

Plataforma de analítica de datos para la caracterización poblacional y la evaluación del riesgo cardiovascular en pacientes del centro-occidente de Colombia

Data analytics platform for population characterization and cardiovascular risk assessment in patients from the west-center of Colombia

Genaro Daza S¹ ; Jhon Jairo Castañeda* ; Jorge Iván Castaño¹ 

* investigador01@angiohealth.co

Forma de citar: Daza G, Castañeda J, Castaño JI. Plataforma de analítica de datos para la caracterización poblacional y la evaluación del riesgo cardiovascular en pacientes del centro-occidente de Colombia. Salud UIS. 2022; 54: e22042. doi: <https://doi.org/10.18273/saluduis.54.e:22042> 

Resumen

Introducción: las enfermedades cardiovasculares son la primera causa de muerte en el mundo. Por tanto, muchas investigaciones han sido dirigidas hacia la predicción del riesgo cardiovascular, con el fin de poder evitarlo. Asimismo, se ha buscado la implementación de sistemas que involucren el análisis de datos automatizados que permita que la información se ponga a disposición, no solo del personal administrativo y directivo, sino también del personal clínico, para mejorar el control de las patologías. **Objetivo:** construir una herramienta para la caracterización poblacional y la evaluación del riesgo cardiovascular en pacientes del centro-occidente de Colombia. **Materiales y métodos:** se propone la construcción de una plataforma de análisis de datos sociodemográficos y clínicos. El modelo general de diseño de la plataforma es el desarrollo evolutivo, que entrelaza actividades de especificación, desarrollo y validación. La plataforma presenta un modelo vista-controlador que permite la creación de plantillas dinámicas distribuidas en módulos de acceso controlados por perfiles de usuario. **Resultados:** se implementó el cálculo automatizado del riesgo de enfermedad cardiovascular y la emisión de alertas tempranas, lo cual mejoró la gestión de los procesos clínicos, así como el apoyo a la toma de decisiones administrativas, a través de la conformación de dos módulos interactivos en la plataforma. **Conclusiones:** la unión de saberes clínicos, administrativos y de ingeniería permitió la consolidación de una herramienta que contribuye en el monitoreo y trazabilidad de los pacientes, orientando la priorización de posibles intervenciones que impacten en la salud de estos.

Palabras clave: Enfermedades cardiovasculares; Factores de riesgo; Ciencia de datos; Desarrollo tecnológico; Colombia.

¹Unidad Vasculard Cardiológica y Neurológica S. A. S., Cartago. Valle del Cauca, Colombia.

Abstract

Introduction: Cardiovascular diseases are the leading cause of death in the world. Countless research has been directed towards the prediction of cardiovascular risk, in order to avoid the threat. Furthermore, the implementation automated data analysis tools have been sought to allow for information to be made readily available, not only to administrative and managerial staff, but also to clinical staff to improve the control of pathologies. **Objective:** To build a tool for the characterization of the population and the evaluation of cardiovascular risk in patients from central-western Colombia. **Materials and methods:** The construction of a platform for the analysis of sociodemographic and clinical data is proposed. The overall platform design model is evolutionary development, which intertwines specification, development, and validation activities. The platform presents a Vista-Controller model, which allows the creation of dynamic templates distributed in access modules controlled by user profiles. **Results:** The automated calculation of cardiovascular disease risk and the issuance of early warnings were implemented, which improved the management of clinical processes, as well as support for administrative decision-making, through the creation of two interactive modules on the platform. **Conclusions:** The union of clinical, administrative and engineering knowledge allowed the consolidation of a tool that contributes to the monitoring and traceability of patients, which guides the prioritization of possible interventions that impact the health of patients.

Keywords: Cardiovascular diseases; Risk factors; Data science; Technological development; Colombia.

Introducción

Las enfermedades cardiovasculares (ECV), que incluyen la cardiopatía coronaria, las enfermedades y accidentes cerebrovasculares (ACV), las arteriopatías periféricas, la cardiopatía reumática, las cardiopatías congénitas y las trombosis venosas profundas y embolias pulmonares¹, son la principal causa de muerte en todo el mundo. Cada año mueren más personas por ECV que por cualquier otra causa. Se calcula que en 2016 murieron por esta razón 17,9 millones de personas, lo cual representa el 31% de todas las muertes registradas en el mundo, de las cuales el 85% se deben a cardiopatía coronaria y ACV¹. Más de 3 millones de estas muertes ocurrieron en personas menores de 60 años. El porcentaje de muertes prematuras para ECV va desde 4% en países de ingresos altos hasta 42% en países de ingresos bajos².

En 2017, a causa de ECV fallecieron aproximadamente 2,8 millones de residentes en los Estados Unidos³, y según la Sociedad Europea de Cardiología, por ECV hubo 2,2 millones de muertes en mujeres y 1,9 millones de muertes en hombres, lo que equivale al 47% y al 39% de todas las muertes por género, respectivamente⁴. En la región de las Américas, la proporción de defunciones debidas a enfermedades cardiovasculares es de 28,1% (150,7 por cada 100 000 habitantes), y en Colombia, particularmente, es del 29,7%, lo que corresponde a la muerte de aproximadamente 185,6 personas por cada 100 000 habitantes^{5,6}, valor que supera claramente las cifras promedio de la región. Adicionalmente, las enfermedades cardiovasculares representan la mayor cantidad de hospitalizaciones, los más altos costos de

atención en salud y la mayor pérdida de productividad en los países desarrollados⁷.

Uno de los estudios más ampliamente conocido, el estudio del corazón de Framingham, ha transformado la manera como la población entiende las enfermedades cardíacas⁸. Los investigadores del estudio en mención identificaron que, entre otras cosas, el colesterol elevado y la presión arterial alta son factores importantes en la aparición de las enfermedades cardiovasculares. El estudio de Framingham acuñó la expresión “factor de riesgo” como un elemento o una característica medible que tiene una relación causal con un aumento de frecuencia de una enfermedad, y constituye un factor predictivo independiente y significativo del riesgo de contraer una enfermedad⁹. En tal sentido, este estudio demostró que la medición y el conocimiento de otras variables fisiológicas permitía establecer e inclusive predecir las condiciones de riesgo futuras de la población frente a una patología. La investigación en enfermedades cardiovasculares basa su avance y sus éxitos en la posibilidad de evaluar los datos de las poblaciones tratadas para encontrar elementos comunes de protección o de exposición que permitan el desarrollo de nuevas estrategias de diagnóstico, tratamiento y pronóstico.

En tal virtud, la investigación soportada en datos juega un papel fundamental para mejorar los resultados de los pacientes con enfermedades cardiovasculares y para la sostenibilidad de los sistemas de salud. Los avances en diagnósticos, terapias y pronósticos de los pacientes con ECV son generalmente frustrados por el acceso limitado a los datos de los pacientes, por la no

estandarización de las mediciones de las variables y por la falta de resultados centrados en los pacientes⁷.

Por otra parte, en las organizaciones de salud, la gestión de la información tiene una relación directa con el mejoramiento de la calidad asistencial y la eficiencia de los servicios, con una alta incidencia en la toma de decisiones, por lo que debe ser objeto de atención¹⁰. Es fundamental disponer de herramientas ágiles para el análisis y visualización de la información que contribuyan en la caracterización de la población, en el entendimiento de sus condiciones de salud y, particularmente en este trabajo, en la evaluación del riesgo cardiovascular, de manera que se posibilite mejorar la calidad de vida de los pacientes a través de una oportuna y adecuada atención.

Recientemente, los investigadores han mostrado interés en el análisis automatizado de datos aplicado al campo de la salud. Múltiples trabajos han identificado la aplicación del análisis automatizado de datos de salud en tres áreas específicas: soporte a decisiones clínicas, cuidados personalizados y operaciones clínicas¹¹. En este contexto organizacional, el análisis automatizado de datos se conoce como inteligencia de negocios, y permite que el proceso de toma de decisiones sea más efectivo, dado que los usuarios pueden acceder a la información de manera rápida y consistente. Las empresas del sector salud usan la inteligencia de negocios para construir tableros de mando que ayudan en la administración, el monitoreo financiero y el desempeño clínico. La implementación de sistemas que involucran el análisis automatizado de los datos ha permitido que la información se ponga a disposición, no solo del personal administrativo y directivo, sino también para el personal clínico, quienes hacen su mayor uso¹².

Existen múltiples referencias del uso del análisis automatizado de datos enfocado en temas de salud. Por ejemplo: Diez *et al.*¹³ presentan el desarrollo de un sistema de georreferenciación para la gestión, movilidad y monitoreo de atención primaria de la salud comunitaria, el cual implica la implementación de una aplicación, para celular o *tablet* y acceso web, que considere las condiciones básicas, de ambiente y movilidad histórica de los usuarios, y que vincule metodologías de actualización cartográfica a partir de mapas sociales colaborativos en plataformas libres. En Leopold *et al.*¹⁴ se resaltan los retos actuales en relación con el genotipado y fenotipado de enfermedades

cardiovasculares que pueden ser resueltos por medio de la integración del *big data* e interpretados usando metodologías analíticas novedosas como el análisis de red. En Villarreal y Samudio¹⁵ se presenta una plataforma que permite la gestión de datos de pacientes con enfermedades crónicas no transmisibles, como la hipertensión arterial; esta almacena los datos de los pacientes, el registro histórico y las tendencias de las medidas almacenadas, alertas y recomendaciones de acuerdo con los rangos de mediciones obtenidas. Una perspectiva amplia acerca de las posibilidades del uso del análisis automatizado de datos en salud se encuentra en Sakr y Elgammal¹⁶, en este documento se definen y discuten los principales retos en los sistemas de salud que pueden ser efectivamente enfrentados por medio de los avances en las tecnologías de la información y la comunicación, con énfasis en tecnologías de sensores, computación en la nube, internet de las cosas y *big data*, lo cual fortalece la eficiencia y eficacia de los servicios de salud. De igual manera, Sanches-Acevedo *et al.*¹⁷, proponen una metodología para la detección de ECV y una herramienta web para analizar los datos de manera más efectiva. La metodología incluye extracción, descripción y visualización de datos, y es ejemplificada a partir de un estudio de ECV en Ciudad de México, empleando bases de datos como la del Instituto Nacional de Estadística y Geografía y la Encuesta Nacional de Salud y Nutrición.

En nuestro trabajo se presentan los resultados de una investigación aplicada, a través de la cual se realizó el diseño e implementación de una plataforma informática que permite la captura, el almacenamiento, el tratamiento, la gestión y la visualización de datos a través de estadísticas descriptivas y analíticas de individuos y subgrupos poblacionales; así mismo, la plataforma brinda el análisis del riesgo cardiovascular de los pacientes que pertenecen a programas de seguimiento y monitoreo, atendidos en la Unidad Vasculard Cardiológica y Neurológica.

La Unidad Vasculard Cardiológica y Neurológica es una institución prestadora de servicios de salud, especializada en el diagnóstico, tratamiento y rehabilitación de patologías de origen cerebrocardiovascular, ubicada su sede principal en el municipio de Cartago y una sede satélite en el municipio de Roldanillo, ambos en el departamento del Valle del Cauca, en la región centro occidental de Colombia. En la zona de influencia (Figura 1) de la Unidad Vasculard, Cardiológica y Neurológica hay 18 municipios.

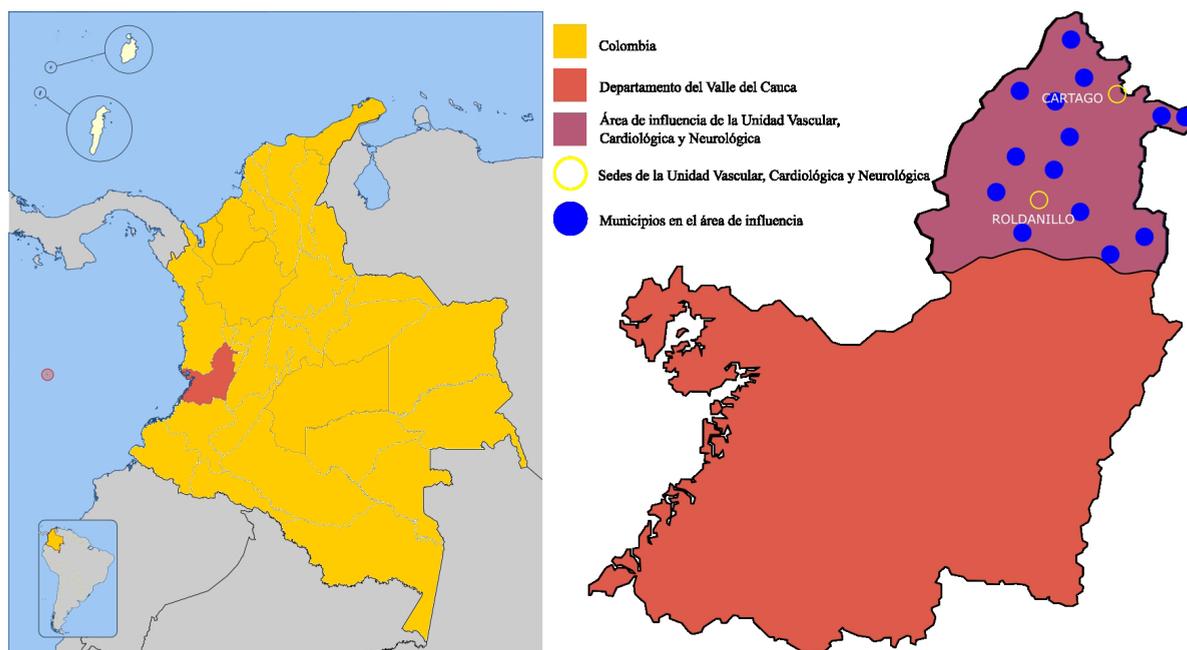


Figura 1. Área de influencia de la Unidad Vascular, Cardiológica y Neurológica en el centro-occidente de Colombia. Fuente: adaptación del mapa realizado por Milenioscuro en Wikipedia.

La población total del área de influencia asciende a 396 473 personas¹⁸, repartida en los siguientes 18 municipios pertenecientes al departamento del Valle del Cauca: Alcalá, Ansermanuevo, Argelia, Bolívar, Caicedonia, Cartago, El Águila, El Cairo, El Dovio, La Unión, La Victoria, Obando, Roldanillo, Sevilla, Toro, Ulloa, Versalles, Zarzal.

Además, es importante mencionar que la pirámide poblacional para estos municipios es, en general, ancha en las edades de 45 a 60 años, lo cual muestra una tendencia al envejecimiento. Según el análisis de la situación de salud del Valle del Cauca del año 2019, en la población general del departamento, las muertes por enfermedades del sistema circulatorio presentan mayor incidencia, según la tasa de mortalidad ajustada por edad¹⁹. En el mismo estudio se afirma que las enfermedades isquémicas del corazón constituyen la principal causa de muerte por enfermedades del sistema circulatorio para todos los años del estudio. La tasa se ha incrementado de 108,26 muertes por 100 000 habitantes en 2015 a 117,83 en 2017. La tendencia general de este evento es al incremento. Adicionalmente, las enfermedades cerebrovasculares son la segunda causa de muerte por enfermedades del sistema circulatorio.

La Unidad Vascular Cardiológica y Neurológica atiende a una población de más de 9.000 pacientes en la región, con patologías crónicas no transmisibles,

donde destacan: diabetes *mellitus*, hipertensión arterial y enfermedad renal crónica. Bajo una visión de trabajo basada en riesgos, en la institución se desarrolla un programa de seguimiento y monitoreo de pacientes, que tiene en cuenta las características epidemiológicas de la población objetivo. Dentro de este enfoque de seguimiento y monitoreo, se cuenta con información histórica de 1.118 pacientes. Las fuentes de información disponibles en la institución con datos de los pacientes son las siguientes:

- **Datos administrativos:** comprende los registros del paciente tales como su identificación completa, sexo, fecha de nacimiento, lugar de procedencia, estado civil, régimen de afiliación al sistema de seguridad social, fechas y códigos de atenciones, consultas y/o exámenes realizados.
- **Datos clínicos:** en los datos clínicos se dispone principalmente de los datos extraídos de la historia clínica del paciente, la cual recopila información de relevancia para la evaluación y seguimiento del paciente. Incluye diagnósticos de riesgo cardiovascular con sus respectivas fechas de detección, farmacoterapia, datos sociodemográficos complementarios y, entre otros datos clínicos, se tiene peso, talla, índice de masa corporal, perímetro abdominal, presión arterial sistólica, presión arterial diastólica, antecedentes personales patológicos, antecedentes familiares patológicos.

- **Datos de laboratorios:** en general, los pacientes bajo monitoreo cuentan mayoritariamente con los siguientes registros, valor de hemoglobina glicosilada, glucemia, creatinina, colesterol total, colesterol de alta densidad, colesterol de baja densidad, triglicéridos, albuminuria, relación albuminuria-creatinuria en orina.
- **Datos de hábitos de vida:** acerca de los hábitos de vida de los pacientes de los programas de monitoreo y seguimiento, se cuenta con hábitos de tabaquismo, hábitos de consumo de alcohol, evaluación de actividad de física bajo la prueba PASE (*The Physical Activity Scale for the Elderly*, la escala de actividad física para adultos mayores)¹⁹, hábitos de dieta de consumo de vegetales, frutas, grasas, azúcares y frituras.

Previo al desarrollo de la plataforma de analítica de datos de la Unidad Vascular, Cardiológica y Neurológica, existía desconocimiento parcial en las relaciones estadísticas entre variables clínicas y sociodemográficas de la población atendida, además de falta de información sistematizada sobre los hábitos de vida de los pacientes, que permitiera la creación de estrategias de asignación de citas y controles, según el perfil de riesgo de los pacientes. La incapacidad para identificar tempranamente los riesgos cardiovasculares en los pacientes era otra dificultad latente en la IPS antes de la plataforma; el acceso a los datos de la población en conjunto implicaba solicitudes adicionales al administrador de la base de datos de historia clínica, que generalmente no podían realizarse en un tiempo prudente, lo cual impedía la implementación de modelos computacionales para caracterizar y analizar a la población. La IPS contaba con un *software* para el registro de historias clínicas y atención del paciente, sin embargo, dicho *software* no tenía la capacidad para unificar y relacionar la información registrada, lo que dificultaba la creación de informes que dieran conocimiento de la población.

La plataforma de analítica de datos descrita en este documento contiene un módulo de informes soportado en estadística descriptiva y un módulo de formularios para la recolección de información administrativa de la empresa y sus colaboradores. En particular, en el módulo de informes, se tienen los siguientes reportes: 1) datos demográficos, 2) datos clínicos, 3) atenciones generales, 4) atenciones detalladas, 5) atenciones en programas de monitoreo y seguimiento, 6) laboratorios clínicos, 7) sistema de gestión de seguridad y salud en el trabajo, 8) seguridad del paciente, 9) reporte de ideas, 10) satisfacción del usuario, 11) descarga de formularios

estandarizados, 12) perfil del paciente (incluye análisis de riesgo cardiovascular).

A lo largo de este documento se presentan los detalles de la información que brinda cada uno de los reportes anteriormente mencionados, su metodología de construcción y operatividad. El desarrollo de la plataforma tiene un claro impacto en la gestión administrativa de la Unidad Vascular, Cardiológica y Neurológica, ya que, entre otras cosas, la programación de citas y estudios de los pacientes puede realizarse teniendo en consideración variables como riesgo cardiovascular, cumplimiento de los parámetros del programa de monitoreo y seguimiento, frecuencia de las atenciones anteriores, etcétera, lo cual impacta en la calidad de vida de los pacientes.

Metodología

El modelo general para el desarrollo de la plataforma de analítica de datos es el desarrollo evolutivo, este enfoque entrelaza las actividades de especificación, desarrollo y validación. Es decir, un sistema inicial se desarrolla a partir de especificaciones generales y se refina con base en las peticiones de los clientes, para producir un sistema que satisfaga sus necesidades²¹. En particular, se trabaja en el desarrollo evolutivo exploratorio con modelado de prototipo, en el cual el objetivo es trabajar con el cliente para explorar sus requerimientos y entregar un sistema final, empezando con las partes de la plataforma que se comprenden mejor, y evoluciona agregando nuevos atributos solicitados. En este sentido, se desarrolla una implementación inicial sujeta a comentarios de los usuarios, la cual realizará los respectivos ajustes y correcciones a través de las diferentes versiones, hasta tener un sistema adecuado¹⁵.

Levantamiento de requerimientos

En una primera fase, se realiza el levantamiento de requerimientos de las necesidades informativas, para la adecuada atención de los pacientes y el análisis de las fuentes de información disponibles. En tal sentido, de manera conjunta con el personal clínico y el personal administrativo, se identifican las principales necesidades de información para mejorar la toma de decisiones en cuanto a la oferta de servicios de salud, para la formulación de estrategias de asignación de citas y, especialmente, para hacer seguimiento y monitoreo integral del estado de salud de los pacientes. La información se recolecta por medio de la técnica “mesas de trabajo”^{23,24,25}.

El proceso inicia estableciendo un lenguaje común entre las personas que conocen el funcionamiento administrativo de la entidad, el equipo clínico que conoce las necesidades de información para la atención de los pacientes y el equipo desarrollador. En el caso particular, bajo el método de mesas de trabajo, se establece un cronograma de actividades con participantes preseleccionados, y se siguen las recomendaciones de buenas prácticas para reuniones efectivas. Por medio de las mesas de trabajo se indaga al personal acerca de sus necesidades de información a nivel operativo y de planificación, para establecer elementos comunes entre los participantes y documentar los resultados. El documento final de requerimientos contiene una descripción de cada uno de los requerimientos señalados por los participantes, en un lenguaje natural, y cuando es posible, con ejemplos que dan claridad del motivo del requerimiento. Adicionalmente, el equipo desarrollador determina las entradas necesarias y las salidas esperadas, de manera concreta y sintética, y las observaciones de gestión que permitan dar solución al requerimiento²⁶.

Uno de los primeros aspectos por resaltar, identificados durante el levantamiento de requerimientos, es la existencia de múltiples fuentes disgregadas de información de los pacientes, entre las cuales se establecen como prioritarias: los registros electrónicos de la historia clínica, los registros de citas y admisiones, los reportes electrónicos de resultados de laboratorios clínicos y los formularios internos acerca de hábitos de vida de los pacientes.

Además, una vez desarrollado el proceso de mesas de trabajo para el levantamiento de requerimientos, se pueden señalar como principales las siguientes necesidades informativas:

- Composición sociodemográfica de la población atendida.
- Análisis gráfico de condición fisiológica básica de la población atendida.
- Conocimiento del riesgo cardiovascular de los pacientes.
- Disponibilidad de alertas tempranas de riesgo cardiovascular para priorizar la atención.
- Información asociada a los hábitos de vida de los pacientes, en temas tales como tabaquismo, consumo de alcohol, estado físico, dieta y disponibilidad de servicios públicos y nevera.
- Cuantificación y representación de la cantidad, tipo y frecuencia de atenciones recibidas por los pacientes.

- Identificación del cumplimiento de los requerimientos mínimos de atención en programas de monitoreo y seguimiento.
- Resultados históricos de los laboratorios clínicos de los pacientes, con semaforización de alteraciones.
- Síntesis gráfica del nivel de satisfacción de los pacientes con el servicio.

Modelo vista controlador (MVC)

El modelo vista controlador es un patrón de arquitectura que separa una aplicación en tres principales componentes lógicos: el modelo, la vista y el controlador. Cada uno de estos componentes se construye para manejar aspectos específicos de desarrollo de una aplicación. MVC es uno de los marcos de desarrollo web estándar de la industria más utilizado para crear proyectos escalables y extensibles²².

La parte superior de la **Figura 2** muestra detalladamente el funcionamiento de este patrón arquitectónico para la plataforma de analítica de datos, acá se puede observar la estructura de cada sección en dos componentes de creación distintos: *FrontEnd* y *BackEnd*. El *FrontEnd* está conformado por el cliente y la visualización. En relación con el cliente, la plataforma se ha diseñado para que su uso sea posible en los navegadores web más empleados, soportada para la visualización mediante el uso de HTML y CSS, y por el uso de Jinja2, que incluye un sistema de mapeo de datos relacional que facilita la creación de plantillas web. Por otro lado, el *BackEnd* de la plataforma está constituido por el modelo (servidor) y por el controlador. En el caso del controlador, la interactividad web es posible mediante el uso de JavaScript (JS) en la relación con el cliente, y con Axios para la conexión de llamados y respuestas del cliente hacia el servidor y viceversa; las peticiones realizadas mediante Axios son del tipo REST-JSONAPI, cuya comunicación se realiza a través del intercambio de datos tipo JSON. Por último, en el modelo o servidor, se presenta como servidor principal el *microframework* Flask de Python, para construir las aplicaciones y mediante el uso de Werkzeug gestionar el enrutado y la integración de las aplicaciones, así como la definición del servidor principal. Además, mediante la integración de MongoDB como administrador de bases de datos NoSQL, las aplicaciones diseñadas en Flask pueden almacenar y compartir información de manera segura. Por último, Flask hace uso de Jinja2, como se mencionó anteriormente, para la generación de plantillas HTML; y el uso de las herramientas de Flask en Python también permite la generación de aplicaciones a partir de un sistema de plantillas definidas en el *microframework*.

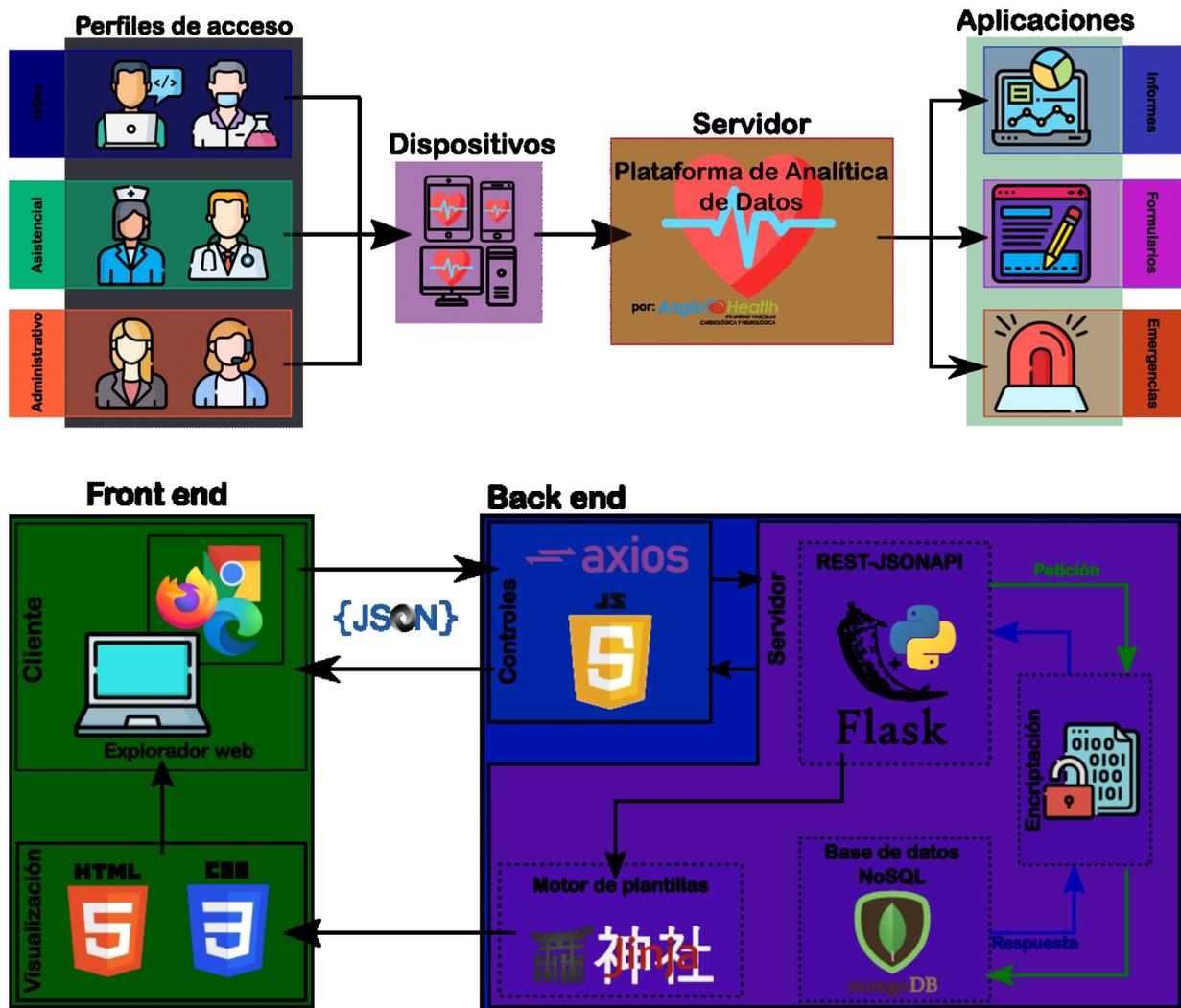


Figura 2. Esquema del modelo vista-controlador y del flujo de información en plataforma.
Fuente: elaboración propia a partir de íconos de Flaticon.

Este sistema de plantillas favorece la creación de código mucho más corto en las vistas. El sistema Flask de Python permite una sintaxis más reducida en su escritura. Los controladores contienen la lógica de la aplicación y permiten organizar el código en clases, sin tener que escribirlo todo en las rutas.

Arquitectura de la plataforma

La plataforma de analítica de datos se diseña para permitir la integración de aplicaciones para el registro de información y el monitoreo de datos, enfocada al control clínico de enfermedades cardiovasculares. La parte inferior de la Figura 2 presenta la arquitectura principal de la plataforma, en un diagrama de flujo que va desde los actores o usuarios hasta los módulos o aplicaciones finales.

La arquitectura que presenta la plataforma está inicialmente enfocada en la usabilidad de los datos por parte de la administración, el personal clínico y el departamento de investigación y desarrollo. La plataforma permite la toma de decisiones gerenciales y el monitoreo del ciclo de atenciones prestado por el personal clínico. Por otra parte, desde el área clínica, la plataforma plantea herramientas efectivas para el análisis de datos de los parámetros fisiológicos de los pacientes, permitiendo a los profesionales médicos y enfermeros verificar, entre otros, el estado general y particular de la población atendida, y llevar un seguimiento ordenado de la atenciones y tratamientos de los pacientes, facilitando la adherencia al proceso. Finalmente, para el departamento de investigación y desarrollo, la plataforma brinda la posibilidad de estudiar patrones en los datos, que permitan la implementación

de métodos para la reducción de los riesgos en salud, particularmente del riesgo cardiovascular.

Como se ve en la parte inferior de la **Figura 2**, se han definido algunos perfiles de acceso general en la arquitectura de la plataforma, los cuales permiten segmentar la información pertinente a cada una de las áreas de la institución. Estos perfiles de usuarios son definidos por la administración y el área de desarrollo, generados de acuerdo con la contratación y cargo actual de cada colaborador, validado a través de un correo institucional que permite el acceso a la plataforma, y, por ende, la visualización de sus respectivos informes y formularios. Es decir, la única forma de acceder a los servicios y a la información que se ofrece a través de la plataforma es ser miembro activo de la institución y tener las respectivas autorizaciones, lo anterior con el fin de garantizar la seguridad e integridad de los datos.

La información obtenida de los pacientes está centralizada en una base de datos NoSQL MongoDB, como se muestra en la parte superior de la **Figura 2**, la cual cuenta con dos métodos de recopilación principales. El primero de ellos es la sincronización periódica (diaria) con la base de datos, estructurada en MySQL, que administra el *software* principal de los registros clínicos de atención de los pacientes, la cual contiene aproximadamente 330 000 registros asociados a 14 000 pacientes, 5 tablas accesibles con 94 variables informativas. La segunda fuente de alimentación de información de la base de datos corresponde a formularios presentes en la plataforma, diseñados con base en las necesidades de los profesionales que interactúan con los pacientes; estos formularios son creados como elementos interactivos HTML, controlados a través de *scripts* de interacción en JS integrados al servidor principal de la plataforma, simulando el comportamiento de una API web. Toda la información en la base de datos NoSQL se encuentra protegida por credenciales de acceso que se verifican en una base de datos alterna de usuarios creada por el administrador de la plataforma. Además, todos los datos en la base de datos NoSQL se encuentran almacenados mediante un proceso de encriptación simétrica, donde cada petición de datos al servidor solo puede ser ejecutada de manera local. La base de datos NoSQL tiene un tamaño de almacenamiento en disco de 33,6 MB y contiene la información de 14 000 usuarios, en 15 colecciones de datos distribuidas en 207 variables, y se proyecta que en los próximos 5 años crezca a un tamaño aproximado de 200 MB garantizando de esta manera la viabilidad de crecimiento de la plataforma sin cambios mayores de infraestructura. Los datos

registrados por medio de los formularios pueden ser consultados por el personal de la institución, a través de los informes específicamente creados. En particular, los informes se soportan en la creación de objetos JSON creados en Flask a partir de los datos y renderizados en HTML mediante el uso de JS utilizando Plotly como librería gráfica interactiva.

Finalmente, es importante resaltar la compatibilidad de la plataforma web en cuanto al acceso desde diversos dispositivos, pues permite el funcionamiento tanto en teléfonos móviles como en computadores personales, desde un servidor físico alojado en la sede administrativa principal de la institución.

Resultados

Módulos de la plataforma

Tal como se mencionó previamente, la plataforma de analítica de datos para la caracterización poblacional y la evaluación del riesgo cardiovascular contiene 2 módulos principales:

- **Módulo de formularios:** este módulo permite la recolección de información de pacientes, que no es registrada en el sistema integrado de historia clínica. Por medio de este módulo se registran los resultados de laboratorios clínicos, condición de alfabetización de los pacientes, eventos adversos, hábitos de vida (tabaquismo, alcohol, actividad física y dieta) y satisfacción del usuario.
- **Módulo de informes:** este módulo presenta reportes de información y estadísticas acerca de datos demográficos, datos clínicos, atenciones generales, atenciones detalladas, atenciones en programas de monitoreo y seguimiento, laboratorios clínicos, sistema de gestión de seguridad y salud en el trabajo, seguridad del paciente, reporte de ideas, satisfacción del usuario, descarga de formularios estandarizados, perfil del paciente (incluye análisis de riesgo cardiovascular). A lo largo de este documento se presentan algunos de estos reportes y su funcionalidad dentro de la plataforma.

En general, todos los reportes contienen filtros interactivos de variables como periodo del informe, sexo, edad, sede de atención, entidad promotora de salud afiliada, régimen de afiliación, diagnóstico principal, municipio, zona urbana o rural, estado de analfabetismo. Algunos reportes cuentan con filtros específicos asociados al tipo de reporte, como estado de laboratorios clínicos, antecedentes personales, índice de masa corporal, tipo de atención, entre otros.

En el reporte de datos demográficos se condensa la caracterización de la población atendida en la Unidad Vasculuar, Cardiológica y Neurológica, por medio de representaciones gráficas interactivas y estadísticas descriptivas por sexo, edad, diagnóstico principal, régimen de afiliación, sede, zona urbana o rural, georreferenciación del municipio de procedencia. Algunos ejemplos de este reporte son presentados en la **Figura 3**.

El reporte de datos clínicos recopila, mediante un diagrama de dispersión, la relación de las siguientes variables: edad, peso, talla, perímetro abdominal, frecuencia respiratoria, presión arterial sistólica, presión arterial diastólica e índice de masa corporal.

Por otra parte, los reportes de atenciones generales, atenciones detalladas y atenciones en programas de monitoreo y seguimiento son reportes de apoyo a la gestión administrativa. En estos reportes se presenta la cuantificación diaria de las actividades realizadas a nivel de consulta externa, estudios diagnósticos e intervenciones mínimamente invasivas. Los reportes relacionados con atención permiten identificar la frecuencia de atenciones en salud, recibida por cada paciente, dentro de un período establecido.

El reporte de laboratorios clínicos permite clasificar los resultados de los laboratorios entre: normal, alterado bajo, alterado medio y alterado alto, lo cual contribuye con la priorización en la atención de los pacientes que presenten mayor nivel de alteración en sus pruebas. La clasificación de los resultados de laboratorios se realiza con base en escalas numéricas, internacionalmente aceptadas, que dan cuenta de la normalidad del resultado; es decir, cada variable de laboratorio medida es asociada a un rango que determina la clasificación del resultado de la prueba de laboratorio. Por ejemplo, los triglicéridos se consideran normales en un rango de 0 a 150 mg/dl, alterado bajo en el rango mayor a 150 y menor o igual que 200 mg/dl, alterado medio en el rango mayor a 200 y menor o igual que 500 mg/dl y alterado alto si es mayor a 500 mg/dl. Igualmente, el reporte de laboratorios contribuye en la gestión administrativa, dado que permite tener el registro de la cantidad de pruebas aplicadas.

El reporte de satisfacción al usuario permite interpretar la percepción de los pacientes acerca de los servicios de salud que reciben, para esto se analizan variables como la sede de atención, el tipo de servicio prestado,

los niveles de satisfacción y los posibles motivos de insatisfacción. En la **Figura 4** se presentan algunos gráficos que ejemplifican este reporte.

Finalmente, el reporte de perfil del paciente es una síntesis gráfica del historial médico del paciente, en el que se incluye una ficha de identificación; datos de su atención médica más reciente; gráfica en simultáneo de dos variables clínicas o de dos variables de laboratorio a lo largo del tiempo; normalidad de los laboratorios; cuantificación y fecha de las atenciones realizadas; resumen histórico de antecedentes personales; hábitos de vida, tales como consumo de licor, tabaquismo, actividades físicas evaluadas por cuestionario PASE²⁰, dieta. Adicionalmente, se encuentra la evaluación del riesgo cardiovascular del paciente, la cual es un elemento de representación esencial para establecer la conducta del paciente y su nivel de priorización. En la **Figura 5** se presentan algunos ejemplos de los diagramas interactivos que dispone el informe del perfil del paciente.

Por otra parte, es importante considerar el nivel de usabilidad de la plataforma desarrollada, que indica la facilidad con la que un usuario puede usar una aplicación de *software* o hace referencia a cómo los usuarios se pueden desenvolver al interactuar con una aplicación^{27,28}. Con el fin de evaluar la usabilidad de la plataforma desarrollada, se emplea la escala de usabilidad del sistema (SUS)²⁹. La escala SUS genera una puntuación de usabilidad entre 0 y 100 puntos; los valores promedio de una muestra poblacional pueden interpretarse como excelente (mayor o igual a 90), buena (mayor o igual que 80 y menor que 90), aceptable (mayor o igual que 70 y menor que 80), marginalmente aceptable alto (mayor o igual que 62,5 y menor que 70), marginalmente aceptable bajo (mayor o igual que 50 y menor que 62,5) y no aceptable (menor que 50)³⁰. Para la plataforma, la metodología SUS se aplicó sobre una muestra de 12 usuarios frecuentes, conformada por 5 personas del equipo administrativo, 5 personas del equipo clínico y 2 personas del equipo de investigación; se obtuvo:

- Media \pm desviación estándar: 92,08 \pm 7,37
- Mediana [cuartil 1; cuartil 3]: 95 [83,75; 99,37]

Estos resultados indican a nivel de usabilidad que la muestra de usuarios considera, en promedio, que la plataforma desarrollada es excelente.

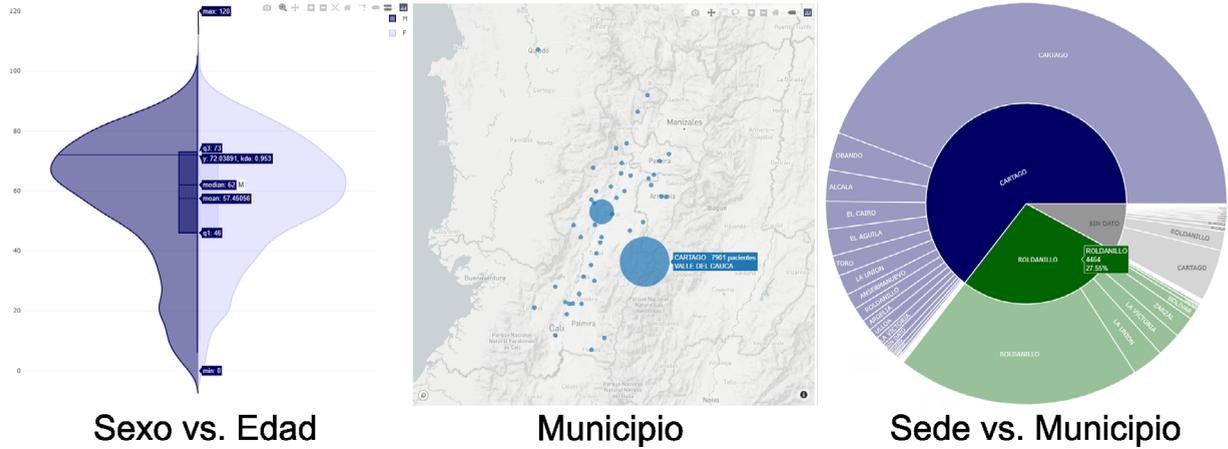


Figura 3. Diagramas del reporte de datos sociodemográficos. Izquierda: gráfico de violín que representa la distribución poblacional por sexo y edad; centro: georreferenciación de la procedencia de los pacientes; derecha: gráfico interactivo de relación entre sede y municipios.

Fuente: elaboración propia a partir de la plataforma de analítica de datos de la Unidad Vasculár Cardiológica y Neurológica.

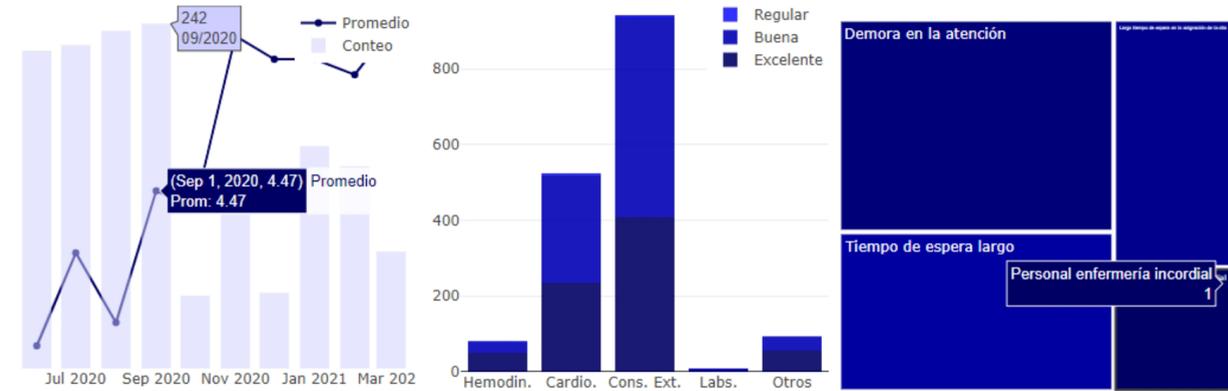


Figura 4. Gráficos del reporte de satisfacción del usuario. Izquierda: cantidad de evaluaciones de satisfacción del servicio y calificación promedio (1 a 5) otorgada por los pacientes con frecuencia mensual; centro: calificaciones de los usuarios por tipo de servicios; derecha: conteo de frecuencia de los motivos de insatisfacción de los pacientes.

Fuente: elaboración propia a partir de la información y resultados de la plataforma de analítica de datos de la Unidad Vasculár Cardiológica y Neurológica.

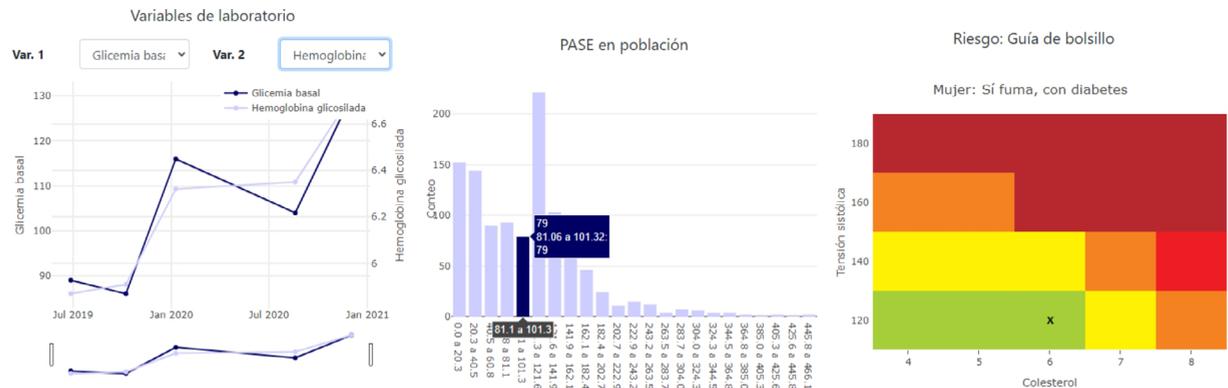


Figura 5. Gráficos ejemplo del reporte del perfil del paciente. Izquierda: diagrama comparativo de variables de laboratorio; centro: diagrama de resultado general de la prueba PASE respecto a la población; derecha: diagrama de riesgo cardiovascular de cada paciente.

Fuente: elaboración propia a partir de la información y resultados de la plataforma de analítica de datos de la Unidad Vasculár Cardiológica y Neurológica.

Discusión

La plataforma de datos se propone como una herramienta que permite la articulación de información clínica y demográfica para el estudio y seguimiento de las patologías de los pacientes. Analizando los datos clínicos históricos de las consultas realizadas, así como los laboratorios practicados y la información de hábitos de vida de los pacientes. Este aspecto diferenciador proporciona nuevas estrategias de análisis y toma de decisiones, congruentes y coherentes con el estado de salud particular de cada paciente.

Al inicio del proyecto, la información de los pacientes no estaba debidamente centralizada ni agrupada en una única base de datos; en igual sentido, no existía estandarización de la información en cuanto a nomenclatura de actividades y tipo de reportes, por ende, la posibilidad de tener análisis estadísticos era escasa y los resultados podrían carecer de representatividad y confiabilidad. Gracias a la implementación de una base de datos NoSQL, los datos pudieron ser reestructurados y depurados correctamente, lo cual permitió la integración de múltiples fuentes de información, previamente desagregadas, así como la extracción de información pertinente. De hecho, por medio de la plataforma de analítica de datos se consolida la posibilidad de contar con datos actualizados, centralizados y depurados, lo que repercute en significativas mejoras al momento de implementar metodologías para el cálculo y prevención del riesgo cardiovascular, monitorear la adherencia de los pacientes a sus tratamientos y llevar una trazabilidad de las atenciones necesarias y requeridas por cada uno de ellos.

Desde otra perspectiva, la plataforma brinda la posibilidad de diseñar nuevos módulos de caracterización, análisis y evaluación de los problemas de salud de la población; así mismo, inspira y brinda datos concretos para la construcción de nuevos protocolos de intervención. Adicionalmente, con la plataforma se motiva la creación de planes administrativos para el control y monitoreo desde un enfoque gerencial.

Algunos de los procesos de generación de informes, hasta antes del desarrollo de la plataforma, eran realizados manualmente por los profesionales de la institución, implicaban tiempos de ejecución largos, tareas repetitivas e inducían en muchas ocasiones errores en reportes e informes; estas barreras operativas son superadas gracias a la automatización y agilidad que ofrece la herramienta descrita.

Dada la metodología de desarrollo implementada, las funcionalidades de la plataforma siguen en crecimiento con base en la demanda de los usuarios. En particular, desde el punto de vista administrativo, se trabaja para la gestión de la seguridad y la salud en el trabajo; desde el punto de vista clínico se construyen nuevos índices de riesgo cardiovascular, soportados en inteligencia artificial, lo cual permite el ajuste de los resultados a las condiciones socioeconómicas y culturales de los pacientes de la región centro-occidental de Colombia.

Conclusiones

El desarrollo de una plataforma de analítica de datos para la caracterización poblacional y la evaluación del riesgo cardiovascular en pacientes del centro-occidente de Colombia permite mejorar la calidad asistencial y la eficiencia de los servicios ofrecidos en la Unidad Vasculard Cardiología y Neurológica. La plataforma claramente contribuye con información relevante y actualizada que soporta la toma de decisiones clínicas y administrativas. En particular, con la herramienta presentada, ha sido posible tener un monitoreo más preciso de las condiciones fisiológicas de los pacientes y, por tanto, brindar un tratamiento ajustado a los cambios particulares que cada paciente vive. A nivel administrativo, la plataforma de analítica de datos facilita, al cuerpo directivo, revisar indicadores poblacionales, conocer el comportamiento de las patologías atendidas e identificar posibilidades de mejora en la atención, así como identificar situaciones atípicas en el desempeño de la institución. Con esta información ha sido posible mejorar los términos y condiciones de relacionamiento con las empresas encargadas del aseguramiento de los pacientes. A través de la plataforma de analítica de datos se crearon estrategias de asignación de citas y control de patologías, mediante la caracterización de asignación de perfiles de riesgo poblacionales haciendo uso de la información sistematizada sobre los hábitos de vida de los pacientes, recolectada a través de los formularios de datos creados en la plataforma; además, la plataforma permitió determinar relaciones estadísticas entre las variables clínicas y sociodemográficas de los pacientes, lo cual permitió la creación de informes sobre el estado general de la población.

Un aspecto fundamental en el desarrollo de la plataforma es el valioso y enriquecedor enfoque multidisciplinario, dado que médicos, enfermeras, administradores e ingenieros han contribuido con sus ideas para la construcción de cada uno de los informes diseñados.

Es importante señalar que este tipo de desarrollos presenta a diario nuevos retos, lo que implica la necesidad de una evolución constante en la información brindada, en la herramienta de análisis y en el alcance de nuevos usuarios. Como reto a futuro cercano es necesario brindar información directa a los pacientes, con la intención de que ellos comprendan mejor el estado de su condición de salud y puedan ser apoyados por medio de consejos personalizados (educación personalizada en salud) asociados con hábitos de vida saludable, y, así mismo, con recordatorios específicos orientados a lograr una mayor adherencia al tratamiento.

Agradecimientos

La realización del desarrollo tecnológico presentado en este documento y su investigación asociada ha sido posible gracias al apoyo de la Unidad Vasculard Cardiológica y Neurológica S. A. S. y del Ministerio de Ciencias, Tecnología e Innovación de Colombia por medio de los proyectos “Plataforma de analítica de datos para la caracterización poblacional y la identificación de riesgos cardiovasculares en pacientes que atienden a la Unidad Vasculard Cardiológica y Neurológica SAS”, con código 8967-854-76382, y del proyecto “Sistema automatizado de identificación y evaluación de riesgo cardiovascular, por medio de una plataforma de analítica de datos, para la caracterización poblacional y la emisión de alertas tempranas”, con código 8967-869-79742.

Referencias

1. Organización Mundial de la Salud. Enfermedades cardiovasculares. OMS [Internet]. 2020. Disponible en: [https://www.who.int/es/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/es/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
2. Mendis S, Puska P, Norrving B. Global atlas on cardiovascular disease prevention and control [Internet]. 1st ed. France. World Health Organization; 2011. Available from: <https://apps.who.int/iris/handle/10665/44701>
3. Virani SS, Alonso A, Benjamin EJ, Bittencourt MS, Callaway CW, Carson AP, et al. Heart disease and stroke statistics—2020 Update: A report from the American Heart Association. *Circulation*. 2020; 141(9): e139–e596. doi: <https://doi.org/10.1161/CIR.0000000000000757>
4. Timmis A, Townsend N, Gale CP, Torbica A, Lettino M, Petersen SE, et al. European Society of Cardiology: Cardiovascular disease statistics 2019. *Eur Heart J*. 2020; 41(1): 12–85. doi: <https://doi.org/10.1093/eurheartj/ehz859>
5. Organización Panamericana de la Salud. Las ENT de un vistazo: Mortalidad de las enfermedades no transmisibles y prevalencia de sus factores de riesgo en la Región de las Américas. OPS; 2019. Disponible en: <https://iris.paho.org/handle/10665.2/51752>
6. Organización Panamericana de la Salud. Enfermedades no transmisibles: hechos y cifras. OPS; 2019. Disponible en: <https://iris.paho.org/handle/10665.2/51482>
7. Silverio A, Cavallo P, De Rosa R, Galasso G. health data and cardiovascular diseases: A challenge for research, an opportunity for clinical care. *Front Med (Lausanne)*. 2019; 6: 36. doi: <https://doi.org/10.3389/fmed.2019.00036>
8. Dawber TR, Moore FE, Mann GV. II. Coronary heart disease in the Framingham study. *Int J Epidemiol*. 2015; 44(6): 1767–1780. doi: <https://doi.org/10.1093/ije/dyv346>
9. O'Donnell CJ, Elosua R. Factores de riesgo cardiovascular. Perspectivas derivadas del Framingham Heart Study. *Rev Esp Cardiol*. 2008; 61(3): 299–310.
10. Torres Fernández JP, Gallo Mendoza JG, Hallo Alvear RF, Abcarius JJ, Muriel Páez MH, Fernández Lorenzo A. Gestión de la información como herramienta para la toma de decisiones en salud: escenarios más probables. *Rev Cub Inv Biomédicas*. 2017; 36(3): 1–10.
11. Verduzco Reyes G, Bautista Thompson E, Ruiz Vanoye JA, Fuentes Penna A. Modelos de tecnologías del Big Data Analytics y su aplicación en salud. *Rev Pistas Edu*. 2017; 128: 1583–1599.
12. Khedr A, Kholeif S, Saad F. An integrated business intelligence framework for healthcare analytics. *Int J Adv Res Comput Sci Softw Eng*. 2017; 7(5): 263–270. doi: [10.23956/ijarcsse/SV715/0163](https://doi.org/10.23956/ijarcsse/SV715/0163)
13. Diez Tetamanti JM, Rocha E, Munsberg G, Peixoto Castro JH, Neutzling AS, Jaime SF, et al. Desarrollo de un sistema georreferenciado para la gestión, movilidad y monitoreo de atención primaria de la salud comunitaria. *Salud Colect*. 2018; 14(1): 121–137. doi: <https://doi.org/10.18294/sc.2018.1210>
14. Leopold JA, Maron BA, Loscalzo J. The application of big data to cardiovascular disease: paths to precision medicine. *J Clin Invest*. 2020; 130(1): 29–38. doi: <https://doi.org/10.1172/jci129203>
15. Villarreal V, Samudio M. Plataforma para la autogestión de datos de pacientes hipertensos en Panamá. *I+D Tecnológico*. 2018; 14(1): 83–93.
16. Sakr S, Elgammal A. Towards a comprehensive data analytics framework for smart healthcare services. *Big Data Research*. 2016; 4(C): 44–58. doi: <https://doi.org/10.1016/j.bdr.2016.05.002>

17. Sánchez-Acevedo MA, Acosta-Chi ZA, Morales-Salgado MR. Cardiovascular risk detection through big data analysis. *J Big Data Anal Transp.* 2020; 5(2): 1-11.
18. Departamento Administrativo Nacional de Estadística. Censo Nacional de Población y Vivienda, 2018-Colombia. DANE; 2018. Disponible en: <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/censo-nacional-de-poblacion-y-vivenda-2018>
19. Secretaría de Salud, Gobernación del Valle del Cauca. Análisis de situación de salud Valle del Cauca. Gobernación del Valle del Cauca; 2019. Disponible en: <https://www.valledelcauca.gov.co/documentos/12672/asis-municipios-2019/>
20. Washburn RA, Smith KW, Jette AM, Janney CA. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol.* 1993; 46(2): 153–162. doi: [https://doi.org/10.1016/0895-4356\(93\)90053-4](https://doi.org/10.1016/0895-4356(93)90053-4)
21. Sommerville I. Ingeniería del software. Madrid: Pearson educación; 2005.
22. Burbeck S. Applications programming in smalltalk-80 (TM): How to use Model-View-Controller (MVC). Smalltalk-80 v2. ParcPlace. 1992; 5: 1–11.
23. Rondón Suárez LM. Calidad en el levantamiento de requerimientos en proyectos de software. Universidad Militar Nueva Granada; 2019. Disponible en: <https://repository.unimilitar.edu.co/handle/10654/31989#~:text=E1%20levantamiento%20de%20requerimientos%20es,con%20los%20requisitos%20del%20cliente>
24. López-Caballero CA, Cuello SR. Efectividad metodológica para el levantamiento de requerimientos de aplicaciones web. *Rev Ontare.* 2016; 4(2): 71–98.
25. Pérez-Virgen HL, Salamando-Mejía CA, Valencia-Ayala LS. Levantamiento de requerimientos basados en el conocimiento del proceso. *Rev Científica.* 2012; 2(16): 42–51. doi: <https://doi.org/10.14483/23448350.4022>
26. Hansen GW, James VH. Diseño y administración de bases de datos. Madrid: Prentice Hall. 1998.
27. Enríquez JG, Casas SI. Usabilidad en aplicaciones móviles. *Informes Científicos – UNOA.* 2013; 5(2): 25-47.
28. Paniagua LA, Bedoya RD, Mera C. Un método para la evaluación de la accesibilidad y usabilidad en aplicaciones móviles. *TecnoLógicas.* 2020; 23(48): 98-116. doi: <https://doi.org/10.22430/22565337.1553>
29. Brooke J. SUS – A quick and dirty usability scale. Usability evaluation in industry. London: Taylor and Francis; 1996. p. 198-194.
30. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. *Int J Human-Comp Interact.* 2008; 24(6): 574-594. doi: <https://doi.org/10.1080/10447310802205776>