to detect patients with hepatitis C without a suitable medical control. Although Savana needs amelioration, we detected 89 patients that could need treatment. Furthermore, the creation of a protocol or an alert system in the electronic medical record, to refer patients to specialized Units would be interesting.

Keywords: hepatitis C, Savana, artificial intelligence, Big Data, control.



1. Introducción

1.1. Historia de la hepatitis C

La hepatitis C es una enfermedad provocada por el virus de la hepatitis C (VHC). Inicialmente se denominó hepatitis no A y no B, era un diagnóstico de exclusión. En 1989, Choo y cols. aislaron el ARN del VHC, dando con la causa de esta enfermedad^(1,2). Hoy en día, hay descubiertos hasta 6 genotipos y 50 subtipos de este virus⁽³⁾.

1.2. Epidemiología de la hepatitis C

Esta enfermedad se transmite principalmente por vía sanguínea, por lo que sus vías de contagio son el consumo de drogas por vía venosa, las transfusiones no controladas (en países desarrollados prácticamente inexistente), el uso de material médico sin la correcta esterilización, o las relaciones sexuales (sobre todo en hombres que tienen sexo con hombres), traumáticas y con múltiples parejas sexuales.

Una vez contraída la infección, la mayoría de los pacientes sufren una hepatitis crónica, y de estos un 15-30% pueden acabar desarrollando cirrosis⁽⁴⁾ y hepatocarcinoma. De hecho, en la actualidad, la hepatitis C es una causa importante de cirrosis y carcinoma hepatocelular, y la 1ª indicación de trasplante hepático en el mundo.

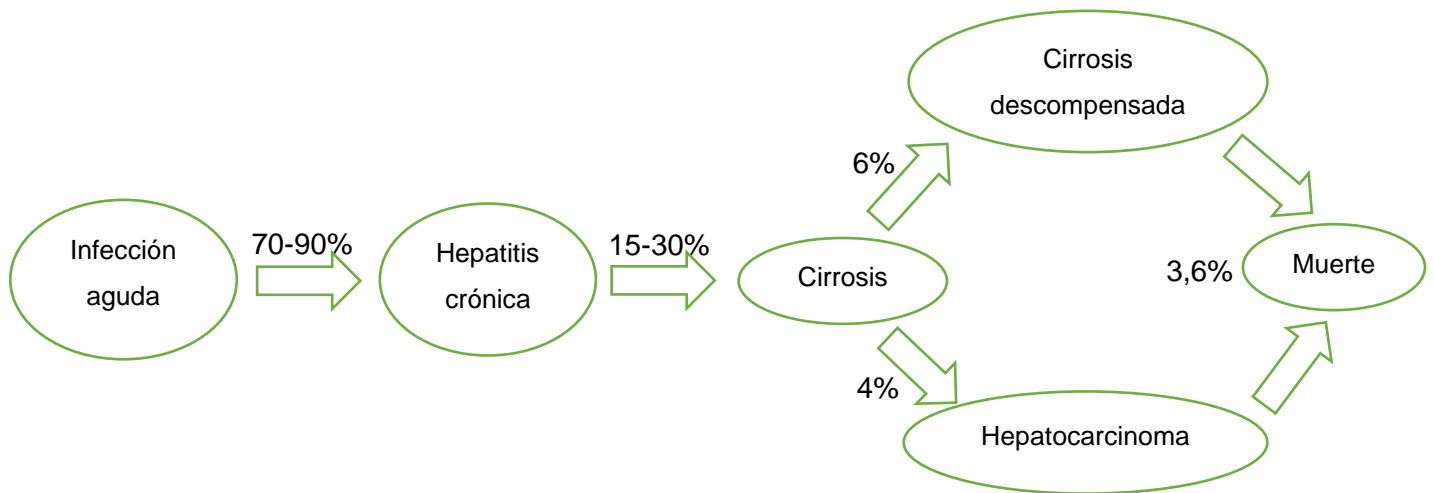


Figura 1. Historia natural de la hepatitis C.

Se estimaba hace unos años, que en el mundo había más de 70 millones de personas afectadas. En nuestro país estudios de prevalencia de hepatitis C poblacionales, realizados entre 1993 y 2017 (**figura 2**) muestran prevalencias de Ac anti-VHC entre el 0,7% en Zamora en 2001⁽⁵⁾ y el País Vasco en 2009⁽⁶⁾ y el 2,64% en Girona y Tarragona en 1995⁽⁷⁾.

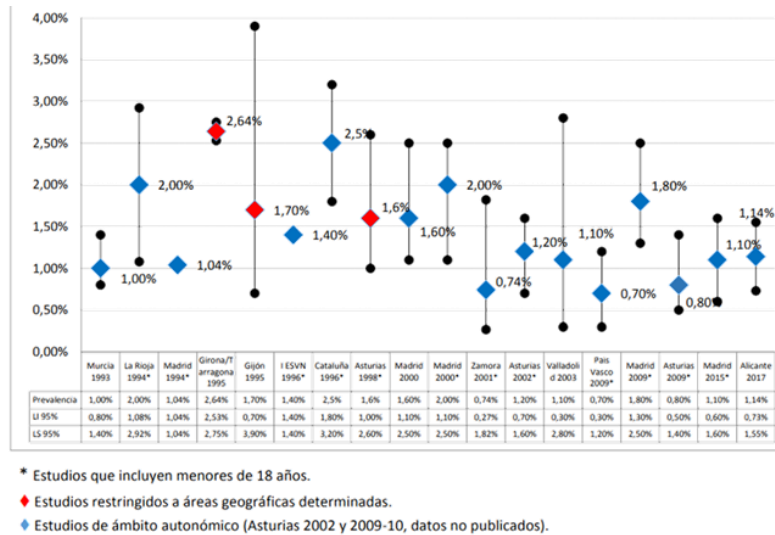


Figura 2. Estimación de seroprevalencia de Ac frente a la hepatitis C en España: revisión de encuestas de seroprevalencia realizadas en España entre 1993-2018⁽⁸⁾.

El Ministerio de Sanidad ha publicado recientemente una encuesta de seroprevalencia⁽⁸⁾ que muestra que en la población de 20 a 80 años la prevalencia de Ac anti-VHC es de 0,85% (IC 95%: 0,64%-1,08%) (**figura 3**) y la prevalencia de infección activa, es decir con ARN positivo, fue de 0,22% (IC 95%: 0,12%-0,32%) (**figura 4**)⁽⁸⁾.

En los casos con RNA +, el genotipo más frecuente fue el 1b (41,18%). Menos frecuentes, el 1a (23,53%), 3a (11,76%), 2c (5,88%), 4a (5,88%) y no concluyente (11,76%)⁽⁸⁾. Se estima que el virus causa aproximadamente 4.800 muertes anuales en España, lo que da cuenta de la necesidad de un diagnóstico para poder administrar los tratamiento efectivos⁽⁹⁾.

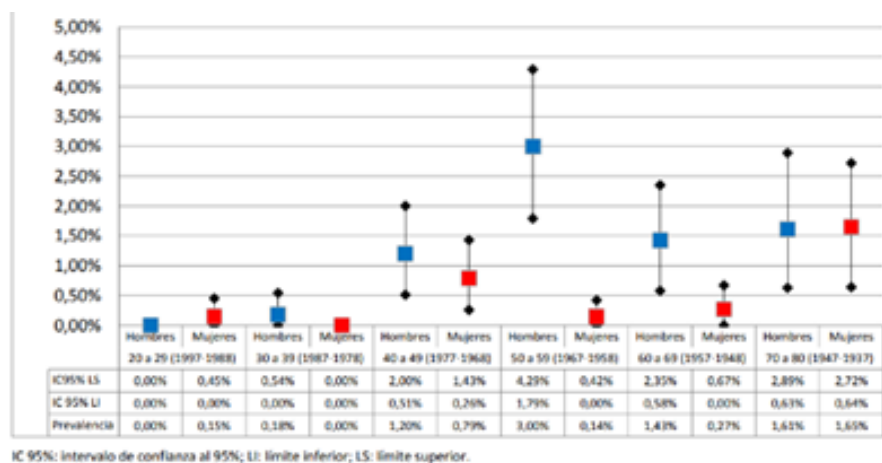
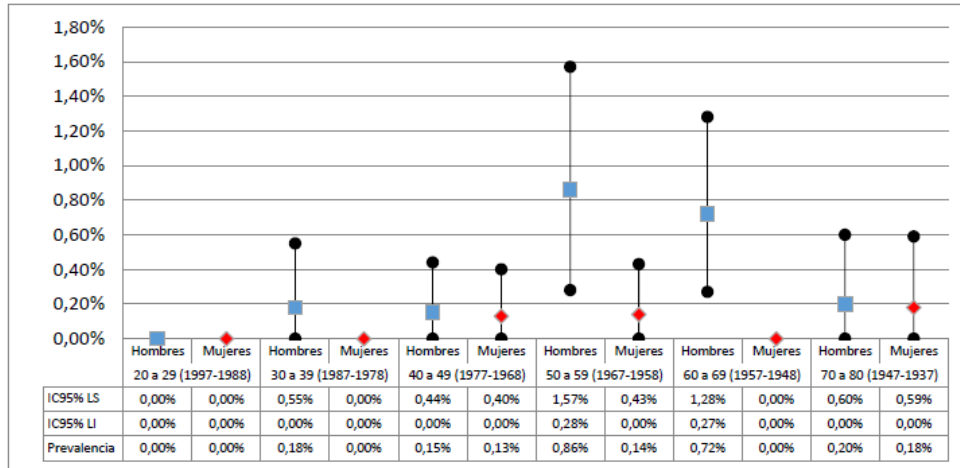


Figura 3. Prevalencia de Ac por grupo de edad y sexo de 20 a 80 años. Prevalencia de la infección por hepatitis C en población general en España: 2017-2018⁽⁸⁾.



IC 95%: intervalo de confianza al 95%; LI: límite inferior; LS: límite superior.

Figura 4. Estimación de la prevalencia de infección activa por grupo de edad y sexo de 20 a 80 años. Prevalencia de la infección por hepatitis C en población general en España: 2017-2018⁽⁸⁾.

1.3. Clínica y diagnóstico de la hepatitis C

Es excepcional que el VHC se manifieste clínicamente como una hepatitis aguda y frecuentemente pasa desapercibida de forma asintomática hasta fases más avanzadas. Este es el motivo por el que el diagnóstico en estadios precoces es difícil.

El diagnóstico de esta enfermedad se basa en pruebas serológicas, que detectan anticuerpos frente al virus. Estos anticuerpos nos informan del contacto con el virus, y permanecen siempre positivos, por lo que son útiles para detectar pacientes que tienen o han tenido hepatitis C; pero no son protectores. Sin embargo, el diagnóstico de enfermedad activa, solo lo da la determinación de la PCR (Reacción en Cadena de la Polimerasa) para la detección del ARN del virus. Esta prueba permite distinguir a los pacientes que padecen activamente la enfermedad de aquellos que se han curado. Se detecta en sangre de forma precoz, a partir de las 2 semanas de la infección, y siempre que haya ARN presente será indicativo de infección por VHC⁽¹⁰⁾.

1.4. Tratamiento de la hepatitis C

El tratamiento de la hepatitis C ha sufrido una gran revolución en los últimos años con el desarrollo de los fármacos antivirales de acción directa (AAD). Hasta entonces, la elección del tratamiento dependía de múltiples factores (genotipo viral, grado de fibrosis hepática, o enfermedades concurrentes como infección por VIH o insuficiencia renal). En la actualidad dichos factores tienen menos peso para elegir el tratamiento, dada la elevada eficacia y seguridad de estos fármacos.



Cuando un paciente se cura y elimina el virus se considera que ha alcanzado la respuesta viral sostenida, definida como la no detección de ARN en el plasma del paciente 12 semanas tras haber completado el tratamiento antiviral⁽¹¹⁾.

Inicialmente la hepatitis C fue tratada con Interferón asociada a Ribavirina. Este era un tratamiento poco eficaz (tasas de curación entre el 18 y el 45%), y tenía muchos efectos secundarios mal tolerados, así como contraindicaciones en perfiles de pacientes con enfermedad mental o adicciones. Desde el año 2015 disponemos de los AAD. Consiguen tasas de respuesta viral sostenida de más del 97%, y los de última generación son pangenotípicos y altamente eficaces en cualquier estadio de fibrosis, incluso cirróticos⁽¹²⁾, trasplantados hepáticos, o en pacientes con disfunción renal. Además, no presentan interacciones farmacológicas y son muy bien tolerados.

En varios estudios se ha visto que la adherencia a los tratamientos con antivirales de acción directa es muy alta, con valores superiores al 90%, y mejor cuantas menos pastillas hubiese que ingerir al día o menos semanas durase el tratamiento⁽¹³⁾ (**figura 5**).

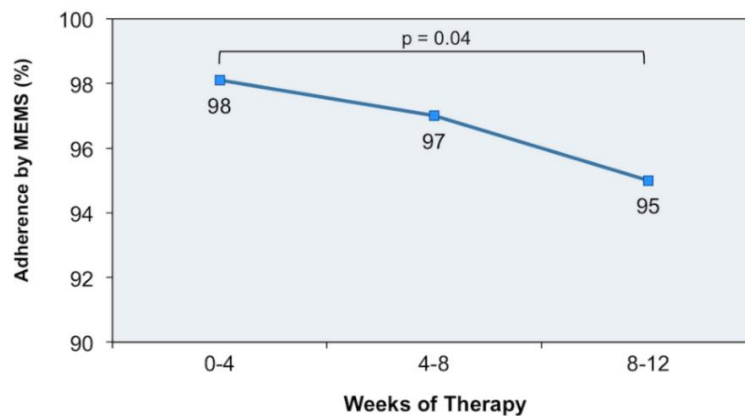


Figura 5. Evolución de la adherencia terapéutica a antivirales de acción directa durante un tratamiento de 12 semanas⁽¹⁴⁾.

Se asocia a una mayor adherencia: genotipo 1, no estar desempleado, no estar deprimido, no estar coinfectado por VIH o no ser primerizo en recibir el tratamiento. La edad, el sexo o la etnia no han demostrado alterar la adherencia⁽¹⁵⁾.

1.5. Eliminación de la hepatitis C

Lo anteriormente mencionado, muestra que es una enfermedad con gran trascendencia a nivel sanitario por las graves consecuencias para los pacientes a nivel individual, así como por el gran uso de recursos a nivel colectivo.

Nuestro país, en línea con la OMS, tiene como objetivo la eliminación de la hepatitis C para el año 2030, o al menos, diagnosticar y tratar al 90% de los pacientes VHC, consiguiendo además una disminución del 90% de la incidencia⁽¹⁶⁾.



Aunque existieron inicialmente restricciones en el acceso a los fármacos (se priorizó el acceso para fibrosis avanzadas, trasplantados hepáticos, etc.), actualmente existe un acceso total para todos los pacientes diagnosticados en nuestro país.

Por tanto, puede parecer que el problema de la hepatitis C está resuelto por completo, con fármacos muy eficaces, muy bien tolerados y sin interacciones. Y, sin embargo, nada más lejos de la realidad.

Como hemos comentado, en España el 0,85% de la población tiene anticuerpos para la hepatitis C y el 0,22% tiene la infección activa, la mayoría asintomáticos. Por este motivo es fundamental detectar a los pacientes enfermos, para poderlos tratar en fases precoces y evitar que su enfermedad progrese a estadios de cirrosis.

La AEEH recientemente ha publicado su posicionamiento sobre eliminación de la hepatitis C, recomendando la búsqueda activa de pacientes diagnosticados previamente de hepatitis C como una de las estrategias más eficientes.

En el HURH se han puesto en marcha varias iniciativas de cara a la búsqueda de pacientes diagnosticados y perdidos en el sistema, en coordinación con Atención Primaria, o una campaña de detección voluntaria de la hepatitis C. Algunas de ellas han sido implementadas a nivel de todo SACYL.

En Estados Unidos hay unos 4 millones de personas con hepatitis C; en el artículo de Terrault⁽¹⁷⁾ se centran en el gran número de pacientes que no están diagnosticados de hepatitis C aunque la padecen, o que aun estando diagnosticados no están tratados, ya que esta enfermedad presenta un gran infradiagnóstico e infratratamiento. En su estudio, informa que del total de pacientes con hepatitis C, solo un 27% de ellos obtienen un diagnóstico confirmatorio, y tan solo el 9% llegan a la curación completa. En su caso, aboga por un cribado de la enfermedad llevado a cabo por Atención Primaria.

John W. Ward⁽¹⁸⁾ por su parte, incide en la importancia del uso de test diagnósticos para hacer cribado poblacional, así como de la educación de la población sobre esta enfermedad. Una vacuna contra la hepatitis C sería clave para conseguir su eliminación; hasta entonces, hay que potenciar educación y screening. A pesar de los costes del tratamiento, su alta eficacia permite frenar la evolución de la enfermedad, evitando los gastos sobrevenidos posteriormente por las complicaciones de la enfermedad.

En un estudio en Alemania, analizaron mediante inteligencia artificial y redes neuronales a 2544 pacientes con hepatitis C entre 1,819,646 de pacientes vistos en los hospitales de una compañía aseguradora⁽¹⁹⁾. Se observó que los pacientes con hepatitis C asocian diagnósticos como “cirrosis” (68%), “hepatitis C crónica” (66%) o “hepatocarcinoma” (48%). Además, intentaron crear un modelo predictivo para captar



pacientes con un riesgo alto de padecer hepatitis C usando estas redes neuronales, para poder así llevar a cabo un cribado, aunque no pudieron validar este modelo.

En otro estudio⁽²⁰⁾, también en Estados Unidos, se apuesta también por sistemas de vigilancia y búsqueda activa de pacientes según factores de riesgo, ya que muchos de ellos forman parte de grupos marginales (como drogadictos).

Para conseguir erradicar la enfermedad, es necesario conocer mejor su epidemiología. Según Juan Berenguer⁽²¹⁾, en España se han hecho varios estudios de prevalencia de enfermedad, obteniendo resultados variados, lo que dificulta conocer la situación de la hepatitis C en España, aunque esto ha quedado más solventado con la encuesta de seroprevalencia del Ministerio de Sanidad⁽⁸⁾.

Nuestro objetivo es continuar la búsqueda de pacientes con hepatitis C utilizando una herramienta de inteligencia artificial disponible en el Hospital U. Río Hortega.

1.6. Inteligencia artificial y Big Data

Los hospitales son fuente importante de datos, aunque su análisis es escaso por la dificultad de acceder a ellos. Actualmente, con la integración de la historia clínica electrónica (HCE), es una oportunidad para evaluar estos datos, detectar asociaciones e incluso realizar modelos predictivos para ayudar a mejorar la atención al paciente.

Savana© es una herramienta de inteligencia artificial (Big Data), que analiza el lenguaje natural de la HCE. Puede detectar pacientes según diagnóstico, analíticas, identificar relaciones desconocidas; podría ayudar a tomar decisiones clínicas basadas en resultados obtenidos previamente en otros pacientes o por otros profesionales⁽²²⁾.

Con la inteligencia artificial surge la duda sobre la privacidad de los datos, y en este caso datos importantes como los incluidos en la historia clínica. Es fundamental que los datos sean anonimizados para preservar los derechos de los pacientes y al tiempo, obtener información útil para mejorar su salud. Los datos deben ser propiedad del centro hospitalario que los custodia, y las empresas de Big Data nunca manejan datos individuales de pacientes, sino que recoge únicamente la información clínica agregada.

El uso del Big Data en el sector sanitario está en pleno desarrollo, ya que hasta hace unos 10 años era inexistente, ahora realmente se está empezando a dar uso a esta herramienta tan potente y con tantas utilidades. El desarrollo en tecnología durante los últimos años ha permitido un aumento de publicaciones científicas sanitarias relacionadas con Big Data, muchas dirigidas a estudios genéticos o microbiológicos. Una utilidad más podría ser la búsqueda de pacientes con hepatitis C, o incluso de pacientes con factores de riesgo que sugieran que pueden tener hepatitis C.



2. Objetivos

2.1. Objetivos principales del estudio

- Valorar la utilidad de un sistema de análisis masivo de datos en la búsqueda de pacientes diagnosticados de hepatitis C sin control en una Unidad especializada (U. Hepatología, Sv. Medicina Interna) y por tanto sin acceso al tratamiento.

2.2. Objetivos secundarios

- Detectar pacientes VHC positivos en los pacientes atendidos en el HURH.
- Valorar la fiabilidad de Savana, evaluando la correlación entre los datos obtenidos en Savana y los encontrados tras revisar la historia clínica electrónica.
- Recuperar para el tratamiento a pacientes con el diagnóstico de hepatitis C sin acceso al tratamiento antiviral.

3. Material y métodos

Estudio observacional retrospectivo en el que se realiza una búsqueda activa de pacientes con hepatitis C entre los pacientes atendidos en el hospital Universitario Río Hortega (HURH, Valladolid) desde el 01/01/2013 hasta el 31/12/2018 no seguidos por una Unidad especializada en hepatitis C (U. de Hepatología, Sv. de Medicina Interna). Para ello utilizaremos un sistema de análisis masivo de Big Data llamado Savana®.

3.1. Selección de candidatos

Tras la descarga de todos los datos de la historia clínica electrónica del HURH (SI Clínica), en Savana, se realizó una búsqueda con el término "hepatitis C" desde el 01/01/2013 hasta el 31/12/2018 de manera global. Posteriormente dicha búsqueda se realizó en cada uno de los servicios hospitalarios.

Se identificaron a todos los pacientes que según Savana estaban diagnosticados de hepatitis C y sin embargo no eran seguidos en Unidades especializadas (Unidad de Hepatología o Servicio de Medicina Interna), y, por tanto, probablemente sin tratamiento.

Tras identificar dichos pacientes, en una segunda parte del estudio se revisó la historia clínica de todos los pacientes con hepatitis C identificados, para comprobar si efectivamente el paciente tenía una hepatitis C y si estaba siendo o no seguido por una Unidad especializada y/o tratado para la infección.

Por último, se contactaría con los pacientes por carta para comunicarles la detección de una enfermedad hepática con posibilidad de tratamiento en la actualidad. Al tratarse de una enfermedad con repercusión en la salud pública, esta acción está permitida por



la ley de autonomía del paciente⁽²³⁾ y la ley de protección de datos⁽²⁴⁾. Esta parte final, quedo truncada debido a la pandemia por el SARS-CoV-2.

3.2. Recogida de datos clínicos

Primero se realizará en Savana la búsqueda con el diagnóstico: hepatitis C desde enero 2013 hasta diciembre 2018.

Se realizará la búsqueda con el mismo término, por servicio hospitalario.

Se excluyeron los pacientes con hepatitis C que estaban siendo controlados en la Unidad de Hepatología y del Servicio de Medicina interna.

Se revisará la historia de cada paciente para confirmar la sospecha de VHC.

Se seleccionarán los pacientes vivos de nuestro área sanitaria en los que conste la sospecha de hepatitis C, para localizarles, confirmar VHC y proponer tratamiento.

Todos estos datos serán almacenados en una base de datos anonimizados, donde se recogerán los datos clínicos y analíticos correspondientes para su posterior análisis.

3.3. Variables a estudio

- Edad
- Sexo
- Diagnóstico en el servicio en que se le sigue
- GOT en la primera analítica en el servicio que es seguido
- GPT en la primera analítica en el servicio que es seguido
- Bilirrubina total en la primera analítica en el servicio que es seguido
- INR en la primera analítica en el servicio que es seguido
- Anticuerpos anti-VHC en la primera analítica en el servicio que es seguido
- PCR VHC en la primera analítica en el servicio que es seguido

3.4. Búsqueda bibliográfica

La búsqueda bibliográfica se ha llevado a cabo en la base de datos MEDLINE a través de PubMed, recurso desarrollado y mantenido por el Centro Nacional de Información Biotecnológica (National Center for Biotechnology Information, NCBI), de la Biblioteca Nacional de Medicina (National Library of Medicine, NLM) de los EE.UU, asentada en los Institutos Nacionales de Salud (National Institutes of Health, NIH).

Para la búsqueda inicial utilizamos términos libres, y la definitiva a través de términos controlados tomados del tesoro de MEDLINE, la base de datos MeSH. Los términos de búsqueda fueron relacionados mediante operadores lógicos e implementados mediante truncadores, utilizando como límites el tiempo (15 últimos años), el idioma



(español e inglés) y las condiciones “investigación en humanos” y “artículos con abstract”, para confeccionar la estrategia más adecuada. Esta estrategia se planteó con formulación PICO (Patients, Intervention, Control, Outcome), adaptada al diseño del estudio.

Entre los artículos recuperados, se seleccionaron los que aportaban la información más relevante y pertinente según criterio personal. Los límites temporales no fueron considerados para trabajos históricos de referencia. La revisión bibliográfica ha sido realizada en dos tiempos: antes del diseño del estudio, para obtener una idea del estado actual del tema, y con posterioridad al mismo, con el fin de priorizar los datos obtenidos y establecer modelos comparativos.

3.5. Aspectos éticos

Este trabajo se realizará siguiendo las recomendaciones de la Declaración de Helsinki de 1964 (última enmienda, 2013). Se cuenta con la aprobación de la Comisión de Investigación y el Comité de Ética Asistencial y Ensayos Clínicos (CEIC) del Hospital Universitario Río Hortega-Área Oeste de Valladolid.

No precisa consentimiento informado ya que solo precisa la revisión de historias clínicas.

No existe conflicto de intereses ni se dispone de financiación por organismos públicos o privados.

3.6. Análisis estadístico

La normalidad de las variables cuantitativas será establecida con la prueba de Kolmogorov-Smirnov. Las variables de distribución normal serán descritas como media \pm desviación estándar (DE). Por su parte, las variables cualitativas serán descritas mediante la tabla de frecuencias absolutas y relativas (porcentajes) de sus categorías.

Para estudiar la asociación entre variables cualitativas se utilizará la prueba de Chi cuadrado con test exacto de Fisher (tablas 2x2) o razón de verosimilitud (más de 2 categorías), dependiendo de sus condiciones de aplicación (frecuencia esperada <5 en más del 20% de las casillas).

Para estudiar las diferencias entre medias independientes y dependientes se utilizará la prueba de la T de Student o la U de Mann-Whitney, para muestras independientes y/o pareadas dependiendo de las condiciones de aplicación (normalidad), para 2 grupos, y el ANOVA o la H de Kruskal-Wallis, dependiendo asimismo de las condiciones de aplicación (normalidad), para más de 2 grupos.

El nivel de significación para todas las pruebas se considerará para una $p \leq 0,05$.



4. Resultados

4.1. Resultados generales

En una primera búsqueda en Savana con el término “hepatitis C” se encontraron 1535 pacientes (**figura 6**). De ellos, 398 no eran seguidos por una U. especializada (Hepatología o Medicina Interna), y por tanto cumplían los criterios de inclusión.

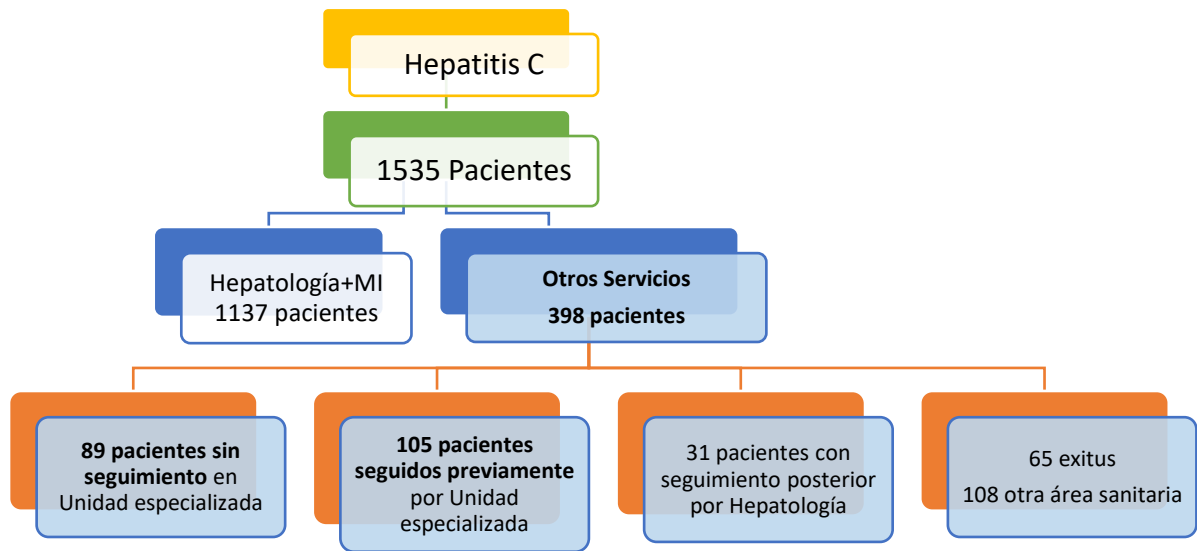


Figura 6. Pacientes obtenidos de Savana y tras la revisión de sus historias clínicas. Se entiende por Unidad especializada: Hepatología y Medicina Interna.

En este grupo de pacientes no controlados por una U. especializada según Savana, los varones eran predominantes, representando un 63% (250 de 398) de todos los pacientes (**figura 7**). La edad media de los todos los pacientes revisados fue de 61,78 años \pm 14,35 años (DE) (**figura 8**).

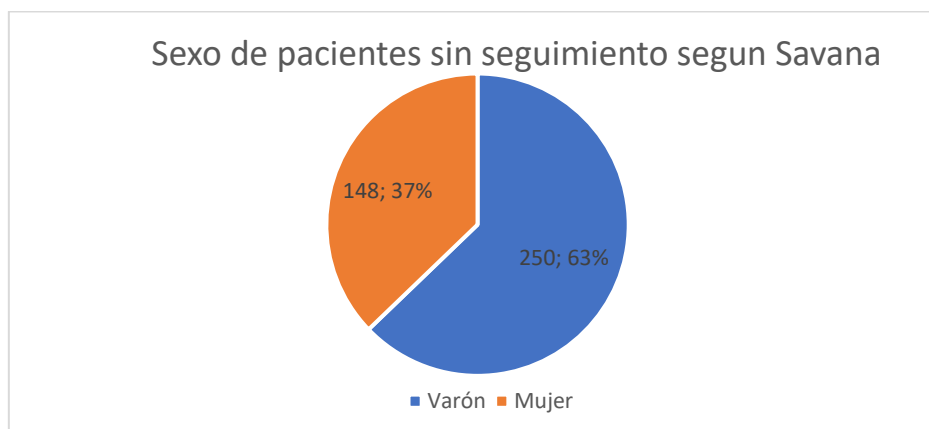


Figura 7. Distribución por sexo de los pacientes VHC + encontrados en Savana sin seguimiento por la Unidad de Hepatología o la Unidad de Medicina Interna.

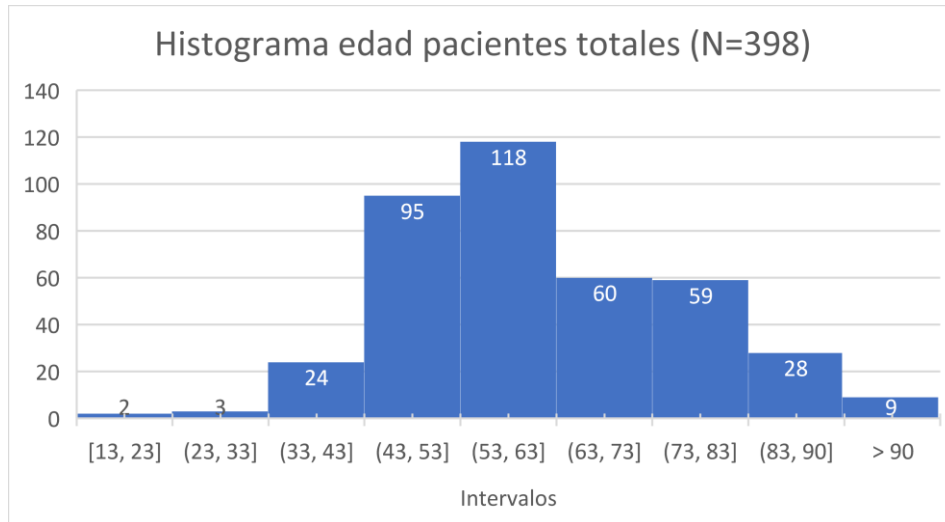


Figura 8. Histograma de frecuencias de la edad de los pacientes VHC+ encontrados en Savana sin seguimiento por la Unidad de Hepatología o la Unidad de Medicina Interna.

De estos 398 sin seguimiento en Unidades especializadas según Savana la mayoría eran vistos por los servicios de Gastroenterología (51 pacientes), Nefrología (49), Cirugía General (34), Neurología (27) y Traumatología (25). Posteriormente se analizará la distribución por servicios.

Tras la revisión de las 398 historias clínicas electrónicas, se observaron 89 pacientes (22,4%) que no eran seguidos por una Unidad especializada (**figura 9**).

En los 398 pacientes, se detectaron además 105 (26,4%) que sí estaban siendo controlados previamente por una U. especializada por su hepatitis C y que Savana no había registrado correctamente el servicio al que pertenecían, clasificándolos como no seguidos por Hepatología cuando en realidad algunos sí que lo eran.

Por otro lado, 31 (7,8%) se comprobó que fueron recuperados para seguimiento por una U. especializada posteriormente; por lo que incluyendo los 105 que eran vistos con anterioridad, 136 pacientes (34,2% de los 398) fueron vistos por una Unidad especializada en algún momento.

En cuanto a los 173 pacientes restantes detectados por Savana, 65 pacientes habían fallecido (exitus) y 108 pacientes pertenecían a otra área sanitaria (52 a Valladolid Este y 56 otras áreas sanitarias).

Dentro de los 89 pacientes no seguidos por una U. especializada, la mayoría estaban en los servicios de Nefrología (24), Neurología (10), Traumatología (9) y Cirugía General (9) (**figura 9**).



Figura 9. Número de pacientes no seguidos por una Unidad especializada el según servicio por el que sí tiene seguimiento.

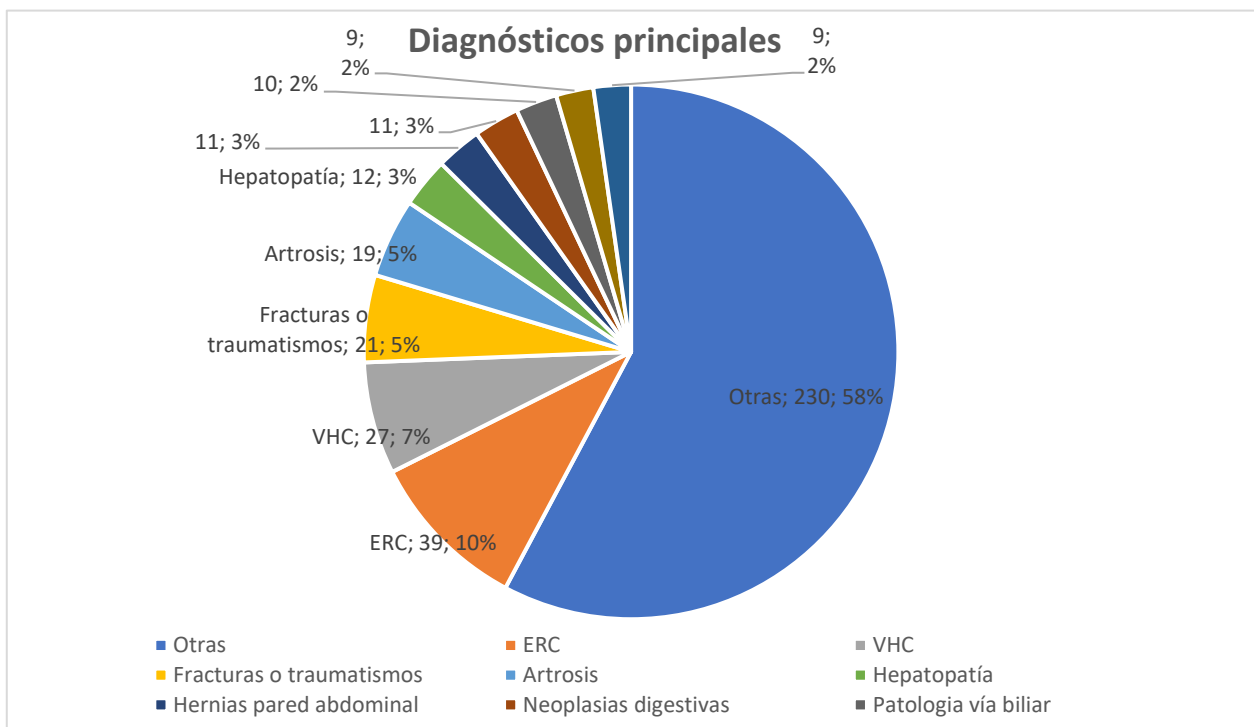


Figura 10. Diagnósticos principales de todos los pacientes revisados.

Por otro lado, se analizó el diagnóstico por el que eran seguidos (**figura 10**) en los servicios respectivos. El diagnóstico más frecuente fue la Enfermedad Renal Crónica (ERC) en 39 pacientes (10%). Seguido de 27 pacientes (7%) con el diagnóstico de infección por VHC, y 21 pacientes (7%) seguidos por alguna fractura o traumatismo.



En cuanto a los estudios específicos para estudiar la hepatitis C en todos los pacientes revisados:

- **Anticuerpos anti-VHC:** tan solo 56 (14,5%) pacientes tenían realizada la determinación de anticuerpos VHC, teniendo positivo 44 (78,6%), mientras que 10 de ellos tenían un resultado negativo (17,9%), 2 de ellos indeterminado (3,5%), y los 342 (85,5%) restantes no tenían este dato registrado en su primera analítica en el servicio en el que era seguido.
- **PCR:** solo 32 de los 398 (8%) pacientes contaban con la determinación de PCR. En 16 el resultado era positivo (50%), y otros 16 con un resultado negativo (50%), mientras que los 366 (92%) restantes no tenían este dato registrado en su primera analítica en el servicio en el que era seguido.

También se recogieron los datos analíticos de estos pacientes, centrándonos en valores de GOT, GPT, INR y bilirrubina total, siempre en la primera analítica obtenida en el servicio de seguimiento (**tabla 1**).

	Media(±DS)	% pacientes con datos	Valores de referencia	p valor
GOT (U/L)	53,37±110,36	178 (44,7%)	12-38	p<0,05
GPT (U/L)	60,82±115,35	209 (52,5%)	7-41	p<0,05
INR	1,08±0,26	133 (33,4%)	0,9-1,1	N.S
Bilirrubina total (mg/dL)	1±1,29	168 (44,2%)	0,1-1,2	N.S

Tabla 1. Valores analíticos de todos los pacientes revisados.

El 44,7% de los pacientes tenían GOT en la primera visita, el 52,5% disponían de GPT, en el 33,4% había INR y en un 44,2% disponíamos de bilirrubina total.

Se observaron diferencias estadísticamente significativas en los valores de GOT y GPT respecto a los valores de referencia (**tabla 1**). Sin embargo, no se encontraron estas diferencias en las cifras de INR y BRT.

4.2. Resultados por servicios

En total, hasta 136 pacientes fueron vistos por una U. especializada antes o después de ser vistos en estos otros servicios. Los servicios de Gastroenterología y Cirugía General fueron los que derivaron una proporción mayor de pacientes (**tabla 2**), a parte de la U. de trasplante. Los 262 restantes son exitus (65), pertenecientes a otra área sanitaria (108), o pacientes sin seguimiento por una U. especializada (89).



ESPECIALIDAD	CONTROL POSTERIOR		ESPECIALIDAD	CONTROL POSTERIOR	
	Si	No		Si	No
Alergias	3	19	Neurología	7	20
Anestesia	4	6	Oncología	1	6
Cardiología	7	10	ORL	4	2
Cirugía general	12	22	Pediatría	1	1
Cx maxilofacial	0	17	Psiquiatría	5	15
Cx plástica	1	19	Rehabilitación	0	1
Endocrino	3	10	Reumatología	3	13
Gastroenterología	42	9	Traumatología	4	21
UCI	1	6	Hopitalización domiciliaria	1	1
Nefrología	8	41	Unidad de trasplante	15	0
Neumología	9	12	Urología	4	5
Neurocirugía	1	6	TOTAL	136	262
			398		

Tabla 2. Número de pacientes sí seguidos o no seguidos posteriormente por una U. especializada.

Posteriormente, se llevó a cabo un análisis diferenciado por cada servicio hospitalario, analizando los mismos parámetros, que podemos ver en las **tablas 3 y 4**.

En la **tabla 3** observamos que tan solo 56 pacientes de los 398 pacientes (14,1%) tenían una determinación de anticuerpos VHC; y solo 32 de los 398 pacientes (8%) contaban con la determinación de PCR, de los que en 16 el resultado era positivo (50%).

Algunos servicios como Gastroenterología, Nefrología o Psiquiatría presentan un porcentaje mayor de anticuerpos VHC o PCR-VHC.

En lo que se refiere a los datos analíticos (**tabla 4**), el 44,7% de los pacientes tenían GOT en la primera visita, el 52,5% disponían de GPT, en el 33,4% había INR y en un 44,2% disponíamos de bilirrubina total.

Además, también podemos observar una gran variabilidad entre los diferentes servicios (**tabla 4**), destacando Cirugía Plástica, Gastroenterología, Psiquiatría o la Unidad de trasplante como los servicios con datos superiores a los valores de referencia, y destacando los servicios de Gastroenterología, Unidad de trasplante, Oncología y Reumatología como los que mayor porcentaje de datos tenían recogidos.



ESPECIALIDAD	SEXO		EDAD		AntiVHC				PCR		
	Varones	Mujeres	Media	Desv estandar	+	-	Indeterminado	No registrado	+	-	No registrado
Alergias	10	12	61,27	10,13	1	0	1	20	0	0	22
Anestesia	3	7	63,80	16,72	0	0	0	10	0	0	10
Cardiología	9	8	64,82	16,25	0	0	0	17	0	0	17
Cirugía general	29	5	60,79	10,48	3	0	0	31	2	0	32
Cx maxilofacial	13	4	54,53	10,79	3	0	0	14	1	0	16
Cx plástica	15	5	56,33	13,41	2	0	0	18	0	0	20
Endocrino	6	7	56,77	14,91	0	0	0	13	0	0	13
Gastroenterología	26	25	62,41	14,85	10	1	0	40	9	11	31
UCI	4	3	61,14	15,85	0	0	0	7	0	0	7
Nefrología	30	19	67,14	15,84	6	7	0	36	0	0	49
Neumología	16	5	66,86	16,32	2	0	0	19	2	0	19
Neurocirugía	4	3	65,14	9,26	0	0	0	7	0	0	7
Neurología	17	10	63,19	13,21	2	1	0	24	1	0	26
Oncología	4	3	65,57	14,34	0	0	0	7	0	0	7
ORL	4	2	57,17	3,92	0	0	0	6	0	0	6
Pediatría	1	1	15,00	2,83	1	0	0	1	0	0	2
Psiquiatría	16	4	53,25	10,38	6	0	0	14	1	2	17
Rehabilitación	1	0	50,00	0,00	0	0	0	1	0	0	1
Reumatología	4	12	58,19	14,36	3	0	1	12	0	1	15
Traumatología	17	8	66,44	16,98	4	1	0	20	0	1	24
Hopitalización domiciliar	1	1	63,33	7,23	0	0	0	2	0	0	2
Unidad de trasplante	11	4	61,27	6,80	0	0	0	15	0	1	14
Urología	9	0	65,22	9,67	1	0	0	8	0	0	9
TOTAL	398	250	148		44	10	2	342	16	16	366

Tabla 3. Sexo, edad, Ac anti-VHC y PCR según cada servicio.

ESPECIALIDAD	GOT			GPT			INR			BRT		
	Media	DE	No registrados	Media	DE	No registrados	Media	DE	No registrados	Media	DE	No registrados
Alergias	44,43	19,88	16 (72,7%)	39,48	26,10	16 (72,7%)			22 (100%)	0,94	0,66	16 (72,7%)
Anestesia			10 (100%)			10 (100%)			10 (100%)			10 (100%)
Cardiología	33,15	15,77	13 (76,5%)	33,03	18,84	13 (76,5%)	1,07	0,10	12 (70,6%)	0,78	0,23	14 (82,3%)
Cirugía general	34,39	22,05	20 (58,8%)	34,89	26,03	11 (32,3%)	1,04	0,14	7 (20,6%)	0,82	0,47	20 (58,8%)
Cx maxilofacial	34,15	7,14	15 (88,2%)	42,90	11,91	13 (76,5%)	1,09	0,08	12 (70,6%)	1,10	1,04	15 (88,2%)
Cx plástica	207,40	126,43	18 (90%)	69,23	94,48	14 (70%)	1,00	0,13	11 (55%)	1,36	0,00	19 (95%)
Endocrino	26,35	8,94	5 (38,5%)	30,11	19,89	5 (38,5%)	1,01	0,00	12 (92,3%)	0,78	0,35	6 (46,1%)
Gastroenterología	90,33	210,73	10 (19,6%)	101,79	192,82	10 (19,6%)	1,12	0,28	22 (43,1%)	1,41	1,81	10 (19,6%)
UCI	60,80	0,00	6 (85,7%)	58,30	0,00	6 (85,7%)	0,85	0,00	6 (85,7%)	0,93	0,00	6 (85,7%)
Nefrología	22,37	12,82	23 (46,9%)	18,70	9,82	23 (46,9%)	0,93	0,12	43 (87,7%)	0,73	0,36	26 (53%)
Neumología	34,55	18,89	10 (47,6%)	31,84	19,86	10 (47,6%)	1,37	0,66	13 (61,2%)	0,62	0,26	11 (52,4%)
Neurocirugía			7 (100%)	23,00	0,00	6 (85,7%)	0,98	0,05	4 (57,1%)			7 (100%)
Neurología	33,89	23,76	13 (48,1%)	34,20	32,92	13 (48,1%)	0,98	0,08	23 (85,2%)	0,72	0,23	13 (48,1%)
Oncología	32,78	11,47	1 (14,3%)	22,20	7,86	1 (14,3%)	0,98	0,00	6 (85,7%)	0,40	0,08	1 (14,3%)
ORL	34,90	0,00	5 (83,3%)	53,77	8,79	3 (50%)	0,89	0,05	3 (50%)	0,74	0,00	5 (83,3%)
Pediatría	27,6	0,00	1 (50%)	44,2	0,00	1 (50%)	0,91	0,00	1 (50%)	0,35	0,00	1 (50%)
Psiquiatría	61,90	47,51	12 (60%)	45,75	37,29	12 (60%)	1,47	0,00	19 (95%)	0,73	0,51	13 (65%)
Rehabilitación			1 (100%)			1 (100%)			1 (100%)			1 (100%)
Reumatología	25,15	11,42	5 (31,2%)	20,24	8,99	5 (31,2%)			16 (100%)	0,62	0,32	5 (31,2%)
Traumatología	25,55	4,88	23 (92%)	41,94	48,88	13 (52%)	1,00	0,09	12 (48%)	0,32	0,00	23 (92%)
Hopitalización domiciliar			2 (100%)			2 (100%)			2 (100%)			2 (100%)
Unidad de trasplante	98,49	93,05	0 (0%)	226,45	188,72	0 (0%)	1,22	0,27	2 (13,3%)	2,02	2,74	1 (6,7%)
Urología	33,00	18,22	4 (44,4%)	33,05	27,54	1 (11,1%)	0,78	0,22	6 (66,6%)	0,68	0,28	5 (55,5%)
TOTAL	398		220 (55,3%)			189 (47,5%)			265 (66,6%)			230 (57,8%)

Tabla 4. GOT (U/L), GPT (U/L), INR y bilirrubina total (BRT, mg/dL) según cada servicio.



4.3. Resultados en pacientes sin seguimiento por U. especializada

Dentro de los pacientes no seguidos por Hepatología o Medicina Interna, nos encontramos que, de los 89 pacientes, 55 son varones (61,8%) y las 34 restantes mujeres (38,2%) (**figura 11**) y una edad media de $62,27 \pm 16,07$ años (**figura 12**), sin diferencias significativas respecto al resto de los grupos ($p > 0,05$).

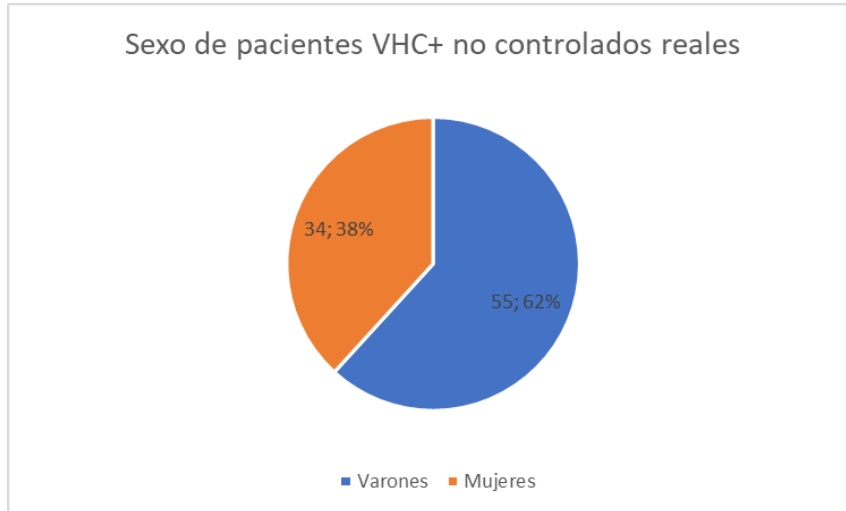


Figura 11. Distribución por sexo de los pacientes VHC+ realmente no seguidos por Unidad especializada.

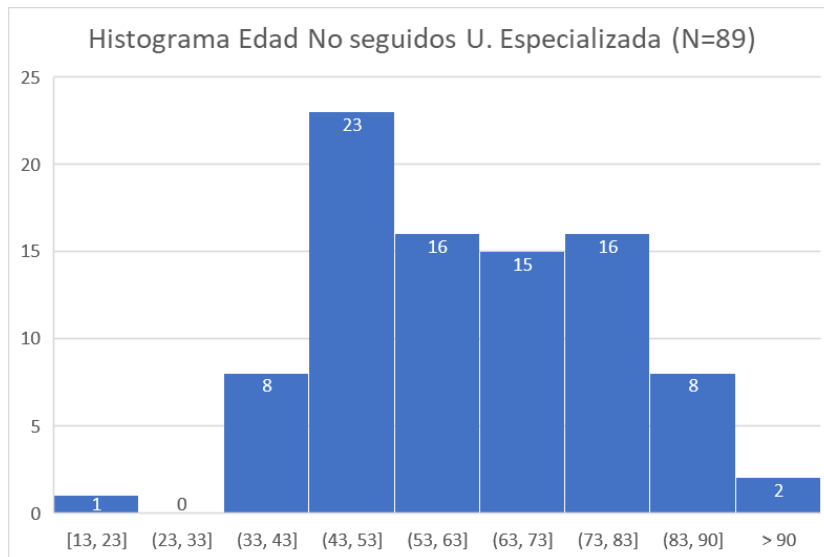


Figura 12. Histograma de frecuencias de la edad de los pacientes VHC+ realmente no seguidos por Unidad especializada.

En cuanto a los estudios específicos para estudiar la hepatitis C (**tabla 5**):

- **Anticuerpos anti-VHC:** 16 pacientes tenían el dato recogido (18%): 10 pacientes presentaban un resultado positivo (63%), 5 pacientes un resultado negativo



(31%), 1 un resultado indeterminado (6%), y los 73 restantes (82%) no tenían este dato registrado en su primera analítica en el servicio en el que eran seguidos.

- **PCR:** solo 4 pacientes tenían el dato recogido (4,5%): 3 pacientes contaban con un resultado positivo para esta prueba (75%), 1 paciente con un resultado negativo (25%) y los 85 restantes (95,5%) no tenían este dato registrado en su primera analítica en el servicio en el que eran seguidos.

	Anti- VHC				PCR		
	Positivo	Negativo	Indeterm	No registrado	Positivo	Negativo	No registrado
General	44 (78,5%)	10 (18%)	2 (3,5%)	342 (85,5%)	16 (50%)	16 (50%)	366 (92%)
NO seguidos	10 (63%)	5 (31%)	1 (6%)	73 (82,1%)	3 (75%)	1 (25%)	85 (95,5%)

Tabla 5. Resultados anticuerpos anti-VHC y PCR en pacientes totales y pacientes no seguidos.

En cuanto a la media de transaminasas, podemos observarlos en la siguiente tabla (**tabla 6**):

	GOT (U/L)	GPT (U/L)	INR	BRT (mg/dL)
	Media±DS	Media±DS	Media±DS	Media±DS
Total Pacientes	53,37±110,36	60,82±115,35	1,08±0,26	1±1,29
Pacientes sin seguimiento	35,01±24,21	33,49±26,18	0,99±0,11	0,65±0,35
P	<0.05	<0.05	N.S.	N.S.

Tabla 6. Resultados analíticos de GOT, GPT, INR y bilirrubina total (BRT) en pacientes totales y en los no seguidos.

Los pacientes no seguidos por un servicio especializado tienen valores significativamente inferiores ($p < 0.05$) que los pacientes totales, y menor variabilidad (desviaciones estándar inferiores), aunque muchos pacientes no tenían datos a valorar.

4.4. Resultados en pacientes seguidos previamente por U. especializada

Dentro de los pacientes que sí eran seguidos previamente por el servicio de Hepatología o Medicina Interna (Unidades especializadas) del HURH por su hepatitis C, pero que Savana los había clasificado como no seguidos, nos encontramos 105 pacientes, 64 varones (61%) y 41 mujeres (39%) (**figura 13**). La edad media es algo



inferior, aunque no estadísticamente significativo, a los grupos anteriores, con una media de 60,58 años \pm 12,87 (**figura 14**).

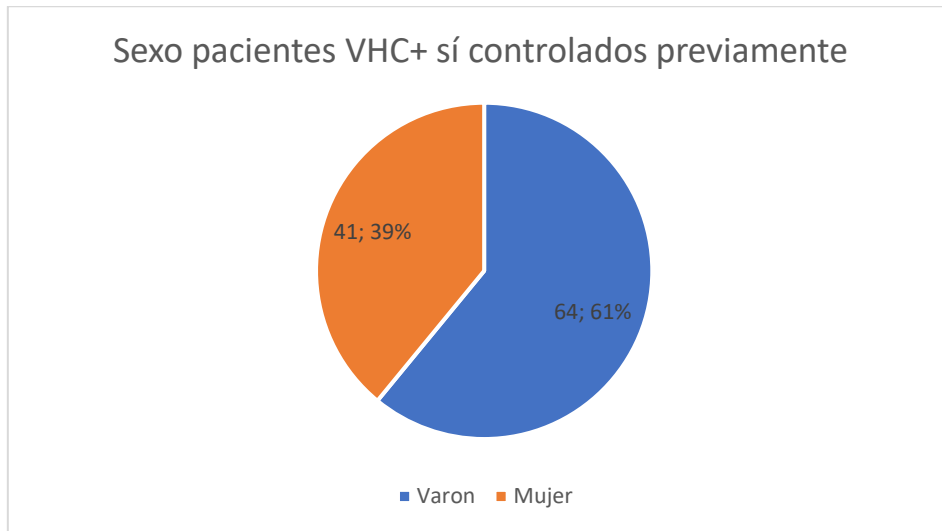


Figura 13. Distribución por sexo de los pacientes VHC+ realmente sí seguidos previamente por Unidad especializada.

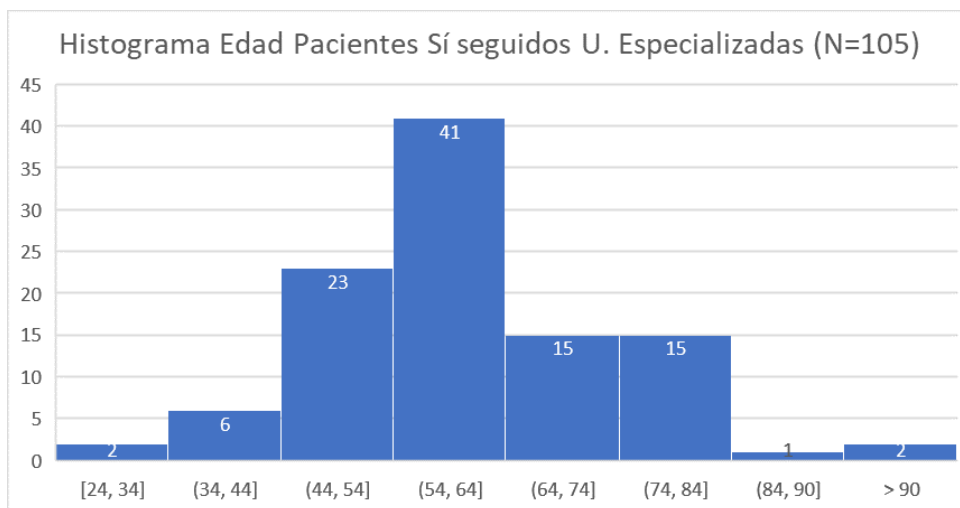


Figura 14. Histograma de frecuencias de la edad de los pacientes VHC+ realmente sí seguidos previamente por Unidad especializada.

En cuanto a los estudios específicos en este grupo de pacientes (**tabla 7**):

- Anticuerpos anti-VHC: 13 pacientes tenían el dato recogido (12,4%): 12 pacientes obtuvieron un resultado positivo en la prueba (92,3%), mientras que solo 1 de ellos obtuvo un resultado negativo (7,7%); los 92 restantes (87,6%) no tenían este parámetro registrado en su primera analítica en el servicio en el que eran seguidos.



- **PCR:** 4 pacientes tenían el dato recogido (4%): 2 pacientes tenían un resultado positivo (50%), 2 un resultado negativo (50%), mientras que los 101 restantes (96%) no tenían este parámetro registrado en su primera analítica en el servicio en el que eran seguidos.

	Anti- VHC				PCR		
	Positivo	Negativo	Indeterm	No registrado	Positivo	Negativo	No registrado
General	44 (78,5%)	10 (18%)	2 (3,5%)	342 (85,5%)	16 (50%)	16 (50%)	366 (92%)
NO seguidos	10 (63%)	5 (31%)	1 (6%)	73 (82,1%)	3 (75%)	1 (25%)	85 (95,5%)
SÍ seguidos	12 (92,3%)	1 (7,7%)	0 (0%)	92 (87,5%)	2 (50%)	2 (50%)	101 (96%)

Tabla 7. Resultados anticuerpos anti-VHC y PCR en pacientes en general, pacientes no seguidos y pacientes sí seguidos previamente por una U. especializada.

En cuanto a los resultados analíticos (**tabla 8**) en este grupo de pacientes sí seguidos previamente, los valores de todos los parámetros van a alcanzar valores superiores a los de referencia, acompañados de una variabilidad aumentada.

	GOT (U/L)	GPT (U/L)	INR	BRT (mg/dL)
	Media±DE	Media±DE	Media±DE	Media±DE
Pacientes total	53,37±110,36	60,82±115,35	1,08±0,26	1±1,29
NO seguidos	35,01±24,21	33,49±26,18	0,99±0,11	0,65±0,35
SÍ seguidos	77,66±185,12	104,40±184,98	1,16±0,40	1,24±1,76

Tabla 8. Resultados de GOT, GPT, INR y bilirrubina total (BRT) en pacientes totales, en los no seguidos y en los sí seguidos previamente por U. especializada

Como vimos en la **tabla 1**, los porcentajes de datos no registrados en los pacientes en general son del 55,3% para la GOT, 47,5% la GPT, 66,6% el INR y 57,8% la bilirrubina total. En la siguiente **tabla 9** vamos a comparar estos porcentajes con los pacientes no seguidos y sí seguidos previamente.

	GOT	GPT	INR	BRT
General	220 (55,3%)	189 (47,5%)	265 (66,6%)	230 (57,8%)
NO seguidos	52 (58,4%)	37 (41,6%)	61 (68,5%)	54 (60,7%)
SÍ seguidos	41 (39%)	42 (40%)	65 (61,9%)	56 (53,3%)

Tabla 9. Número de pacientes y porcentaje sin dichos parámetros recogidos en su primera analítica en el servicio que les seguía, en pacientes totales, pacientes no seguidos y pacientes sí seguidos previamente por una U. especializada.



5. Discusión

La relevancia de la hepatitis C es grande, relacionada con el riesgo de desarrollo de cirrosis hepática, hepatocarcinoma y la muerte del paciente. En España, hay unos 4.800 fallecimientos anuales por VHC⁽⁹⁾. Según la encuesta de seroprevalencia del Ministerio Sanidad⁽⁸⁾, en nuestro país hay una prevalencia de Ac anti-VHC del 0,85% (IC 95%: 0,64%-1,08%), con una prevalencia de RNA+ de 0,22% (IC 95%: 0,12%-0,32%). La mayoría son asintomáticos, de ahí la importancia de diagnosticar y tratar a estos pacientes⁽⁹⁾. El diagnóstico es sencillo: una prueba serológica para detectar Ac, y una PCR para detectar ARN y confirmar la infección actual⁽¹⁰⁾. Disponemos de tratamientos eficaces, seguros, que consiguen tasas de respuesta viral sostenida del más del 97%⁽¹¹⁾.

Por otro lado, el uso de Big Data en el sector sanitario está en pleno desarrollo a pesar de sus puntos débiles, como que son aún estudios observacionales en los que los datos pueden no haber sido introducido siempre correctamente. Un inconveniente del Big Data en medicina es la privacidad y seguridad, ya que los datos clínicos son posesión de cada paciente⁽²⁵⁾. La solución que ofrece Savana es la anonimización y el encriptamiento de estos datos; como comenté, Savana no recoge datos clínicos aislados, si no que obtiene información clínica agregada.

Los datos los recoge Savana sin la supervisión humana. Toma datos de historias clínicas electrónicas escritas por sanitarios, pero cada persona se expresa de forma diferente, lo que dificulta homogeneidad en los datos que después permita un buen análisis. Nosotros, de cara a valorar el análisis hecho por Savana, hemos revisado las historias clínicas electrónicas, para confirmar si los datos obtenidos mediante Savana son fiables y si los pacientes están correctamente clasificados. Un problema añadido es que muchas veces no se recogen todos los datos de la historia clínica⁽²⁶⁾.

Como vimos⁽¹⁶⁻²¹⁾, la hepatitis C está infradiagnosticada e infratratada. En consecuencia, con este proyecto intentamos ayudar en esa búsqueda de pacientes sin el diagnóstico ni tratamiento necesarios, para contribuir al objetivo de la OMS de eliminar la hepatitis C en el año 2030⁽¹⁶⁾.

Mientras que en los estudios mencionados por Berenguer⁽²¹⁾ buscan enfermos en los grupos de riesgo, nosotros hemos buscado pacientes diagnosticados pero sin seguimiento especializado. Para ello, aportamos el uso innovador de la inteligencia artificial, con objeto de buscar pacientes diagnosticados pero perdidos en el sistema; un paso más hacia la eliminación de la hepatitis C, y detectando en una muestra de 398 pacientes hasta 89 (22,4%) en esa situación, en comparación con Terrault⁽¹⁷⁾, que aboga por un cribado poblacional llevado a cabo por Atención Primaria.



En Savana encontramos 398 pacientes no seguidos por una U. especializada por su hepatitis C, pero Savana no había clasificado adecuadamente 105 (26,4%) pacientes, por lo que probablemente necesitará algún tipo de mejora o aprendizaje en el reconocimiento del diagnóstico o del servicio asignado al paciente.

Y aunque el diagnóstico constaba en la historia clínica, destaca el alto porcentaje de pacientes en los que no se había confirmado el diagnóstico en el servicio en el que eran seguidos (**tablas 1 y 9**). Esto indica que la enfermedad pasa desapercibida a pesar de ser atendidos por diversos especialistas. Dado que el periodo de estudio es muy amplio, se detectaron un grupo de pacientes fallecidos actualmente.

Finalmente, encontramos 89 pacientes con sospecha de hepatitis C sin control en U. especializada. Dada la aparición de la pandemia del virus SARS-CoV-2, no se pudo completar el estudio contactando con ellos, para haber conocido su estado real y haber comprobado si presentaban PCR positiva para recibir el tratamiento adecuado.

En el estudio de Reiser y cols.⁽¹⁹⁾, único que encontré en el que aplicasen la inteligencia artificial para la localización de pacientes, entrenaron su herramienta de inteligencia artificial con un grupo de pacientes enfermos de hepatitis C, para conocer sus características médicas y sociales, y sus factores de riesgo. Con ese entrenamiento intentarían mejorar la tasa de detección de pacientes gracias al “machine learning”. Así, encontraron 2217 pacientes que tenían características muy similares a los ya diagnosticados de hepatitis C, pero que no estaban diagnosticados de la enfermedad.

En nuestro caso, Savana necesita un aprendizaje o entrenamiento, para comprender mejor las historias clínicas y clasificar de forma más adecuada los pacientes, evitando la tasa de error encontrada en nuestro estudio (26,4%) o disminuyéndola. Savana también permite ver asociaciones entre enfermedades o hábitos de los pacientes (fumador o bebedor, por ejemplo), siempre que vengan reflejados en la historia clínica. Esto podría permitirnos evaluar los factores de riesgo asociados a la hepatitis C, o a otras enfermedades, y poder implantar así un cribado en pacientes que compartan esos factores de riesgo pero que no estén diagnosticados de hepatitis C.

Una variable en su estudio⁽¹⁹⁾ era la edad, asociándose mayor edad a mayor riesgo de hepatitis. En nuestro estudio, no había diferencias de edad estadísticamente significativas entre los diferentes, que apoya que sí sean enfermos de hepatitis C (si la edad media en los no seguidos fuese inferior, podría indicar que no la padecían).

No pudieron realizarles un test para evaluar si tenían hepatitis C debido a restricciones impuestas por el Comité de Ética. En nuestro caso, después sí que



podríamos haberles hecho el test diagnóstico, dado que es un problema de salud pública que justifica la necesidad de tratamiento para evitar el contagio. Sin embargo, esta parte quedó truncada por la pandemia del SARS-CoV-2 y no se pudo completar.

En el estudio de Reiser y cols.⁽¹⁹⁾ los diagnósticos más frecuentes fueron hepatitis crónica y hepatocarcinoma. En nuestra serie la Enfermedad Renal Crónica (ERC) fue el diagnóstico más frecuente, seguida de la infección por VHC (**figura 10**).

Hasta 16 servicios del HURH tenían pacientes con hepatitis C sin seguimiento especializado, aunque 4 de ellos concentraban el 58,4% (**figura 9**). Sería útil la creación de un protocolo, o un sistema de alerta, para evitar que pacientes con una posible hepatitis C siguiesen sin un control especializado. Cabe destacar, que varios servicios prácticamente no tenían datos de los pacientes, ya que eran pacientes pertenecientes a otra área sanitaria, pero acudían al HURH a servicios de referencia (Alergias o Cirugía Plástica). Por ello, tampoco tienen control posterior por una U. especializada del HURH.

Los valores de GOT y GPT en pacientes sin seguimiento son significativamente más bajos que en los pacientes totales y sí seguidos. Aun así, pacientes con hepatitis C, aunque asintomáticos y estables, deben tener un seguimiento para evaluar su situación, o la progresión de las lesiones hepáticas si las tienen. Para ello, es útil realizarles un fibroscan, para clasificar el grado de fibrosis que presentan, y decidir si necesitan un seguimiento más estrecho o si puede dárseles de alta.

La tasa de Ac positivos es superior en los pacientes en seguimiento por Unidades especializadas. Probablemente estos resultados indican que los pacientes en seguimiento estaban correctamente diagnosticados y por ello estaban en seguimiento.

Aunque hay un porcentaje alto de datos sin recoger, el subgrupo de pacientes en seguimiento previo presenta mayor proporción de datos recogidos, probablemente por el seguimiento hospitalario y el mayor número de consultas y controles por parte de los facultativos de los servicios que les siguen, al saber que son pacientes con hepatitis C.

Nuestro objetivo final era encontrar pacientes con hepatitis C sin seguimiento en U. especializadas. Encontramos 89 pacientes subsidiarios de ser localizados y de completar el estudio para valorar la necesidad de tratamiento. En 10 de ellos existían Ac anti-VHC y en 3 pacientes PCR +. En cuanto la pandemia lo permita, el objetivo será localizar a los pacientes para completar estudio y ofrecerles tratamiento.

Como limitaciones en el estudio: se trata de un estudio observacional, con las limitaciones inherentes a este tipo de estudios, que recoge datos desde 2013, por lo que parte de los pacientes habían fallecido, y habría datos sin recoger. Además, hay un alto porcentaje de parámetros no recogidos.



6. Conclusiones

- La hepatitis C es un enfermedad oligosintomática, lo que dificulta su diagnóstico.
- Se sabe que muchos pacientes diagnosticados de hepatitis C no están siendo seguidos en Unidades especializadas y no tienen acceso al tratamiento.
- Mediante una herramienta de inteligencia artificial (Savana), insertada en la historia clínica del hospital universitario Río Hortega, localizamos 89 pacientes diagnosticados de hepatitis C, sin controles en Unidades especializadas (U. de Hepatología o Servicio de Medicina Interna) y sin acceso al tratamiento.
- Sería necesario promover la creación de un protocolo o sistema de alerta para derivar pacientes con sospecha de hepatitis C hacia la Unidad de Hepatología u otra Unidad especializada, ya que hasta 16 servicios del HURH seguían pacientes con hepatitis C que no tenían un seguimiento por una Unidad especializada en enfermedades del hígado.
- El sistema de análisis masivo de datos (Big Data) Savana necesita alguna mejora o aprendizaje, con el objetivo de poder localizar y clasificar a los pacientes de una forma más fidedigna, dado que clasificó de forma incorrecta a 105 pacientes de los 398 resultados obtenidos (26,4%).
- Con el desarrollo del Big Data, habrá que tener máximo cuidado con el manejo de la privacidad y de los datos personales de los pacientes; por ello la anonimización y el encriptamiento de los datos es básico para mantener el derecho de confidencialidad.
- Hubiera sido deseable poder finalizar el estudio localizando a los pacientes y completando el estudio diagnóstico.



7. Bibliografía

1. Editorial: Non-A, non-B? *Lancet Lond Engl.* 12 de julio de 1975;2(7924):64-5.
2. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science.* 21 de abril de 1989;244(4902):359-62.
3. Harrison, Longo. *Principios de medicina interna.* 18.ª ed. Vol. II. Mexico: Mc Graw Hill; 2543 p.
4. Hepatitis C OMS [Internet]. [citado 3 de octubre de 2019]. Disponible en: <https://www.who.int/news-room/fact-sheets/detail/hepatitis-c>
5. Montserrat Chimeno M a, Chocarro Á, Brezmes P, Ochoa C, Pérez J. Seroprevalencia del virus de la hepatitis C en la población general. *Enfermedades Infecc Microbiol Clínica.* enero de 2002;20(2):64-7.
6. Encuesta de Seroprevalencia de la Comunidad Autónoma del País Vasco [Internet]. Departamento de Sanidad y Consumo. Gobierno Vasco; 2011 [citado 18 de octubre de 2019]. Disponible en: http://www.euskadi.eus/contenidos/informacion/publicaciones_departamento/es_def/a_djuntos/salud_publica/seroprevalencia.pdf
7. Solà R, Cruz de Castro E, Hombrados M, Planas R, Coll S, Jardí R, et al. Prevalencia de las hepatitis B y C en diversas comarcas de Cataluña: estudio transversal. *Med Clínica.* 22 de junio de 2002;119(3):90-5.
8. Ministerio de Sanidad, Consumo y Bienestar. Prevalencia de la infección por hepatitis C en población general en España: 2017-2018. [Internet]. 2019 [citado 24 de octubre de 2019]. Disponible en: https://www.msbs.gob.es/ciudadanos/enfLesiones/enfTransmisibles/sida/docs/INFORME_INFICCION_VHC_ESPANA2019.pdf
9. Albillos Martínez A, Cañada Merino JL, Molero García JM, Pérez Cachafeiro S, Pérez Escanilla F, Simón Marco MA, et al. Consenso de recomendaciones para el diagnóstico precoz, la prevención y la atención clínica de la hepatitis C en Atención Primaria. *Luzan;* 2017.
10. Webster DP, Klenerman P, Dusheiko GM. *Lancet Seminar – Hepatitis C.* *Lancet Lond Engl.* 21 de marzo de 2015;385(9973):1124-35.
11. Sanjiv Chopra, MD, MACP, Paul J Pockros, MD. Overview of the management of chronic hepatitis C virus infection. En: *UpToDate* [Internet]. UpToDate. Waltham, MA; 2018 [citado 4 de octubre de 2019]. Disponible en: <https://www.uptodate.com/contents/overview-of-the-management-of-chronic-hepatitis-c-virus-infection>
12. Pecoraro V, Banzi R, Cariani E, Chester J, Villa E, D'Amico R, et al. New Direct-Acting Antivirals for the Treatment of Patients With Hepatitis C Virus Infection: A Systematic Review of Randomized Controlled Trials. *J Clin Exp Hepatol.* 1 de julio de 2019;9(4):522-38.
13. Core Concepts - Addressing Adherence Prior to Initiating HCV Treatment - Evaluation and Preparation for Hepatitis C Treatment - Hepatitis C Online [Internet]. [citado 3 de noviembre de 2019]. Disponible en: [https://www.hepatitisc.uw.edu/go/evaluation-](https://www.hepatitisc.uw.edu/go/evaluation-treatment/addressing-adherence-problems/core-concept/all)
14. Petersen T, Townsend K, Gordon LA, Sidharthan S, Silk R, Nelson A, et al. High adherence to all-oral directly acting antiviral HCV therapy among an inner-city patient population in a phase 2a study. *Hepatol Int.* 26 de noviembre de 2015;10(2):310-9.
15. Younossi ZM, Stepanova M, Henry L, Nader F, Younossi Y, Hunt S. Adherence to treatment of chronic hepatitis C: from interferon containing regimens to interferon and ribavirin free regimens. *Medicine (Baltimore)* [Internet]. 18 de julio de 2016 [citado 3 de noviembre de 2019];95(28). Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4956799/>
16. WHO | Global health sector strategy on viral hepatitis 2016-2021 [Internet]. WHO. World Health Organization; [citado 25 de mayo de 2020]. Disponible en: <http://www.who.int/hepatitis/strategy2016-2021/ghss-hep/en/>
17. Terrault NA. Hepatitis C elimination: challenges with under-diagnosis and under-treatment. *F1000Research* [Internet]. 14 de enero de 2019 [citado 24 de mayo de 2020];8. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6338244/>
18. Ward JW. Global Elimination of Hepatitis C Virus. *Gastroenterol Hepatol.* octubre de 2016;12(10):632-5.
19. Reiser M, Wiebner B, Hirsch J, the German Liver Foundation. Neural-network analysis of socio-medical data to identify predictors of undiagnosed hepatitis C virus infections in Germany (DETECT). *J Transl Med.* 19 de marzo de 2019;17(1):94.
20. Saab S, Le L, Saggi S, Sundaram V, Tong MJ. Toward the elimination of hepatitis C in the United States. *Hepatol Baltim Md.* 2018;67(6):2449-59.
21. Berenguer J. Eliminación del virus de la hepatitis C en España: un reto pendiente. *Enfermedades Infecc Microbiol Clínica.* 1 de junio de 2018;36(6):323-4.
22. Savana - Transforma el texto de tus historias clínicas en Big Data [Internet]. [citado 6 de octubre de 2019]. Disponible en: <https://savanamed.com/es/>
23. Ley 41/2002, de 14 de noviembre, básica reguladora de la autonomía del paciente y de derechos y obligaciones en materia de información y documentación clínica. Jefatura del Estado «BOE» núm. 274, de 15 de noviembre de 2002 Referencia: BOE-A-2002-22188.
24. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. BOE» núm. 294, de 06/12/2018. BOE-A-2018-16673.
25. Ristevski B, Chen M. Big Data Analytics in Medicine and Healthcare. *J Integr Bioinforma* [Internet]. 10 de mayo de 2018 [citado 3 de noviembre de 2019];15(3). Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6340124/>
26. Lee CH, Yoon H-J. Medical big data: promise and challenges. *Kidney Res Clin Pract.* marzo de 2017;36(1):3-11.



8. Póster

Utilidad de un sistema de análisis masivo de datos (Big Data) insertado en la historia clínica, en la búsqueda activa de pacientes con hepatitis C



Autor: Miguel Bardají Carrillo **Tutora:** Gloria Sánchez Antolín
Grado en Medicina. Universidad de Valladolid. Servicio de Aparato Digestivo del HURH

Introducción

- **Causa** de cirrosis y hepatocarcinoma, 1ª indicación trasplante hepático.
- **Diagnóstico sencillo y tratamiento eficaz.**
- En España, prevalencia Ac+ de 0,85% y RNA + de 0,22%, la mayoría asintomáticos.
- **Objetivo OMS:** eliminación 2030.
- **Big Data** para localizar enfermos sin seguimiento. En pleno desarrollo.

Objetivos

- Evaluar la utilidad del **Big Data** en la **búsqueda de pacientes** diagnosticados de hepatitis C **sin seguimiento por Unidad especializada** (U. Hepatología o Servicio de Medicina Interna).
- Evaluar la **fiabilidad de Savana**.
- **Tratar** a los pacientes detectados

Método

Búsqueda en **Savana** de pacientes con hepatitis C sin seguimiento por Unidad especializada desde enero 2013 a diciembre 2018.

Revisión de las **historias clínicas** electrónicas de dichos pacientes.

Carta a pacientes diagnosticados de hepatitis C sin seguimiento por Unidad especializada.

Resultados

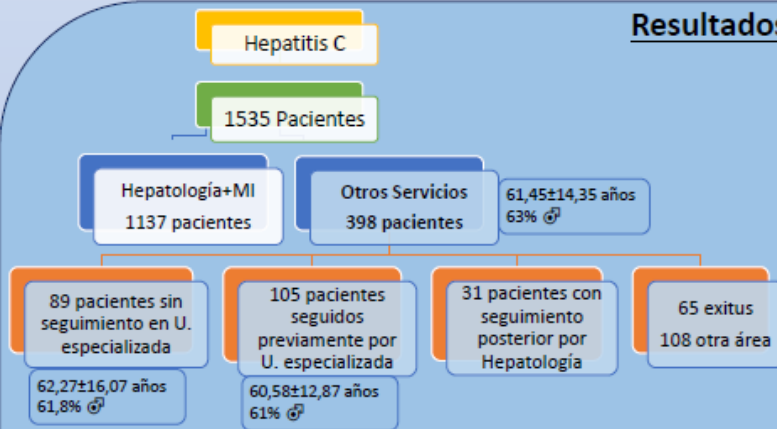


Figura 6. Pacientes obtenidos de Savana y tras la revisión de sus historias clínicas.

- No diferencias significativas de edad entre grupos.
- 105 de 398 (26,4%) pacientes erróneamente clasificados por Savana.

	GOT (U/L)	GPT (U/L)	INR	BRT (mg/dL)
	Media±DE	Media±DE	Media±DE	Media±DE
General	53,37±110,36	60,82±115,35	1,08±0,26	1±1,29
NO seguidos	35,01±24,21	33,49±26,18	0,99±0,11	0,65±0,35
SÍ seguidos	77,66±185,12	104,40±184,98	1,16±0,40	1,24±1,76

Tabla 8. Valores de GOT, GPT, INR y bilirrubina total (BRT) en pacientes en general, en no seguidos, y en sí seguidos previamente por una U. especializada

- GOT y GPT significativamente inferiores en no seguidos vs. sí seguidos.
- Aún así, hay que seguir a los pacientes aún estando asintomáticos.

Conclusiones

- Detectados **89 pacientes** sin acceso a seguimiento especializado ni a tratamiento.
- **Savana necesita aprendizaje** (26,4% clasificados erróneamente).
- Necesidad de un **protocolo o alerta** para derivar pacientes con hepatitis C a U. de Hepatología o especializada.

Nº de pacientes sin seguimiento en Unidad especializada por servicios (N=89)

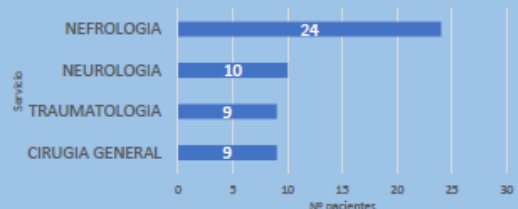


Figura 9. Número de pacientes no seguidos por una Unidad especializada según servicio por el que sí tiene seguimiento.

- Hasta 16 servicios con pacientes sin seguimiento.
- Necesidad de **protocolo o alerta** de derivación.

Diagnósticos principales

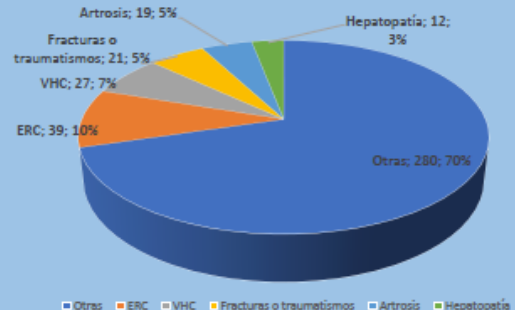


Figura 10. Diagnósticos principales de los pacientes revisados.

- ERC como diagnóstico más frecuente.

Bibliografía

- Reiser M, Wiebner B, Hirsch J, the German Liver Foundation. Neural-network analysis of socio-medical data to identify predictors of undiagnosed hepatitis C virus infections in Germany (DETECT). J Transl Med. 19 de marzo de 2019;17(1):94.
- WHO | Global health sector strategy on viral hepatitis 2016-2021 [Internet]. WHO. World Health Organization; [citado 25 de mayo de 2020]. Disponible en: <http://www.who.int/hepatitis/strategy2016-2021/ghss-hep/en/>
- Savana - Transforma el texto de tus historias clínicas en Big Data [Internet]. [citado 6 de octubre de 2019]. Disponible en: <https://savanamed.com/es/>



9. Resolución comité ético



INFORME DEL COMITÉ DE ÉTICA DE LA INVESTIGACIÓN CON MEDICAMENTOS (CEIm)

Dña. ROSA Mª CONDE VICENTE, Secretario del Comité de Ética de la Investigación con medicamentos (CEIm) del Área de Salud Valladolid Oeste

CERTIFICA:

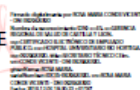
Que este Comité ha tenido conocimiento del Proyecto de Investigación, Trabajo Fin de Grado (TFG), titulado: **“Utilidad de un sistema de análisis masivo de datos (Big Data) insertado en la historia clínica electrónica en la búsqueda activa de pacientes con hepatitis C”**, Ref. CEIm: PI218-19, Protocolo versión 1.0, y considera que:

Una vez evaluados los aspectos éticos del mismo, acuerda que no hay inconveniente alguno para su realización, por lo que emite **INFORME FAVORABLE**.

Este Proyecto de Investigación será realizado por el alumno **D. Miguel Bardají Carrillo**, siendo su tutora en el Hospital Universitario Río Hortega la **Dra. Gloria Sánchez Antolín**.

Lo que firmo en Valladolid, a 4 de Diciembre de 2019

ROSA MARIA
CONDE VICENTE
- DNI
09296839D



Fdo. Dña. Rosa Mª Conde Vicente
Secretario Técnico CEIm