Pontificia Universidad JAVERIANA Bogotá

editorial Pontificia Universidad JAVERIANA

David Barrera Ferro
Elena Valentina Gutiérrez
Andrés Felipe Osorio Muriel
Editores académicos

APLICACIONES DE INVESTIGACIÓN DE OPERACIONES EN SISTEMAS DE SALUD EN COLOMBIA



Fotografía de *Jose Gil* en Unsplash

Aplicaciones de investigación de operaciones en sistemas de salud en Colombia

Aplicaciones de investigación de operaciones en sistemas de salud en Colombia

DAVID BARRERA FERRO
ELENA VALENTINA GUTIÉRREZ
ANDRÉS FELIPE OSORIO MURIEL

EDITORES ACADÉMICOS

Pontificia Universidad
JAVERIANA
Bogotá

Pontificia Universidad JAVERIANA Bogotá | Facultad de Ingeniería

editorial Pontificia Universidad JAVERIANA

Reservados todos los derechos
© Pontificia Universidad Javeriana
© David Barrera Ferro (Pontificia Universidad Javeriana, Bogotá), Elena Valentina Gutiérrez (Universidad de Antioquia, Medellín), Andrés Felipe Osorio Muriel (Universidad Icesi, Cali), editores académicos

Primera edición: Bogotá, D. C., noviembre de 2020
ISBN (digital): 978-958-781-575-7
DOI: https://doi.org//10.11144/Javeriana.9789587815757
Impreso y hecho en Colombia
Printed and made in Colombia

Editorial Pontificia Universidad Javeriana
Carrera 7.ª n.° 37-25, oficina 1301
Edificio Lutaima
Teléfono: 320 8320 ext. 4205
www.javeriana.edu.co/editorial
editorialpuj@javeriana.edu.co
Bogotá, D. C.

Corrección de estilo:
Ella Suárez
Diagramación:
Kilka Diseño Gráfico
Diseño de cubierta:
Kilka Diseño Gráfico

MIEMBRO DE LA RED DE EDITORIALES UNIVERSITARIAS DE AUSJAL
www.ausjal.org

Pontificia Universidad Javeriana | Vigilada Mineducación. Reconocimiento como Universidad: Decreto 1297 del 30 de mayo de 1964. Reconocimiento de personería jurídica: Resolución 73 del 12 de diciembre de 1933 del Ministerio de Gobierno.

Pontificia Universidad Javeriana. Biblioteca Alfonso Borrero Cabal, S. J.
Catalogación en la publicación

Aplicaciones de investigación de operaciones en sistemas de salud en Colombia / editores académicos David Barrera Ferro, Elena Valentina Gutiérrez, Andrés Felipe Osorio Muriel. -- Primera edición. -- Bogotá : Editorial Pontificia Universidad Javeriana, 2020.
Incluye referencias bibliográficas.
ISBN:(digital) 978-958-781-575-7
1. Sistemas de salud - Colombia 2. Servicios de salud - Colombia 3. Investigación operativa en sistemas de salud 4. Atención primaria de la salud - Colombia 5. Programación (Matemáticas) 5. Procesos de optimización 6. Servicios médicos - Innovaciones tecnológicas 7. Servicios de salud - Innovaciones tecnológicas I. Barrera Ferro, David, editor académico II. Gutiérrez, Elena Valentina, editora académica III. Osorio Muriel, Andrés Felipe, editor académico IV. Pontificia Universidad Javeriana.
CDD 614 edición 21

inp 11/11/2020

Prohibida la reproducción total o parcial de este material sin la autorización por escrito de la Pontificia Universidad Javeriana. Las opiniones expresadas son responsabilidad exclusiva de los autores y las autoras y no comprometen a la Pontificia Universidad Javeriana.

Contenido

Prólogo. Aplicaciones de investigación de operaciones en sistemas de salud: nuevas tendencias

Juan G. Villegas

Capítulo 1. Investigación de operaciones en sistemas de salud: problemas de decisión y retos en Colombia

Elena Valentina Gutiérrez, David Barrera Ferro y Andrés Felipe Osorio Muriel

Capítulo 2. Localización de un centro médico especializado en la ciudad de Bogotá integrando el proceso analítico jerárquico y la programación matemática

Juan Esteban Paredes Bayona, Carlos Andrés Bolívar Polanía y William J. Guerrero

Capítulo 3. Evaluación de rutas de personal asistencial en el cuidado a la salud domiciliaria mediante programación matemática y simulación multiagente

Julián Alberto Espejo-Díaz, Eduyn López-Santana y William J. Guerrero

Capítulo 4. Mejora de los tiempos de espera en la sala de urgencias de un hospital público, por medio de la simulación de eventos discretos

Álvaro Junior Caicedo-Rolón, Deisy Paola Leal Hernández, Liceth Paola Ortiz López y Leonardo Rivera Cadavid

Capítulo 5. Impacto de la inoportunidad de ayudas diagnósticas e interconsultas en la duración de la estancia hospitalaria

Tatiana María Ceballos Acevedo, Paula Andrea Velásquez Restrepo y Juan Sebastián Jaén-Posada

Capítulo 6. Árboles de modelos aditivos generalizados de localización, escala y forma para la predicción de tiempo de estancia hospitalaria

Juan Camilo España Lopera, Olga Cecilia Úsuga Manco y Juan Sebastián Jaén Posada

Capítulo 7. Marco de trabajo para la vigilancia epidemiológica en Santander, bajo el paradigma *big data*

Yuly Andrea Ramírez Sierra y Henry Lamos Díaz

Capítulo 8. Metodología para la identificación y seguimiento del riesgo en salud de una cohorte poblacional atendida bajo el modelo Sermás

Edwin Alonso Álvarez Tobón, Paula Andrea Velásquez Restrepo, Marta Cecilia Ramírez Orrego y Alma Karina Rodríguez Quintero

Colaboradores

Prólogo. Aplicaciones de investigación de operaciones en sistemas de salud: nuevas tendencias

Leer el contenido de este libro que tienen en sus manos realmente emociona y leer cada uno de sus capítulos en detalle enseña e inspira. El esfuerzo que han hecho sus editores (Andrés Osorio Muriel, David Barrera y Valentina Gutiérrez) y los más de veinte autores es muy importante. Hasta ahora, las aplicaciones más visibles de la investigación de operaciones en nuestro medio han estado concentradas en la utilización que tradicionalmente se ha hecho de esta para planear y controlar sistemas eléctricos, logísticos y productivos. De esta manera, es una pieza única este libro, dedicado a las aplicaciones de la investigación de operaciones en sistemas de salud, con énfasis en el ámbito colombiano. Para constatar su valor basta con (h)ojear la tabla de contenido. La selección de capítulos da buena cuenta de la diversidad de aplicaciones, herramientas, alcances y retos que se puede tener cuando se estudian los sistemas de salud con las herramientas de análisis que nos provee la investigación de operaciones.

En este libro se reúnen aplicaciones de herramientas tan diversas como el análisis de decisión multicriterio, la programación matemática/optimización y la simulación. Esta última se utilizó (remarcablemente) en tres paradigmas principales: simulación con agentes, simulación de eventos discretos y dinámica de sistemas. Asimismo, mención especial tienen las aplicaciones de análisis de grandes volúmenes de datos (*big data*) y las herramientas de aprendizaje de máquina (*machine learning*), presentadas en dos de los capítulos. Estos últimos muestran cómo las nuevas tendencias de analítica (*business analytics*) e inteligencia artificial (*artificial intelligence*) confluyen y se suman a las herramientas más tradicionales de la investigación de operaciones.

Sin duda, ver los sistemas de salud desde la perspectiva de la investigación de operaciones impone nuevos retos que no son insalvables. Hay en los sistemas de salud desafíos y dilemas a los que quizás antes no nos veíamos enfrentados al trabajar en otras industrias o servicios. Desde dilemas éticos, causados por la naturaleza del servicio de salud como un derecho fundamental, bien sea porque en las decisiones que se modelan muchas veces están involucradas las vidas y el bienestar de los pacientes, o porque los objetivos de los sistemas de salud van más allá de la rentabilidad o la disminución de costos a los que estamos comúnmente acostumbrados (*v. g.*, mejorar la cobertura o lograr equidad).

Por otro lado, entender los sistemas de salud y las decisiones que en ellos se toman (para el tratamiento de los pacientes, para la planeación de las operaciones de los servicios o para la definición de políticas públicas) exige apropiarse

de conocimientos que son ajenos a la formación tradicional en investigación de operaciones. Por fortuna, esto último exige la conformación de equipos multidisciplinarios, con personal asistencial (médicos, enfermeras, instrumentadores, terapeutas, etc.), con epidemiólogos, con funcionarios del Estado, con expertos en sistemas de información, calidad y acreditación en salud, entre otros. En estos equipos es también importante la investigación de operaciones, pues los complementa aportando una visión cuantitativa, analítica e ingenieril que algunas veces no tienen.

Hay todavía algunas barreras que salvar, tal como lo resaltan los editores en su primer capítulo. Quizás la más importante sea lograr que los resultados de estos estudios se utilicen rutinariamente en las entidades promotoras de salud (EPS), las instituciones prestadoras de salud (IPS) y los entes reguladores, de control y generadores de política pública del Estado. Sin embargo, ver entre los coautores de algunos de los capítulos de este libro a funcionarios que trabajan en IPS o EPS es un ejemplo de que esto se puede lograr y que los equipos interdisciplinarios pueden alcanzar resultados valiosos para estas entidades.

Espero que este sea el primer paso (editorial) de una comunidad colombiana dedicada a la investigación de operaciones en salud. El primer capítulo del libro, escrito por los editores, da una muy buena retrospectiva de la evolución y consolidación de este tipo de investigación, la cual se empieza a ver en múltiples manifestaciones. Por ejemplo, con la participación nutrida y sostenida en conferencias nacionales, en las cuales la investigación de operaciones aplicada a sistemas de salud es ya un tema recurrente. Igualmente, la creación del Grupo de Trabajo en Salud en el seno de la Asociación Colombiana de Investigación Operativa es otra de estas manifestaciones.

Para terminar, estoy seguro de que los capítulos de este libro serán una fuente de inspiración, para buscar nuevas aplicaciones o para reproducir y extender las que aquí se presentan. Este libro es una invitación a construir lazos que junten a los expertos en investigación de operaciones del país con los decisores en el ámbito del sistema de salud. Todo esto con el fin de contribuir al desarrollo del país en un asunto tan fundamental para la sociedad como la salud de la población.

Juan G. Villegas
Profesor del Departamento de Ingeniería Industrial,
Facultad de Ingeniería, Universidad de Antioquia, Medellín, Colombia

Capítulo 1.
Investigación de operaciones en sistemas de salud: problemas de decisión y retos en Colombia

Elena Valentina Gutiérrez,[1]
David Barrera Ferro [2] *y Andrés*
Felipe Osorio Muriel [3]

[1] Departamento de Ingeniería Industrial
Universidad de Antioquia
Medellín, Colombia
elena.gutierrez@udea.edu.co

[2] Departamento de Ingeniería Industrial
Pontificia Universidad Javeriana
Bogotá, Colombia
barrera-o@javeriana.edu.co

[3] Departamento de Ingeniería Industrial
Universidad Icesi
Cali, Colombia
afosorio@icesi.edu.co

Resumen

Este capítulo brinda una perspectiva general del estado de las aplicaciones de investigación de operaciones desarrolladas por investigadores colombianos para apoyar decisiones en los sistemas de salud del país. El capítulo se presenta en el marco de un primer intento por proporcionar un espacio de difusión de las experiencias de investigación cuyo objetivo es mejorar la prestación de servicios de salud en el contexto colombiano. En particular, se revisan brevemente los libros disponibles en la literatura sobre el tema, se identifican las brechas de investigación existentes y luego estas últimas se contrastan con el registro de los trabajos presentados en dos conferencias colombianas.

Palabras clave: sistemas de salud, investigación de operaciones, Colombia

Introducción

La relevancia y uso de la investigación de operaciones para apoyar las decisiones en sistemas de salud han llamado la atención de los investigadores en Colombia desde mediados de la década del 2000. Por un lado, las dinámicas socioeconómicas que se han derivado de la legislación del servicio desde la aplicación de reformas como la Ley 100 de 1993 los han motivado al desarrollo de métodos y aplicación de técnicas de la investigación de operaciones, dirigidas a mejorar

la prestación del servicio. Por otro, las iniciativas de la comunidad internacional por reconocer y promover el uso de la investigación de operaciones en el área de la salud han evidenciado las oportunidades de trabajo y le han permitido a la comunidad en el país identificar caminos de acción más claros para estudiar dichos sistemas, generando retos tanto prácticos como metodológicos.

Los gestores y tomadores de decisiones en instituciones de salud en Colombia enfrentan retos asociados al diseño, a la operación, al control y al mejoramiento de los procesos de prestación de servicios de manera más efectiva y eficiente. Adicionalmente, la presión sobre los sistemas de salud ha aumentado no solo por el incremento de la demanda, sino también por el incremento de los costos en salud (Organisation for Economic Co-operation and Development [OECD], 2020). En Colombia, la cobertura del sistema de salud se ha incrementado del 23,5 % en 1993 al 96,6 % en el 2014, y dicho crecimiento se ha dado con mayor rapidez para las personas en el 20 % más pobre de la población (OECD, 2015). La evidencia de la crisis del sistema de salud en el país, así como las múltiples estadísticas epidemiológicas de la población (Bernal y Gutiérrez, 2012), han mostrado la necesidad de abordar los problemas y retos que enfrentan los tomadores de decisiones en Colombia con técnicas y métodos de la investigación de operaciones.

En el mundo, el uso de la investigación de operaciones para apoyar decisiones en sistemas de salud data de 1981 (Fries, 1981). Los problemas y métodos estudiados abarcan un amplio espectro que incluye las decisiones en diferentes servicios de salud o en diferentes horizontes de planeación, mediante variados métodos de investigación de operaciones y distintas etapas del ciclo de prestación del servicio. Las decisiones van desde el diagnóstico y cuidado de los pacientes como eje central del sistema, pasando por las decisiones de asignación y planeación de recursos, hasta el diseño de redes de prestación de servicios y su respectiva cobertura. Esta diversidad de problemas y métodos ha generado un incremento en las publicaciones en el campo (Brailsford y Vissers, 2011; Turner et al., 2010), así como el fortalecimiento de comunidades y grupos de trabajo en Europa mediante el Operational Research Applied to Health Services (ORAHS, 2020) y en Estados Unidos con el capítulo de salud del Institute for Operations Research and the Management Sciences (INFORMS) (Denton et al., 2020), y la aparición de revistas especializadas en la aplicación de la investigación de operaciones en los sistemas de salud. Entre las revistas más reconocidas se encuentran *Health Care Management Science*, *Operations Research for Health Care*, *Health Systems* e *IISE Transactions on Healthcare Systems Engineering*. En Colombia, publicaciones seriadas como la revista *Ingeniería y Universidad* y la *Revista Facultad de Ingeniería Universidad de Antioquia* han editado números especiales que incluyen la aplicación de métodos de investigación de operaciones en sistemas de salud.

La literatura científica demuestra también la relevancia y uso de la investigación de operaciones para apoyar decisiones en sistemas de salud. Las revisiones

y taxonomías más recientes para el estudio del área dan cuenta de la diversidad en problemas y métodos. Hulshof et al. (2012) presentan una de las revisiones más citadas y amplias en la literatura, en la que se reportan más de 450 publicaciones en el campo. En este trabajo, los autores proponen una taxonomía para la clasificación de los problemas de planeación en salud que diferencia seis tipos de servicios, tres horizontes principales de planeación y seis métodos de investigación de operaciones. Brailsford y Vissers (2011) plantean una estructura de análisis en dos dimensiones para reportar el desarrollo de la comunidad europea en el campo. En la primera dimensión de la estructura se consideran nueve etapas del ciclo de vida en servicios de salud y en la segunda se definen tres niveles de clasificación de los procesos en diferentes niveles en salud. Brandeau et al. (2004) también presentan uno de los libros más citados en el estudio de la investigación de operaciones en salud. Los autores presentan 32 capítulos clasificados en tres áreas principales: gestión de operaciones en salud, política pública y análisis económico y aplicaciones clínicas.

En Colombia, al sistema de salud lo regula el Ministerio de Salud y de Protección Social, a través de un sistema de seguridad social financiado por las contribuciones de los empleados y empleadores en el régimen contributivo, y por el Estado, para el régimen subsidiado. Esta estructura, definida por la Ley 100 de 1993 y generada por las reformas a los sistemas de salud comunes en Latinoamérica, ha hecho que el sistema en Colombia adopte una filosofía de seguridad social basada en el aseguramiento de los empleados mediante instituciones privadas, en un esquema de competencia de mercado (Esteves, 2012). Ante este panorama, la estructura de prestación de servicios difiere en parte de las estructuras comunes de prestación de servicios, sobre todo en Europa, y parcialmente de la estructura en Estados Unidos. Sin embargo, y a pesar de dichas diferencias, los gestores y tomadores de decisiones en sistemas de salud en el país enfrentan retos similares a los encontrados con regularidad en la literatura: incrementos en los requerimientos de servicios de salud, con recursos limitados y cada vez más contraídos.

Frente a dicho contexto, la comunidad de investigación de operaciones en Colombia ha abordado diferentes problemas de decisión, con diferentes metodologías. Dentro de las primeras actividades de reunión de la comunidad hay pocos registros iniciales de aplicaciones en el campo. En la primera versión que hubo de una sociedad de investigación de operaciones en el país se realizó un primer congreso en la Universidad Nacional de Colombia (sede Medellín), en septiembre del 2001. A pesar de que en dicha oportunidad de los veintidós trabajos presentados ninguno correspondía a aplicaciones en salud, en septiembre del 2008, esa misma comunidad organizó el que hasta ahora ha sido el único Congreso Latino-Iberoamericano de Investigación Operativa (Claio) realizado en Colombia. En el Claio 2008 hubo siete sesiones dedicadas a sistemas de salud y más de veinticinco trabajos presentados en el congreso estudiaron aplicaciones de investigación de

operaciones a diferentes problemas en sistemas de salud. Puede decirse que, a partir del 2008, la comunidad colombiana de investigación de operaciones empezó a dedicarle más esfuerzo a las aplicaciones en el campo de estudio.

Desde el 2008, la comunidad en Colombia ha trabajado en un conjunto amplio de problemas de investigación de operaciones en sistemas de salud. En noviembre del 2018 se realizó en la Pontificia Universidad Javeriana (sede Bogotá) el primer encuentro del Health Care Engineering Days. En dicho encuentro se presentó un recuento breve de los trabajos adelantados en la comunidad (Gutiérrez et al., 2018) y se contó con la participación de investigadores de dieciséis instituciones, quienes compartieron sus experiencias en el campo. Con este encuentro se identificó que existe una masa crítica creciente en el país y se evidenció la necesidad de generar un espacio de difusión y discusión para que la comunidad comparta sus experiencias y se abran espacios de colaboración. De igual forma, con el encuentro se sentaron las bases para formalizar las actividades de la comunidad y presentar la propuesta de un Grupo de Trabajo en Salud ante la Asociación Colombiana de Investigación Operativa (Asocio).

De este modo, el objetivo principal de este libro, titulado *Aplicaciones de investigación de operaciones en sistemas de salud en Colombia* (*ApIOSS-Col*) es divulgar las experiencias de investigación que han desarrollado aplicaciones de técnicas de la investigación de operaciones para apoyar decisiones y mejorar la prestación de servicios de salud en el contexto colombiano. De igual forma, este libro pretende generar un espacio de difusión para identificar experiencias exitosas y posibles oportunidades de colaboración entre la comunidad de investigación de operaciones del país y el sector de la salud. El resto de este capítulo se encuentra organizado de la siguiente forma. En la siguiente sección se hace una revisión de libros de investigación de operaciones dedicados al estudio de problemas en sistemas de salud. Luego, se analizan los problemas y métodos estudiados por la comunidad y dados a conocer en las ponencias de los congresos de la Asocio y del Congreso Internacional de Industria y Organizaciones. En la cuarta sección se presenta la estructura del libro, brindando una breve reseña de cada uno de los capítulos incluidos. Finalmente, se concluye el capítulo.

Literatura de investigación de operaciones en salud

En esta sección se revisan diferentes libros que abordan aplicaciones de investigación de operaciones en sistemas de salud. Este tipo de publicaciones tuvo sus inicios hacia 1980; sin embargo, se ha evidenciado un crecimiento en la cantidad de libros publicados sobre el tema. Algunos de los últimos años abordan temas específicos en salud, como el flujo de pacientes y las emergencias; otros involucran otras disciplinas, como la gerencia de operaciones y la analítica de datos, y

otros siguen mostrando aplicaciones variadas específicamente de investigación de operaciones en sistemas de salud.

Los primeros libros los publicaron Boldy (1981) y Fries (1981). El primero es una revisión de aplicaciones, algunas de las cuales presentan casos detallados; mientras que en otros se mencionan las técnicas sin aplicaciones específicas. Estas se dividen en estratégicas y tácticas, dependiendo del horizonte de tiempo analizado. Es importante resaltar que el libro está orientado a una audiencia en el área de investigación de operaciones; sin embargo, el autor también busca motivar a administradores y tomadores de decisiones del sector de la salud. El segundo caso, el de Fries (1981), recoge aplicaciones de investigación de operaciones en sistemas de salud. El libro analiza 314 artículos publicados sobre el tema. Cada artículo es descrito brevemente y se agrupan en grandes temas que incluyen: pronósticos, localización de instalaciones, ocupación de hospitales, bancos de sangre, inventario de suministros hospitalarios, entre otros. Para seleccionar los artículos, el autor tuvo en cuenta aplicaciones puras de investigación de operaciones, y no se consideraron técnicas como análisis de costo-efectividad y análisis económicos.

En Duckett (1987) se relacionan diferentes temas, como la gestión de inventarios, la teoría de colas, la simulación y asignación, entre otros. Cada tema tiene un tutorial y se describen ejemplos de posibles aplicaciones. En algunos casos, el autor las menciona, pero el libro está enfocado en describir y explicar las técnicas, en lugar de mostrarlas de manera profunda. Es importante resaltar que, dada la fecha de publicación del libro, pueden estar obsoletas muchas de las técnicas mencionadas.

Después de estos libros, hay un salto en el tiempo y en el 2004 aparece el libro de Brandeau et al. (2004), que se convierte en uno de los más referenciados en esta área. Se estructura en tres partes: gerencia de operaciones en salud, análisis económico y políticas públicas y aplicaciones clínicas. El libro reseña no solo las aplicaciones más relevantes en ese momento, sino también los futuros retos en esta área de conocimiento. Además, son diversas, desde planeación de capacidad, modelos de detección en diabetes y retinopatía hasta modelos de optimización de radioterapias.

En la última década se ha evidenciado un crecimiento importante en el número de textos publicados sobre el tema. En el 2011, la Organización Mundial de la Salud dio a conocer el libro escrito en inglés titulado *Prioridades en investigación de operaciones para mejorar el cuidado y control de la tuberculosis* (World Health Organization, 2011). En este se presentan los resultados de un proceso en el que se reunió a un grupo de expertos en investigación de operaciones con el fin de identificar oportunidades en la lucha contra la tuberculosis. Así se identificaron cinco áreas en las que se requieren estrategias óptimas para actividades de control de tuberculosis: 1) acceso, 2) detección y diagnóstico, 3) colaboración sostenible con proveedores de servicios de salud en tuberculosis, 4) prevención de tuberculosis en personas con VIH y tratamiento conjunto VIH y tuberculosis

más acceso y entrega de tratamientos y 5) creación de capacidad en investigación de operaciones.

El libro editado por Tanfani y Testi (2012) contiene trece capítulos que enmarcan diferentes técnicas —como análisis envolvente de datos, simulación, optimización y análisis multicriterio— empleadas en una gran variedad de problemas. Una de las principales características del libro es que, junto con la aplicación, se explican de modo claro los conceptos metodológicos y teóricos de las técnicas usadas, al igual que los impactos y los beneficios. También cubre diferentes tipos de aplicaciones, incluyendo administración de fuerza laboral, programación de quirófanos, rutas hospitalarias y problemas de asignación.

Un año después aparecieron los textos de Hall (2013) y Zaric (2013). En el primer caso, el libro está enfocado en el uso de técnicas para mejorar el flujo de pacientes en los sistemas de salud. El objetivo del libro es presentarle a la audiencia un conjunto de técnicas y estrategias que el personal médico y los tomadores de decisiones pueden emplear para reducir las demoras en los sistemas de salud. De esta manera, se centra en áreas como la congestión de los sistemas, los programas de citas y las herramientas de modelado. Uno de los aspectos relevantes de este libro es que contiene una sección dedicada a la implementación de las soluciones obtenidas. Por otra parte, en los dieciocho capítulos del libro de Zaric (2013) se encuentran diferentes aplicaciones de investigación de operaciones en el diseño de políticas en sistemas de salud, por ejemplo, optimización, cadenas de Markov, simulación, entre otras, que se utilizan para apoyar el diseño de políticas en diferentes países (*e. g.* Estados Unidos y Canadá). En sus diferentes secciones se incluyen una revisión general de aplicaciones de investigación de operaciones en políticas de salud, aplicaciones enfocadas en problemas específicos (como operaciones, VIH y farmacéuticas) y lineamientos sobre la construcción de modelos de políticas en salud.

En los dos volúmenes de su libro, Mustafee (2016a, 2016b) recoge diferentes contribuciones, en especial respecto al tema de planeación de emergencias; esto incluye técnicas de preparación, planeación y respuesta de accidentes y emergencias en salud pública. El primer volumen aborda el problema desde un enfoque operativo; mientras que en el segundo volumen el enfoque es completamente estratégico. Los libros se estructuran considerando tres perspectivas: las técnicas de investigación de operaciones, el área de aplicación y el contexto de la aplicación.

El libro de Kahraman y Topcu (2018) contiene aplicaciones en diferentes países como Turquía, Estados Unidos, Canadá, entre otros. Una de las secciones incluye una taxonomía para clasificar las aplicaciones de investigación de operaciones en salud. La estructura recorre algunos elementos del flujo de los pacientes, esto incluye preparación, diagnóstico y tratamiento. De igual forma, esto se

complementa con secciones donde se analizan los sistemas, las operaciones y los causantes de las demoras en los procesos de atención.

Otros libros como los de Denton (2013), Dai y Tayur (2018) y Yang y Lee (2016) incluyen temas de aplicaciones en investigación de operaciones en salud; sin embargo, se enfocan en disciplinas relacionadas. En el caso de Denton (2013), se dirige hacia la gerencia de operaciones en salud; mientras que Dai y Tayur (2018) y Yang y Lee (2016) se centran en la analítica de datos. Finalmente, Brailsford y Harper (2007) presentan una colección de contribuciones del Grupo de Trabajo en Salud de la Asociación Europea de Investigación de Operaciones (ORAHS, 2020).

En materia de libros de texto, orientados a estudiantes de pregrado o personal dedicado a la aplicación de investigación de operaciones en sistemas de salud, vale la pena citar el de Ozcan (2009), el cual, en quince capítulos, recorre los principales problemas de la gestión de operaciones, como pronósticos, localización de instalaciones, distribución de planta, toma de decisiones, gerencia de proyectos e inventarios, entre otros. Cada capítulo se desarrolla de forma clara a través de múltiples ejemplos y al final se dejan ejercicios propuestos para los estudiantes. Por su parte, Clarkson et al. (2009) presentan una guía básica que permite elegir entre las diversas técnicas y herramientas de la investigación de operaciones la más apropiada según el contexto. Este libro resulta particularmente útil para los tomadores de decisiones en el ámbito de los sistemas de salud, los cuales en muchos casos tienen un conocimiento limitado de la diversidad de enfoques que la investigación de operaciones ofrece.

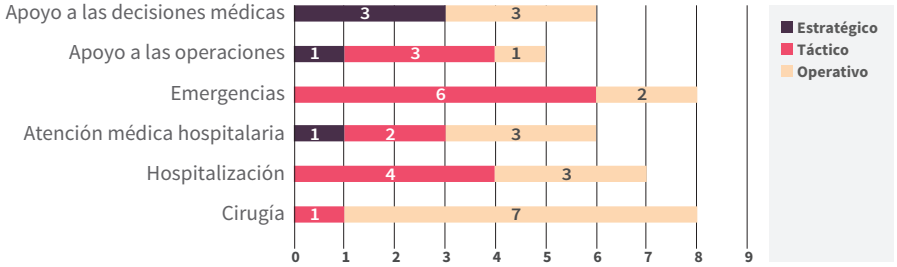
Como se evidencia, existen múltiples aplicaciones; sin embargo, es importante resaltar que no se encontraron libros enfocados en países en desarrollo. En ese orden de ideas, este libro contribuye a entender y mejorar los sistemas de salud teniendo en cuenta las particularidades del sistema de salud colombiano.

Comunidad de investigación de operaciones en salud en Colombia

Con el objetivo de dar un contexto a los trabajos incluidos en el libro, en esta sección se analizan los problemas y los métodos estudiados en algunas aplicaciones locales. Se incluyeron las ponencias presentadas en el Congreso Colombiano de Investigación de Operaciones (versiones 2015, 2017 y 2019) y en el Congreso Internacional Industria y Organizaciones (versiones 2016, 2017 y 2018). En total, se encontraron 42 trabajos. Un total de veintinueve ponencias abordan problemas de planeación y se agruparon usando las categorías definidas por Hulshof et al. (2012). Además, seis trabajos corresponden a modelos de apoyo a decisiones médicas y cinco más analizan problemas de apoyo a las operaciones. Finalmente, en dos ponencias, los autores presentan revisiones de la literatura (Duque Uribe y Sa-

rache Castro, 2017; Ruiz Orjuela y Adarme Jaimes, 2019). La figura 1 resume la distribución de cuarenta trabajos en dichas categorías, considerando el horizonte de planeación de la decisión apoyada. A pesar de no ser una revisión sistemática de la literatura, dos conclusiones son de interés: primero, existe un mayor balance entre las decisiones tácticas y operativas que el encontrado en estudios hechos en otros contextos (Brailsford y Vissers, 2011); segundo, se evidencia un crecimiento del interés de la comunidad local en el tema, comparado con lo encontrado por Velasco et al. (2012).

Figura 1. Clasificación de trabajos con base en la taxonomía propuesta por Hulshof et al. (2012)



Fuente: elaboración propia.

En servicios de cirugía, seis ponencias (86 %) resuelven la versión determinística del problema de programación de procedimientos electivos (Arévalo et al., 2017; Aristizábal Cataño et al., 2016; Díaz et al., 2015; Estupiñán et al., 2015; Noguera et al., 2017; Saavedra et al., 2015). En estos trabajos, los autores asumen que se conoce la lista de procedimientos que deben ejecutarse durante el horizonte de planeación y su respectiva duración; no obstante, a pesar de reportar colaboración con hospitales e incorporar las particularidades del contexto, los resúmenes no brindan información que permita concluir sobre los niveles de implementación de los resultados. Respecto a los métodos, cuatro trabajos proponen reglas de despacho (Aristizábal Cataño et al., 2016; Díaz et al., 2015; Estupiñán et al., 2015; Saavedra et al., 2015) y dos usan estrategias de búsqueda local para mejorar la calidad de las soluciones iniciales (Arévalo et al., 2017; Noguera et al., 2017). Para este problema, no fue posible encontrar trabajos que estudien las decisiones estratégicas o tácticas ni que consideren la incertidumbre relacionada con los parámetros. Finalmente, Avella-Centeno et al. (2016) estudian el problema de la programación de enfermeras en un servicio quirúrgico de alta complejidad. Los autores proponen el uso de una técnica híbrida de simulación-optimización que modela la variabilidad asociada a la duración de los procedimientos quirúrgicos y el impacto que esta tiene en la carga de trabajo.

En los servicios de emergencia, las decisiones de localización y relocalización de ambulancias concentran la mayor parte de la atención en las aplicaciones locales (43 %). Rodríguez et al. (2015) analizan los registros históricos de dos años de funcionamiento de un servicio de emergencias en Medellín y dan a conocer un modelo de simulación de eventos discretos para evaluar diferentes políticas de relocalización de ambulancias. Castañeda et al. (2015) proponen integrar un modelo de optimización multiobjetivo y un modelo de simulación de eventos discretos para tomar decisiones de localización. Con esta metodología es posible cuantificar los sacrificios entre diferentes funciones-objetivo e incorporar la incertidumbre de la demanda y los tiempos de servicio. Finalmente, Rico Carrillo et al. (2018) proponen un modelo de programación lineal para encontrar la ubicación óptima de ambulancias marítimas medicalizadas en la ciudad de Cartagena. En esta categoría, otras decisiones menos estudiadas incluyen el uso de grandes cantidades de información para caracterizar usuarios (Felizzola et al., 2017) o procesos (Aguirre y Rodríguez, 2017), la gestión de inventarios (Carvajal et al., 2015) y la selección del hospital destino (Rojas et al., 2016).

En los servicios de atención médica domiciliaria se encontró también que la decisión más estudiada es el ruteo del personal encargado de la prestación de los servicios. Cuatro aplicaciones (57 %) resuelven la versión determinística de dicho problema (Espejo et al., 2019; López-Santana et al., 2016; Manrique y Gutiérrez, 2019; Vásquez Echavarría y Gutiérrez, 2018). A pesar de que la mayoría de los trabajos propone el uso de métodos exactos, Vásquez Echavarría y Gutiérrez (2018) desarrollaron una herramienta computacional para facilitar la toma decisiones y discutieron sobre las implicaciones de las decisiones tácticas en el problema operativo. Además, existen trabajos que buscan caracterizar la forma en que se toman decisiones de asignación de recursos. Cortés y Gutiérrez (2016) proponen un modelo de medición de madurez de capacidades logísticas que después se usó en el Valle de Aburrá (Cortés et al., 2017a). Finalmente, Cortés et al. (2017b) proponen un modelo de programación entera mixta para el problema de la zonificación, considerando el balance de cargas de trabajo, la contigüidad y las medidas de distancia.

En los servicios hospitalarios, dos temas concentran el 71 % de los trabajos. Por un lado, tres ponencias estudian la gestión de inventarios usando simulación (Amaya Bermúdez et al., 2018; Martínez Sánchez et al., 2018; Moscoso Moreno y Cañón Muñoz, 2016). Por otro, dos ponencias estudian la planeación de turnos de enfermería. Moscoso Moreno y Cañón Muñoz (2016) proponen el uso de simulación de eventos discretos para evaluar diferentes alternativas de distribución de dosis unitarias. Adicionalmente, Martínez Sánchez et al. (2018) y Amaya Bermúdez et al. (2018) usan la dinámica de sistemas para analizar la capacidad instalada y la distribución de medicamentos oncológicos. Giedelman et al. (2015) presentan un modelo matemático para planear turnos de enfermería y

Velásquez et al. (2017) desarrollaron un aplicativo basado en un modelo de redes. Por último, se encontraron trabajos en rediseño de procesos (Buitrago et al., 2015) y análisis de la estancia hospitalaria (Ceballos et al., 2015).

Una categoría de interés, que no fue incluida en la clasificación propuesta por Hulshof et al. (2012), corresponde a los trabajos dedicados a estudiar los procesos de apoyo a las operaciones. En esta se encuentran la gestión de inventarios, la mejora de procesos, el ruteo de vehículos y la localización de instalaciones. Piraban y Labadie (2017) proponen un modelo de programación estocástica para diseñar políticas de inventario y de productos hemocomponentes, para un entorno hospitalario. En la misma línea, Poveda et al. (2017) diseñaron una metodología de simulación-optimización para el diseño de políticas coordinadas entre un banco de sangre y una red de hospitales públicos. Por otro lado, Marín et al. (2017) propusieron mejoras al proceso de facturación de una clínica y evaluaron el impacto de la propuesta usando una simulación de eventos discretos. Así mismo, Bueno et al. (2017) estudiaron el proceso de recolección de residuos sólidos en una red de veintiocho hospitales usando heurísticas constructivas. Finalmente, Mayorga et al. (2017) resolvieron el problema de la localización de centros de distribución para una empresa farmacéutica.

Por último, en una última categoría se agruparon los modelos que brindan soporte a las decisiones médicas. Se encontró que el 83 % de los estudios son desarrollos teóricos, en etapa de validación. En consecuencia, los autores no reportan colaboración con una entidad del sector salud. Por un lado, tres ponencias estudiaron decisiones de tratamiento (Carranza et al., 2015; Coy et al., 2015; Jiménez Lozano y Jiménez Muñoz, 2017). Coy et al. (2015) presentaron una extensión de un proceso de decisión markoviano para apoyar la toma de decisiones en el tratamiento de cáncer de seno. Carranza et al. (2015) propusieron una herramienta para seleccionar el mejor tipo de examen diagnóstico en cáncer de cuello uterino, y Jiménez Lozano y Jiménez Muñoz (2017) analizaron el potencial de las aplicaciones de programación lineal en la planeación de tratamientos oncológicos. Por otro lado, tres estudios apoyaron decisiones de política pública (Angulo et al., 2017; Arboleda et al., 2015; Camacho, 2019). Arboleda et al. (2015) plantearon un proceso de decisión markoviano parcialmente observable para evaluar políticas del examen diagnóstico del cáncer de cuello uterino, Angulo et al. (2017) analizaron la costoefectividad de los tratamientos de la misma enfermedad y Camacho (2019) caracterizó la operación de redes rurales de atención en salud materno-perinatal.

Estructura del libro *ApIOSS-Col*

La literatura sobre el tema ha evidenciado el uso de métodos de la investigación de operaciones para abordar problemas de decisión en el área de la salud. Sin embargo, existe la necesidad de identificar las experiencias de investigación que han

desarrollado aplicaciones de técnicas en este campo para apoyar decisiones y para mejorar la prestación de servicios en el contexto colombiano. De este modo, los objetivos del libro son responder a dicha necesidad e identificar las oportunidades de investigación en el contexto.

Con base en la clasificación propuesta por Brandeau et al. (2004), el libro cubre dos principales áreas de estudio: 1) gestión de operaciones en salud y 2) política pública y análisis económico. La mayoría de los capítulos están orientados a la primera área. Así, cada uno presenta el problema de investigación, los métodos utilizados y los resultados obtenidos. Esta estructura se enmarca en una aplicación en el contexto colombiano, de modo que distintos actores del sistema, como investigadores y tomadores de decisiones, puedan acceder a este material.

El libro cubre temas diversos en la primera área de estudio, como la localización y el diseño de redes de centros de atención en salud, la planeación de servicios de atención médica domiciliaria, la mejora de tiempos de atención en salas de urgencias, el impacto de las ayudas diagnósticas y la estancia hospitalaria. En la segunda área de estudio se presenta un marco para la vigilancia epidemiológica y una metodología para la gestión de riesgo en salud. A pesar de la diversidad en las aplicaciones, cada uno de los capítulos evidencia cómo el uso de la investigación de operaciones puede contribuir al mejoramiento de la prestación de servicios de salud en Colombia.

Gestión de operaciones en salud

La primera sección del libro incluye capítulos que describen la aplicación de modelos de investigación de operaciones en la gestión de operaciones en salud.

En el capítulo 2, Juan Esteban Paredes Bayona, Carlos Andrés Bolívar Polanía y William J. Guerrero describen una propuesta metodológica para la localización de un centro médico especializado en la ciudad de Bogotá, a través de la integración del proceso analítico-jerárquico y la programación matemática. El trabajo está inspirado en el problema de cobertura de conjuntos, y los autores integran además herramientas de administración e ingeniería. Con base en su aplicación en el contexto de la ciudad de Bogotá, los resultados permiten estimar el impacto de los métodos propuestos, en términos de cobertura en los servicios médicos de la ciudad.

En el capítulo 3, Julián Alberto Espejo-Díaz, Eduyn López-Santana y William J. Guerrero presentan un modelo matemático que determina las rutas óptimas del personal asistencial en el cuidado domiciliario de la salud. Para ello, los autores proponen un método que combina la programación matemática y la simulación basada en agentes, el cual busca minimizar los tiempos de transporte del personal médico y de espera de los pacientes. Este método se aplicó en un caso de estudio real de un proveedor del servicio de cuidado a la salud domiciliaria en la

ciudad de Bogotá, y los resultados muestran un *trade-off* entre los niveles de servicio y los costos operativos al optimizar las rutas con mayor o menor frecuencia.

En el capítulo 4, Álvaro Junior Caicedo-Rolón, Deisy Paola Leal Hernández, Liceth Paola Ortiz López y Leonardo Rivera Cadavid describen la mejora de tiempos de espera en la sala de urgencias de un hospital público fronterizo, por medio de la simulación de eventos discretos. Con base en el análisis del proceso, los autores evalúan el impacto de posibles mejoras como la eliminación de un proceso de preadmisiones y la agregación de un médico para consulta general. El modelo evidencia que dichas mejoras permiten, entre otras, determinar el indicador de pacientes que salen del sistema sin ser vistos por un médico, reducir el tiempo promedio de espera de una consulta médica, reducir la ocupación del personal auxiliar de enfermería y aumentar el egreso de pacientes.

En el capítulo 5, Tatiana María Ceballos Acevedo, Paula Andrea Velásquez Restrepo y Juan Sebastián Jaén-Posada estudian el impacto de la inoportunidad de ayudas diagnósticas e interconsultas en la duración de la estancia hospitalaria. Para ello, los autores presentan una metodología que mide el impacto en días de estancia, que tienen los retrasos en la entrega de dichas ayudas en la eficacia y la eficiencia de la ronda médica de un grupo relacionado de diagnóstico específico. Los autores desarrollaron un modelo de dinámica de sistemas, el cual permite identificar cuáles políticas de sincronización entre los tiempos de servicio de ayudas diagnósticas, interconsultas y ronda médica podrían reducir la estancia hospitalaria de pacientes del grupo relacionado estudiado. Dicha mejora implica, además, disminuir el promedio de la estancia hospitalaria de tal grupo.

En el capítulo 6, Juan Camilo España Lopera, Olga Cecilia Úsuga Manco y Juan Sebastián Jaén-Posada desarrollan una metodología de árboles de modelos aditivos generalizados de localización, escala y forma, en conjunto con la técnica de árboles de clasificación y regresión, para predecir tiempos de estancia hospitalaria. Los autores aplicaron la metodología en tres casos de estudio de hospitales que se diferencian en su país de origen, en el número de observaciones y en la cantidad y naturaleza de las variables explicativas. Los resultados muestran que el nivel de ajuste de los modelos varía según las características de los casos de estudio, pero se evidencia también que el desempeño del modelo generado con la técnica de árboles de decisión supera los resultados que se obtienen en otros modelos, en particular en los escenarios logrados con regresión lineal.

Política pública y análisis económico

La segunda sección del libro incluye dos capítulos sobre el uso de modelos de investigación de operaciones para el análisis de política pública y el análisis económico de sistemas de salud en Colombia.

En el capítulo 7, Yuly Andrea Ramírez Sierra y Henry Lamos Díaz construyen un marco de trabajo para la vigilancia epidemiológica en Santander, bajo

el paradigma de la *big data*. Para ello, y según las etapas del proceso de descubrimiento de conocimiento en bases de datos, los autores adoptaron técnicas de minería de datos y de aprendizaje automático, con el fin de apoyar la toma de decisiones en salud pública. La metodología propuesta se aplicó al estudio de caso del dengue, y se evidenciaron las ventajas de integrar la analítica descriptiva y la predictiva.

En el capítulo 8, Edwin Alonso Álvarez Tobón, Paula Andrea Velásquez Restrepo, Marta Cecilia Ramírez Orrego y Alma Karina Rodríguez Quintero reseñan una metodología para la identificación y el seguimiento del riesgo en salud de una cohorte poblacional atendida bajo el modelo de pago global prospectivo. La metodología permitió identificar, clasificar, definir y estratificar grupos, según el riesgo, mediante modelos descriptivos y predictivos. Los autores emplearon la metodología a una cohorte poblacional y los resultados muestran la identificación de pacientes que se podrían beneficiar de intervenciones preventivas para reducir eventos no deseados. De igual forma, los resultados muestran que es posible tener una mejor planificación y gestión de los recursos.

Conclusiones

Este capítulo introductorio ofrece una perspectiva general del estado de las aplicaciones de investigación de operaciones desarrolladas por investigadores colombianos para apoyar decisiones en sistemas de salud del país. Para ello, exploró tres fuentes de información: 1) una revisión de libros en el tema, 2) una revisión de trabajos desarrollados en Colombia a través de memorias de conferencias y 3) una descripción de cada capítulo de este libro. De acuerdo con el análisis de cada sección, se presentan dos conclusiones principales, en términos del enfoque de libros existentes y el número y enfoque de aplicaciones divulgadas por la comunidad. Se identificaron también para cada punto los retos que enfrenta la comunidad y la necesidad de generar redes de colaboración.

Primero, la mayoría de los libros encontrados en la literatura científica documentan experiencias en países con un alto grado de desarrollo económico e industrial, que mantienen sólidos sistemas de salud pública. Este hecho hace que las problemáticas, los recursos físicos y tecnológicos e, incluso, la cultura de pacientes y personal de salud tengan características distintas cuando se comparan con países en vías de desarrollo. Por esto, es necesario centralizar los esfuerzos para promover las aplicaciones de investigación de operaciones en el sistema de salud colombiano. En el mundo existen comunidades y grupos de trabajo de investigación de operaciones en salud que, a través del fortalecimiento de su masa crítica de investigadores, publicaciones, conferencias y asociaciones, han evidenciado madurez y oportunidades de colaboración en sistemas de salud. En Colombia, es necesario continuar el esfuerzo en la creación de redes para promover el

tema. La evidencia de los congresos muestra el desarrollo de varios tutoriales y la diversidad en las aplicaciones; sin embargo, es necesario fortalecer los lazos de colaboración entre los investigadores en el tema. En el mismo sentido, es necesario motivar al personal médico y a los gestores de instituciones prestadoras de servicios de salud a conocer y participar de la aplicación de investigación de operaciones en sus decisiones.

Segundo, el número de publicaciones de aplicaciones de investigación de operaciones en sistemas de salud en Colombia ha evidenciado una tendencia creciente en los últimos diez años. Los trabajos de la comunidad nacional muestran un mayor balance entre las decisiones tácticas y operativas que el encontrado en estudios hechos en otros contextos. No obstante, se identificó que es necesario que los practicantes e investigadores en el tema también colaboren con entidades e instituciones gubernamentales, de modo que los resultados de las investigaciones contribuyan a la definición de políticas públicas, y que el impacto trascienda las decisiones tácticas y operativas, y alcance los niveles estratégicos de planeación de salud en Colombia. De igual forma, la evidencia de los trabajos presentados en conferencias muestra una tendencia creciente, pero se ha encontrado que es limitada la visibilidad de dichas memorias. El reto que enfrentan los investigadores, en este sentido, es convertir dichos trabajos en publicaciones científicas (*i. e.*, artículos en revistas indexadas y volúmenes editados como este), de tal forma que los desarrollos tengan más visibilidad y circulación y, por ende, más oportunidades de reproducirse en otros contextos.

Este libro es un primer intento por proporcionar un espacio formal de difusión de las experiencias que ha tenido la comunidad en Colombia en la aplicación de investigación de operaciones en sistemas de salud nacionales. Al igual que la comunidad internacional, aunque en tiempos y ritmos diferentes, la comunidad de investigadores y gestores de sistemas de salud en el país ha mostrado un interés por el estudio y la aplicación de métodos de la investigación de operaciones para el mejoramiento de la prestación de servicios de salud. Por ello, en este libro se pretende divulgar las experiencias de investigación desarrolladas en Colombia, con miras además a identificar oportunidades de trabajo y redes de investigación. El proceso de convocatoria, evaluación y edición del libro se ha realizado en el contexto del recientemente formalizado Grupo de Trabajo en Salud (GTS) de Asocio. Los miembros del GTS, así como la comunidad de Asocio, han desarrollado un conjunto de actividades, cursos, tutoriales, encuentros y congresos, entre otros, que han permitido el reconocimiento de los trabajos e investigadores en el país. De este modo, esperamos que este libro cumpla su objetivo y promueva el trabajo colaborativo en la comunidad, y el desarrollo y madurez del área en nuestro país.

Referencias

Aguirre, S. y Rodríguez, C. (2017). *Mejoramiento del servicio de urgencias en una red de clínicas con minería de procesos* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Eafit, Colombia.

Amaya Bermúdez, C. D., Barreto Benítez, M. F. y Zamora Aguas, J. P. (2018). *Análisis de capacidades de almacenamiento de medicamentos oncológicos en una IPS de Bogotá* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad Nacional de Colombia, Colombia.

Angulo, D., Mura, I. y AkhavanTabatabaei, R. (2017). *Análisis de costo-efectividad de las políticas públicas colombianas contra el cáncer de cuello uterino* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Eafit, Colombia.

Arboleda, J., Namen, I., Akhavan-Tabatabaei, R. y Gel, E. (2015). *Un enfoque POMDP a políticas de detección primarias para el cáncer de cuello uterino en Colombia* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Arévalo, S., Palma, J., Abdallah, R. N. y Rojas, J. A. (2017). *Eficiencia en la planeación de cirugías a través de la aplicación de dos métodos heurísticos* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones. Universidad Eafit, Colombia.

Aristizábal Cataño, M. C., Parra Flórez, L. A., Millán González, A. T., Román Londoño, M. A., Bolaños Carmona, M. A. y Rojas Trejos, C. A. (2016). *Aplicación de reglas de despacho para la programación de cirugías. Caso estudio: Hospital del Norte del Valle del Cauca* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad del Valle, Colombia.

Avella-Centeno, N., Moreno, N., Riaño, C. D., Barrera, D. y Suárez, D. R. (2016). *Herramienta computacional para la programación de turnos de enfermería en un servicio quirúrgico de alta complejidad* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad del Valle, Colombia.

Bernal, O. y Gutiérrez, C. (2012). *La salud en Colombia*: *logros, retos y recomendaciones*. Ediciones Uniandes.

Boldy, D. (1981). *Operational research applied to health services*. Croom Helm.

Brailsford, S. y Harper, P. (Eds.). (2007). *Operational research for health policy: Making better decisions*. Documento procedente de la 31st Annual Conference of the European Working Group on Operational Research Applied to Health Services, Peter Lang.

Brailsford, S. y Vissers, J. (2011). OR in healthcare: A European perspective. *European Journal of Operational Research, 212*(2), 223-234.

Brandeau, M. L., Sainfort, F. y Pierskalla, W. P. (Eds.). (2004). *Operations research and health care: A handbook of methods and applications* (vol. 70). Kluwer Academic Publishers.

Bueno, E., Murillo, A. y Manyoma, P. (2017). *Ruteo de vehículos para la recolección de residuos sólidos hospitalarios peligrosos considerando ventanas de tiempo* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad de Antioquia, Colombia.

Buitrago, M., Varela, L. y Perea, N. (2015). *Aplicación del* lean manufacturing *y la simulación de eventos discretos en el análisis de procesos de prestación de servicios de salud* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Camacho, M. (2019). *Caracterización de los procesos de integración institucional para la planeación y operación de redes rurales de atención en salud materno perinatal en Cundinamarca* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Industrial de Santander, Colombia.

Carranza, V., Wilches, M. y Akhavan-Tabatabaei, R. (2015). *Desarrollo de una herramienta de apoyo a la decisión para el diagnóstico de cáncer cervical* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Carvajal, V., Maya, P. y Osorno, G. (2015). *Modelo de gestión de inventarios de medicamentos para un sistema de atención prehospitalaria* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Castañeda, C., Villegas, J. y Maya, P. (2015). *Diseño de un modelo de optimización/simulación para la localización de vehículos en sistemas de emergencia* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Ceballos, T., Velásquez, P. y Jaén-Posada, J. (2015). *Diseño de políticas para la reducción de la estancia hospitalaria en pacientes de la IPS Universitaria, Sede Clínica León XIII* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Clarkson, J., Slote Morris, Z., Jun, G. T., Eldabi, T., Jahangirian, M., Naseer, A. y Berger, S. (2009). *Modelling and simulation techniques for supporting healthcare decision making: A selection framework*. Research into Global Healthcare Tools. https://www-edc.eng.cam.ac.uk/downloads/right.pdf

Cortés, S. y Gutiérrez, E. V. (2016). *Un modelo de medición de madurez de capacidades para servicios de atención médica domiciliaria* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad del Valle, Colombia.

Cortés, S., Gutiérrez, E. V. y Jaén-Posada, J. (2017a). *Caracterización y diagnóstico de la gestión logística en la prestación de servicios de atención médica domiciliaria en el Valle de Aburrá* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad de Antioquia, Colombia.

Cortés, S., Gutiérrez, E. V., Villegas, J. G. y Palacio, J. D. (2017b). *Estudio e implementación de un modelo matemático para solucionar el problema de zonificación en el contexto de la atención médica domiciliaria* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Eafit, Colombia.

Coy, J. D., Akhavan-Tabatabaei, R. y Zhang, S. (2015). *Modelo de toma de decisiones para el tratamiento de pacientes colombianas con cáncer de seno* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Dai, T. y Tayur, S. (Eds.). (2018). *Handbook of healthcare analytics: Theoretical minimum for conducting 21st century research on healthcare operations*. John Wiley & Sons.

Denton, B. T. (Ed.). (2013). *Handbook of healthcare operations management: Methods and applications*. Springer-Verlag.

Denton, B., Schrage, L. y Keskinocak, P. (Eds.). (2020). *Health care*. Institute for Operations Research and the Management Sciences (Informs). https://www.informs.org/Explore/History-of-O.R.-Excellence/O.R.-Application-Areas/Health-Care

Díaz, L., Fuquén, L., Barrera, D., González-Neira, E. M., García, L. G. y Suárez, D. (2015). *Impacto de la variabilidad artificial de la programación de casos quirúrgicos en un hospital de Bogotá* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Duckett, S. J. (1987). *Operations research for health planning and administration*. Springer-Verlag.

Duque Uribe, V. y Sarache Castro, W. (2017). *Perspectivas de la investigación en la gestión de cadenas de abastecimiento, gestión logística hospitalaria y desempeño sostenible: una revisión de literatura* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad de Antioquia, Colombia.

Espejo, J., López, E. y Guerrero, W. (2019). *Diseño y evaluación de políticas de re-optimización para equipos médicos en el cuidado a la salud domiciliaria* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Industrial de Santander, Colombia.

Esteves, R. J. F. (2012). The quest for equity in Latin America: A comparative analysis of the health care reforms in Brazil and Colombia. *International Journal for Equity in Health, 11*(1), 1-16.

Estupiñán, A. M., Torres, M. J., Pérez, N., González-Neira, E. M., Barrera, D., Barbosa, J., … Suárez, D. (2015). *Mejora en ocupación, oportunidad y variabilidad*

en la programación de un servicio de cirugía [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Felizzola, H. A., Parra Jiménez, J. A. y Marulanda Roncancio, A. A. (2017). *Analítica visual para indicadores de servicios de urgencias en hospitales públicos de Bogotá* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Eafit, Colombia.

Fries, B. E. (1981). *Applications of operations research to health care delivery systems*. Springer-Verlag.

Giedelman, N., Sarmiento, A., Ruiz, C. y Guerrero, W. (2015). *Metaheurística para el problema de asignación de turnos de enfermería* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Gutiérrez, E. V., Barrera, D. y Osorio, A. F. (2018). Grupo de Trabajo en Salud. En *Health care engineering* (pp. 1-5). Pontificia Universidad Javeriana.

Hall, R. (Ed.). (2013). *Patient flow: Reducing delay in healthcare delivery* (2.ª ed.). Springer.

Hulshof, P. J. H., Kortbeek, N., Boucherie, R. J., Hans, E. W. y Bakker, P. J. M. (2012). Taxonomic classification of planning decisions in health care: A structured review of the state of the art in OR/MS. *Health Systems, 1*(2), 129-175.

Jiménez Lozano, G. y Jiménez Muñoz, X. A. (2017). *Linear programming applications: Continuous - cancer treatment. Diffuse - electric power* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Eafit, Colombia.

Kahraman, C. y Topcu, Y. I. (Eds.). (2018). *Operations research applications in health care management*. Springer.

Ley 100 de 1993, Sistema de Seguridad Social Integral de Colombia. Congreso de la República de Colombia. http://www.secretariasenado.gov.co/senado/basedoc/ley_0100_1993.html

López-Santana, E., Espejo-Díaz, J. y Méndez, G. (2016). *Modelo de programación entera mixta para programación y ruteo en ciudad a la salud domiciliaria considerando la promesa de servicio* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad del Valle, Colombia.

Manrique, J. y Gutiérrez, E. V. (2019). *Prototipo de soporte a decisiones logísticas de ruteo en atención médica domiciliaria: Aplicación en el Valle de Aburrá, Colombia* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Industrial de Santander, Colombia.

Marín, K., Morales, J., Buitrago, M. y González, J. (2017). *Simulación del proceso de facturación de citas y atención médica en un hospital del norte del Valle del Cauca* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Eafit, Colombia.

Martínez Sánchez, J. F., Pardo Rincón, H. A. y Zamora Aguas, J. P. (2018). *Análisis de la distribución de medicamentos oncológicos en una IPS de Bogotá* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad Nacional de Colombia, Colombia.

Mayorga, O., Torres, B. y Sánchez, D. (2017). *Modelo de multilocalización de centros de distribución para una empresa de distribución de medicamentos en Colombia* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad de Antioquia, Colombia.

Moscoso Moreno, X. A. y Cañón Muñoz, M. (2016). *Evaluación de alternativas de distribución de medicamentos por dosis unitarias en el Hospital Universitario de la Samaritana a partir de la aplicación de simulación de eventos discretos* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad del Valle, Colombia.

Mustafee, N. (Ed.). (2016a). *Operational research for emergency planning in healthcare* (vol. 1). Palgrave Macmillan.

Mustafee, N. (Ed.). (2016b). *Operational research for emergency planning in healthcare* (vol. 2). Palgrave Macmillan.

Noguera, M., Reynales, L. y Abdallah, R. N. (2017). *Diseño de un aplicativo para optimizar la programación en salas de cirugía* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad de Antioquia, Colombia.

Organisation for Economic Co-operation and Development (OECD). (2015). *OECD reviews of health systems: Colombia 2016*. http://www.oecd.org/health/oecd-reviews-of-health-systems-colombia-2015-9789264248908-en.htm

Organisation for Economic Co-operation and Development (OECD). (2020). *Health.* http://www.oecd.org/health

Operational Research Applied to Health Services (ORAHS). (2020). European Working Group on Operational Research Applied to Health Services. https://www.euro-online.org/web/ewg/3/ewg-orahs-euro-working-group-on-health-services

Ozcan, Y. A. (2009). *Quantitative methods in health care management: Techniques and applications* (2.ª ed.). John Wiley & Sons.

Piraban, A. y Labadie, N. (2017). *Stochastic inventory control and distribution of blood products* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Eafit, Colombia.

Poveda, M. C., Vargas, D., Urrego, J., Barrera, D., Gutiérrez, E. V., Osorio, A., … Camacho, B. (2017). *Diseño de una política de coordinación de inventarios entre el IDCBIS y hospitales de la red pública de Bogotá* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Eafit, Colombia.

Rico Carrillo, M. Y., Rico Carrillo, K., Barón Villamizar, J. y Coronado-Hernández, J. R. (2018). *Localización de ambulancias marítimas para atender emergencias de poblaciones vulnerables en la bahía de cartagena de indias* [presentación

de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad Nacional de Colombia, Colombia.

Rodríguez, A., Osorio, G. y Maya, P. (2015). *Modelo de simulación para analizar el problema de relocalización de las ambulancias de un Servicio de Emergencia Médico (SEM)* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Rojas, V., Romero, L., Paredes, G., Barrera, D. y Suárez, D. R. (2016). *Diseño de una herramienta de selección de Hospital destino para el traslado de pacientes que requieren servicio de ambulancia en la ciudad de Bogotá D. C.* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad del Valle, Colombia.

Ruiz Orjuela, E. T. y Adarme Jaimes, W. (2019). *Cadena de suministro hospitalaria desde un enfoque* lean [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad Industrial de Santander, Colombia.

Saavedra, C., Castaño, F. y Velasco, N. (2015). *Una herramienta para la planeación y programación de cirugías en ambientes con limitada disponibilidad de especialistas* [presentación de ponencia]. Congreso Colombiano de Investigación de Operaciones, Universidad de La Sabana, Colombia.

Tanfani, E. y Testi, A. (Eds.). (2012). *Advanced decision making methods applied to health care*. Springer.

Turner, J., Mehrotra, S. y Daskin, M. S. (2010). Perspectives on health-care resource management problems. En M. S. Sodhi y C. S. Tang (Eds.), *A long view of research and practice in operations research and management science: The past and the future* (vol. 148, pp. 231-247). Springer.

Vásquez Echavarría, V. y Gutiérrez, E. V. (2018). *Diseño de una herramienta computacional para la programación de personal asistencial de atención médica hospitalaria* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad Nacional de Colombia, Colombia.

Velasco, N., Barrera, D. y Amaya, C. A. (2012). Logística hospitalaria: lecciones y retos para Colombia. En *La salud en Colombia: logros, retos y recomendaciones* (pp. 309-344). Ediciones Uniandes.

Velásquez, G., Tajc, V., Zambrano, J. D., Barrera, D. y Otero, R. (2017). *Programación de turnos de enfermeras en una clínica de Bogotá* [presentación de ponencia]. Congreso Internacional Industria y Organizaciones, Universidad de Antioquia, Colombia.

World Health Organization. (2011). *Priorities in operational research to improve tuberculosis care and control*. World Health Organization. https://apps.who.int/iris/handle/10665/44662

Yang, H. y Lee, E. K. (Eds.). (2016). *Healthcare analytics: From data to knowledge to healthcare improvement*. Wiley.

Zaric, G. S. (2013). *Operations research and health care policy*. Springer-Verlag.

Capítulo 2.
Localización de un centro médico especializado en la ciudad de Bogotá integrando el proceso analítico jerárquico y la programación matemática

Juan Esteban Paredes Bayona,[1]
Carlos Andrés Bolívar Polanía[2]
y William J. Guerrero[3]

[1] Facultad de Ingeniería
Universidad de La Sabana
Chía, Colombia
juanpaba@unisabana.edu.co

[2] Facultad de Ingeniería
Universidad de La Sabana
Chía, Colombia
carlosbopo@unisabana.edu.co

[3] Facultad de Ingeniería
Universidad de La Sabana
Chía, Colombia
william.guerrero1@unisabana.edu.co

Resumen

Este capítulo presenta una propuesta metodológica que busca determinar la ubicación de un nuevo centro médico en la ciudad de Bogotá, para el mercado de medicina prepagada, mediante el uso de herramientas cuantitativas y cualitativas. La metodología se fundamenta en tres fases que se integran de forma secuencial. La primera es una fase de diseño donde se aplican conceptos de administración y estadísticos, utilizando la herramienta Canvas, que permite describir cómo una organización crea, entrega y captura el valor. La segunda fase es la priorización de factores y definición de modelo de negocio por medio de un proceso analítico jerárquico. Y la tercera fase es el diseño del modelo de localización utilizando programación matemática, específicamente programación entera, y analizando la información. El estudio está inspirado en el problema de cobertura de conjuntos que integra herramientas de administración e ingeniería que permitirán diseñar un modelo para localizar un centro médico. Por último, la metodología se aplica al contexto de la ciudad de Bogotá y con los resultados se analiza la decisión en el contexto del mercado de medicina prepagada y se estima el impacto en términos de cobertura de los servicios médicos de la ciudad.

Palabras clave: localización de hospitales, programación matemática, Canvas, logística hospitalaria, sistemas de salud

Introducción

En Colombia, el sector de la salud es uno de los menos competitivos y protagonista de numerosos cambios tributarios y de crisis financieras (Santos Ospina et al., 2016). Además, tiene una dinámica única en el país, por su complejidad y estructura (Guerrero et al., 2011). Dentro del sector de la salud existen dos ramas: la privada y la pública. La primera se financia a través de cuotas y primas moderadoras que pagan los usuarios a aseguradoras privadas (que llamaremos *entidades de medicina prepagada*) que, a su vez, subcontratan servicios con instituciones prestadoras de servicios de salud (IPS). La segunda rama se financia mediante contribuciones obligatorias, impuestos, fondos públicos y cuotas de carácter obligatorio para los ciudadanos colombianos. Estos recursos los administra el Fondo de Solidaridad y Garantía que, al mismo tiempo, es el encargado de distribuirlos a empresas aseguradoras (que llamaremos *entidades promotoras de salud* [EPS]). Las EPS se encargan de administrar estos recursos para garantizar la cobertura, la accesibilidad y la calidad del servicio.

Las empresas de medicina prepagada son empresas aseguradoras que comercializan planes complementarios de medicina especializada, a través de IPS. Cuentan con estrategias orientadas hacia la fidelización de sus usuarios y atracción de nuevos clientes, los cuales se ven beneficiados en el momento de acceder a servicios adicionales de salud. Entre tanto, las EPS son empresas aseguradoras de salud que ofrecen planes básicos, a los cuales todo colombiano tiene derecho; las empresas de medicina prepagada son planes más amplios de salud de naturaleza optativa para los colombianos.

El mercado de la medicina prepagada en Colombia es reducido; hasta septiembre del 2016 había alcanzado tan solo 1 098 996 de afiliados (Departamento Administrativo Nacional de Estadística, 2016). Se diferencia de una EPS por su complementariedad de servicios, por requerimientos de agilidad e inmediatez a la hora de consultar el servicio que impone la competencia del mercado y por su mayor costo. Cuando se acude a un servicio de medicina prepagada, hay centros médicos específicos y especializados, con un portafolio de servicios amplios, que puede ofrecer agilidad en la atención.

Las empresas de medicina prepagada buscan mejorar la situación de los usuarios, ofreciéndoles opciones de acceso y calidad de servicio. Estas oportunidades provocan que estas organizaciones estén en procesos de mejora continua. Así, buscan la expansión de centros médicos en nuevas zonas geográficas para garantizar el acceso a la prestación de sus servicios y cubrir una mayor demanda que les genere mayores ingresos. Poseer centros médicos propios tiene ventajas

operativas que son resaltadas por las empresas de medicina prepagada. Por ejemplo, los clientes perciben mejor el servicio y hay menos retrasos al transmitir la información entre entidades.

Cuando se planea la construcción de un nuevo centro médico, se debe considerar que se está discutiendo una alta inversión. Por ello, escoger el lugar idóneo, que satisfaga todas las necesidades del usuario, es una tarea crítica (Daskin y Dean, 2004). Dentro de la toma de decisiones de las empresas de medicina prepagada, es perentorio garantizar la accesibilidad, la sostenibilidad, el cubrimiento de usuarios y el mejor posicionamiento para atraer nuevos clientes y ganar terreno dentro del mercado. Son varios los elementos que deben sopesarse en esta decisión. Por ejemplo, el estado del sector y su crecimiento proyectado, el modelo de negocio del centro médico y su complejidad. Es decir, no es lo mismo ubicar uno especializado, a uno que tenga atención prioritaria, como el servicio de urgencias.

Teniendo en cuenta estos elementos y la región geográfica delimitada por la ciudad de Bogotá, en este capítulo se quiere dar una aproximación al problema de la localización de instalaciones médicas, a fin de encontrar la mejor y de optimizar los recursos de la organización, al tiempo que se está de acuerdo con las necesidades del mercado. Así, se espera que, dentro de la perspectiva de ingeniería, se logren estrategias de fidelización y atracción de nuevos usuarios exitosas, que se reflejen en las rentabilidades de las empresas, al cumplir con los objetivos de planeación estratégica establecidos. Con este objetivo, se propone un modelo basado en programación entera, que maximiza la cantidad de pacientes que puede atender un centro médico, teniendo en cuenta la distancia que están dispuestos a viajar los pacientes para acceder a los servicios y un presupuesto limitado para hacer la inversión de construcción. Con esta aproximación, la empresa de medicina prepagada podrá evaluar la viabilidad financiera del proyecto en una ciudad donde actualmente no tiene operaciones.

En la literatura científica se ha hecho evidente que es un reto científico e industrial el problema de diseñar y localizar centros médicos. En especial, Hwang y Christensen (2008) describen las dificultades, de cara a ofrecer servicios de salud económicos y accesibles para la población americana, fundamentados en el problema de diseñar modelos basados en innovaciones disruptivas para el cuidado de la salud, identificando la barrera de la falta de apropiación de tecnologías y nuevos modelos de negocio que generen valor. A efectos del análisis de modelos de negocio, asegurando su coherencia, Canvas permite sintetizar y analizar su estructura. Esta herramienta fue propuesta por Osterwalder y Pigneur (2010) y es ampliamente utilizada en la industria.

Además, con frecuencia, en las investigaciones del sector de la salud se requieren técnicas de análisis de decisión de criterios múltiples, ya que en el proceso de toma de decisiones permiten incorporar las perspectivas de factores humanos, médicos, económicos, sociales y ambientales (Adunlin et al., 2015). Por ejem-

plo, un estudio realizado en la ciudad de Bogotá para la trasferencia de pacientes (Guerrero et al., 2012) utilizó modelos multicriterio de optimización con el propósito de encontrar la mejor forma de tratar al paciente cuando precisa ser trasladado de un hospital a otro.

A efectos de resolver el problema de localización de hospitales, la literatura científica ha tenido avances significativos. Por ejemplo, un estudio realizado en India muestra la utilidad del proceso analítico jerárquico (AHP, por sus siglas en inglés) en el momento de localizar una instalación médica (Chatterjee y Mukherjee, 2013). Según el estudio, el acceso físico es un factor importante y, por lo tanto, seleccionar la ubicación se convierte en un aspecto estratégico, ya que está relacionado con la calidad del servicio. El éxito de una instalación de este tipo depende de factores ambientales, económicos, geográficos (la distancia a los clientes) y de conveniencias sociales. Los autores utilizan análisis de sensibilidad y el método AHP cuando seleccionan la ubicación de los hospitales para garantizar una ventaja competitiva. En Irán se usan técnicas de decisión de criterios múltiples con el fin de establecer centros de atención de salud y esquemas de priorización de lugares deseables para localizar hospitales en el sur de Teherán (Salehi et al., 2016). En el estudio se evalúa la proximidad a centros de salud, a espacios verdes, a otros centros médicos, a complejos residenciales y a centros educativos. Finalmente, el acceso a la comunicación y la ubicación de centros clínicos en lugares con una alta densidad de población hace que las áreas terapéuticas estén cerca de los puntos de gravedad y puedan satisfacer las necesidades clínicas de la población.

En la literatura científica también existen otras técnicas como los *modelos de optimización*, que ayudan a mejorar los procesos de decisiones frente a una oportunidad de localización de hospitales (Daskin y Dean, 2004). Estos se utilizan como referencia en investigaciones para la localización de centros médicos de diferentes servicios (Ahmadi-Javid et al., 2017). Por ejemplo, se usó un modelo de optimización con el propósito de definir la localización de un centro médico especializado en el condado de Hampshire, en Inglaterra (Meskarian et al., 2017). Con el modelo matemático se maximizó la demanda cubierta, considerando el tiempo de transporte y los servicios demandados por los pacientes. También Griffin et al. (2008) propusieron un modelo de optimización basado en el problema de cobertura de conjuntos, a fin de identificar los mejores lugares para localizar centros de salud comunitarios en el estado de Georgia (Estados Unidos). Estos ofrecen a la comunidad medicina preventiva, servicios dentales, medicina primaria, entre otros. El modelo desarrollado maximiza la cobertura promedio de la demanda de la población, de acuerdo con el presupuesto y la capacidad de las instalaciones posibles. Para hacer esto posible se tuvieron en cuenta tres factores: el número de personas sin seguro médico, la prevalencia de las condiciones de salud y la importancia de las condiciones generales de salud.

En Taiwán, Lin et al. (2008), respecto a la localización de hospitales, plantearon usar el diamante de Porter y el AHP para explicar cómo se logran ventajas competitivas con una buena localización de un hospital. Los autores determinaron que las categorías relevantes incluyen la inversión, las condiciones de demanda (factores sociodemográficos), la estructura de la estrategia, la rivalidad entre hospitales, la estructura de las industrias que apoyan (proveedores, escuelas de medicina, farmacéuticas, etc.), el gobierno y las oportunidades a partir de cambios significativos en las fluctuaciones en los costos, cambios financieros y cambio en las tasas de cambio.

Un resumen de la literatura sobre metodologías para localizar centros médicos, de acuerdo con patrones jerárquicos donde sí se incluyen disponibilidad, ubicación, propósito, niveles de relacionamiento, tiempo, capacidad, infraestructura en un análisis real se presenta en Zanjirani Farahani et al. (2014).

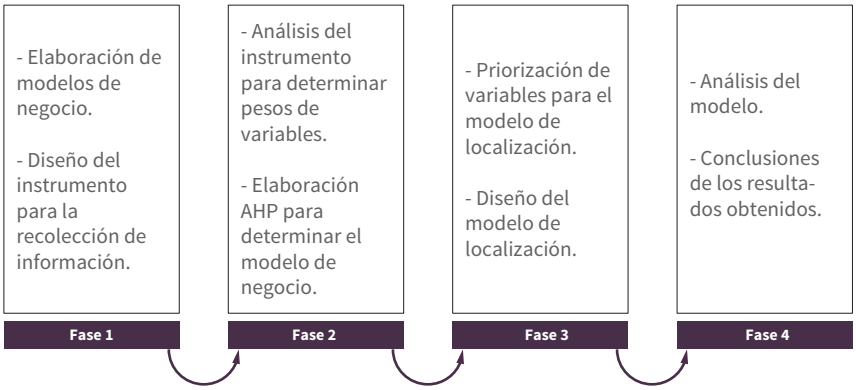
Nuestra propuesta se basa en integrar las metodologías del AHP y la programación matemática para analizar la decisión de localizar un nuevo centro médico de medicina prepagada en la ciudad de Bogotá, incluyendo la etapa de diseño del modelo de negocio, usando herramientas cualitativas y un análisis espacial de su demanda potencial. En el caso de estudio que analizamos en este manuscrito, una empresa de medicina prepagada quiere iniciar sus operaciones en la ciudad de Bogotá, a través de la localización de su primer centro médico en la ciudad que cubra la mayor cantidad de clientes potenciales, dentro de su mercado objetivo.

Por lo anterior, resumimos las contribuciones de este capítulo en dos: 1) se propone una metodología de diseño del modelo de negocio y la localización de un nuevo centro médico, específicamente para una empresa de medicina prepagada, integrando diferentes herramientas de ingeniería de tipo cuantitativo y cualitativo, y 2) se analiza el caso específico del diseño y la localización de centros médicos en la ciudad de Bogotá, mostrando sus resultados. Este documento se estructura de la siguiente manera. En la siguiente sección se presentan los métodos para analizar la decisión. Luego se muestran los resultados de la aplicación de la metodología en el caso específico de la ciudad de Bogotá, y después se analizan y discuten los hallazgos logrados en la investigación. Finalmente, se dan las conclusiones y se plantea una futura investigación.

Métodos

La propuesta metodológica que se analiza consta de cuatro fases para tomar la decisión de localizar nuevas sedes en el caso de la medicina prepagada. En este capítulo se exponen las generalidades de cada fase y en la siguiente sección se plantean los resultados de su implementación. Las fases de la metodología se presentan en la figura 1.

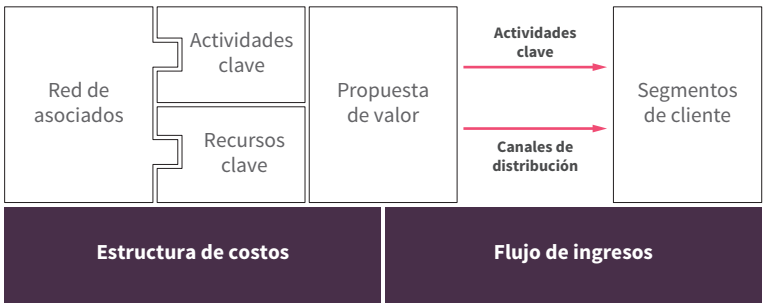
Figura 1. Fases de la metodología propuesta



Fuente: elaboración propia.

El primer paso de la propuesta metodológica es definir el modelo de negocio del centro médico que se proyecta. Este permite identificar los elementos relevantes para ubicar la sede de un centro médico. Así, con la herramienta Canvas se sintetiza y analiza la coherencia de las actividades de un modelo de negocio —véase la figura 2 propuesta por Osterwalder y Pigneur (2010)—.

Figura 2. Modelo de negocio según Canvas



Fuente: Osterwalder y Pigneur (2010).

El modelo Canvas está compuesto por nueve segmentos de bloques que se describen a continuación:

1. Segmento de clientes. Define los diferentes grupos de clientes que la organización quiere alcanzar, teniendo en cuenta los factores sociales y económicos del segmento.

2. Segmento de propuesta de valor. Describe los productos y los servicios creados por la organización con el propósito de generarle algún valor al cliente objetivo.
3. Segmento de relación con el cliente. Describe qué tipo de relación establece la organización con el segmento de cliente.
4. Segmento canales. Describe cómo la compañía se comunica con el segmento de cliente objetivo, para hacerle llegar la propuesta de valor.
5. Segmento de ingresos. Representa los flujos de efectivo y los ingresos de la organización.
6. Segmento de actividades clave. Describe cuáles son los procesos más importantes que la organización debe establecer para hacer funcionar el negocio.
7. Segmento de recursos clave. Describe los bienes más importantes, requeridos para producir los productos y servicios, así como para hacer funcionar la organización.
8. Segmento de red de proveedores y socios que contribuyen al funcionamiento del negocio. Describe los agentes que entregan la materia prima y apoyan procesos tercerizados en la organización.
9. Segmento de estructura de costos. Describe todos los costos incurridos para hacer operar el negocio (Osterwalder y Pigneur, 2010).

Estos nueve bloques deben recopilar la información necesaria para la operación del negocio deseado. En el caso del modelo de negocio del nuevo centro médico, inspirado en Canvas, se diseñan tres alternativas que, según la opinión de expertos, son viables. Luego, se diseña un instrumento de medición, para saber la importancia y los pesos de cada factor definido. Esto permitirá establecer cuál de los tres modelos es el mejor en un contexto específico. Se utiliza, así, la metodología del AHP (Bhushan y Rai, 2004), que determina la importancia de cada factor, dentro de cada modelo diseñado. Luego, bajo la misma metodología AHP (Bhushan y Rai, 2004), se puntualiza el que mejor se ajusta a las necesidades del mercado.

El siguiente paso es modelar la decisión de localización a partir del problema de cobertura de conjuntos, utilizando los factores identificados en el paso anterior, a fin de buscar la mejor cobertura de demanda en la ciudad de Bogotá de acuerdo con el modelo de negocio seleccionado. Para analizar la decisión, se consideran algunos escenarios de demanda futura; esto permite indagar sobre la sensibilidad del modelo matemático en diferentes situaciones.

Paso seguido, se analiza la información obtenida por el proceso metodológico propuesto, recogiendo la información en una matriz DOFA, para identificar debilidades, oportunidades, fortalezas y amenazas. Esta herramienta es fundamental, ya que permite tener una visión holística, al centrarse en los aspectos internos y externos de las organizaciones (David, 2011) y al destacar aspectos

internos de un proceso para la toma de una decisión. También nos permite ver aspectos externos, como la dinámica de un sector, sus mercados y los aspectos que pueden representar una amenaza u oportunidad. Por estas razones, con la matriz DOFA es posible analizar todas las herramientas cuantitativas y cualitativas propuestas desde varias posturas que afectan a las compañías. A continuación, se dan detalles específicos de cada una de las fases de la metodología propuesta.

Primera fase: diseño del modelo de negocio

Utilizando la herramienta Canvas (Osterwalder y Pigneur, 2010), se diseñaron tres alternativas de modelos de negocio diferentes, usando la opinión de expertos y considerando los nueve bloques propuestos en la herramienta. También se realizaron socializaciones preliminares con personas expertas en el sector de la salud, para obtener comentarios y observaciones que mejoraran el diseño de los modelos de negocio. El proceso del diseño del modelo de negocio fue el siguiente:

Primero se realizaron las preguntas guía que se proponen en el anexo. Para el segmento de clientes se definió el perfil estándar del cliente que se muestra en la tabla 1.

Tabla 1. Perfil del cliente de medicina prepagada

Característica	Valor
Edad	30-70 años
Sexo	Masculino/Femenino
Estado civil	Cualquiera
Nivel socioeconómico	Estrato 4 a 6
Personas a cargo	Sí/No
Ingresos mensuales aproximados	Más de COP 3 000 000 4 SMMLV
Régimen	Contributivo

Fuente: elaboración propia.

Se asumió que los ingresos mensuales del cliente potencial deben ser superiores a cuatro salarios mínimos legales vigentes para que pueda pagar una cuota mensual de entre COP 249 179 y COP 651 395, aproximadamente, según el plan seleccionado y según un promedio de los precios en el mercado. Este valor varía según el número de personas en el contrato de medicina prepagada y la edad de los clientes. De esta forma, el gasto en planes complementarios de salud equivaldrá a entre el 8 % y el 20 % del ingreso mensual del cliente.

En cuanto a la propuesta de valor, se realizó un análisis de enfermedades (morbilidad) y causas de muerte (mortalidad) en la ciudad de Bogotá (Secretaría Distrital de Salud de Bogotá, 2017), a fin de identificar las mayores incidencias de estos indicadores y, así, diseñar una mejor propuesta de valor de acuerdo con el mercado. Luego, se investigaron cuáles eran las empresas de medicina prepagada con mayor número de afiliados en la ciudad, para comparar los servicios que ofrecen los centros médicos de la competencia y generar una propuesta de valor atractiva. Las principales especialidades que se identificaron como posibles candidatas para la propuesta de valor, basados en las necesidades de la población y la oferta del mercado, se muestran en la tabla 2.

Tabla 2. Posibles servicios candidatos para las propuestas de valor

Grupo	Servicio
Consulta externa	Neumología
	Cardiología
	Endocrinología
	Odontología general
	Medicina familiar
	Gastroenterología
	Neurología
	Otorrinolaringología
	Dermatología
	Pediatría
	Oftalmología
	Psicología
	Urología
	Fisioterapia
	Medicina interna

Fuente: elaboración propia.

Con varias alternativas de modelo de negocio, se procedió a diseñar el instrumento para recolectar la información, con el objetivo de que uno fuera evaluado por al menos seis expertos del sector de la salud. Los perfiles de las personas

expertas que evaluaron los modelos de negocio y aplicaron el proceso AHP se describen a continuación:

1. Experto 1: *account manager* de industria y comercio. Su principal foco de trabajo son las cuentas del sector de la salud donde se encuentran las principales empresas de medicina prepagada del país como clientes. Con su experiencia, conoce el mercado de la medicina prepagada y facilita soluciones empresariales para cubrir nichos de mercado.
2. Experto 2: profesional del sector de la salud. Experiencia en gestión de recursos de clínicas y hospitales; conocimiento de necesidades de pacientes y contacto directo con las necesidades de estos. Su experiencia en clínicas y hospitales ofrece la posibilidad de observar cómo se mueven los recursos y modelos de negocio desde el corazón de la operación de un centro médico.
3. Experto 3: gerente de producto en el sector de la salud. Su foco de trabajo es en el manejo de necesidades de clínicas, hospitales y centros médicos. Valoración del mercado de sector salud en el país y contacto con agentes directivos de entidades públicas y privadas. Manejo de recursos y *pipeline* con orientación a proyección del mercado.
4. Experto 4: gerente de centros médicos. Gerente actual de un centro médico (sede Palermo, Bogotá) en una empresa de medicina prepagada. Su experiencia se basa en el mercado de la medicina prepagada y en el mercado de la rama contributiva en salud. Ha sido gerente de centros médicos en varias IPS del país.
5. Experto 5: enfermero. Actualmente es director de calidad en salud en una empresa de medicina prepagada, su experiencia profesional se basa en la atención a pacientes en sus primeros años, luego se especializa en el área de calidad en la salud y habilitación de centros médicos y clínicas. Conoce las normas y legislaciones que regulan el sector de la salud en temas de salud, especialmente de la medicina prepagada.
6. Experto 6: gerente de una IPS. Actualmente es consultor del sector de la salud en diferentes IPS y centros médicos. Su experiencia se basa en plantear estrategias e impulsar a las compañías a utilizar metodologías a nivel gerencial que les permita sobrevivir en el mercado. Conoce el mercado de medicina prepagada, ya que ha realizado consultorías a diferentes empresas del sector.

Segunda fase: análisis de alternativas de modelos de negocio

La segunda fase de la metodología propuesta incluye una herramienta utilizada para garantizar un buen proceso de toma de decisiones conocida como el AHP, el cual busca analizar varias alternativas y validar una decisión sin usar prejuicios y, así, proporcionarle una mayor credibilidad a la decisión final con múltiples criterios (Saaty, 2008). Tomar la decisión de manera organizada implica descomponerla en los siguientes pasos (Saaty, 2008):

- *Paso 1*: definir los criterios que se van a utilizar para tomar la decisión. En este caso son los bloques del modelo Canvas.
- *Paso 2*: definir las alternativas que se van a evaluar. En este caso, son tres modelos de negocio diferentes que se diseñan usando información pública y opinión de expertos en la fase 1 de la metodología.
- *Paso 3*: calificar cada una de las alternativas según los criterios establecidos, usando la opinión de los expertos, quienes asignan puntajes. Para las comparaciones se utilizó la escala de Saaty descrita en la literatura.
- *Paso 4*: consolidar resultados y normalizar la matriz. Luego, se calcula el vector promedio de cada uno de los bloques del Canvas. Esto se muestra en la tabla 3.

En la tabla 3 se presenta la ponderación de cada criterio y se concluye que la mayor importancia la tienen el factor "Cliente" (20 %) y la "Propuesta de valor" (14 %). El criterio que menor importancia tiene es la "Red de socios" (4 %).

Tabla 3. Ponderación normalizada para componentes principales de Canvas

Criterio	Ponderación en porcentaje
Cliente	20
Propuesta de valor	14
Relaciones	10
Canales	7
Actividades clave	10
Recursos clave	10
Red de socios	4
Costos	10
Ingresos	14

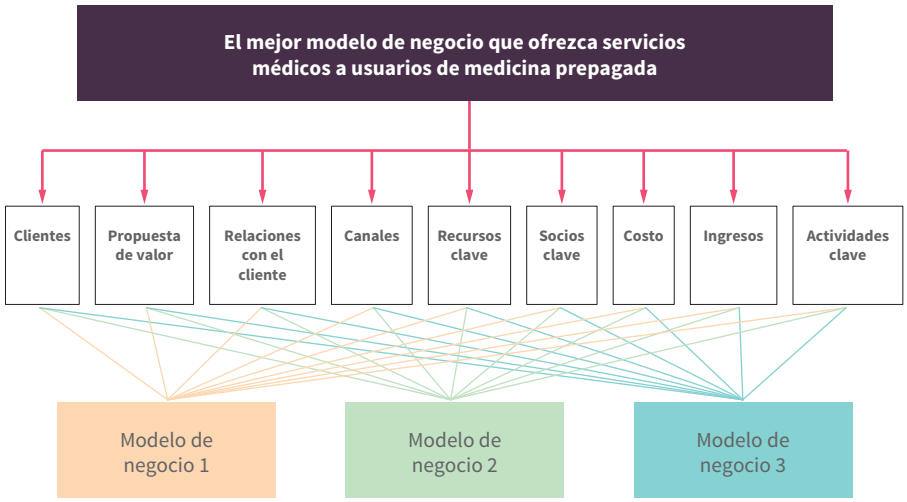
Fuente: elaboración propia.

- *Paso 5*: evaluar los tres modelos de negocio usando la opinión de cada experto, de acuerdo con los criterios de decisión principales. Estos ex-

pertos son los mismos que hacen parte del proceso de la fase anterior. En esta ocasión, cada uno ofrece su evaluación de forma independiente en una única ronda.

- *Paso 6*: consolidar resultados de cada criterio, de acuerdo con la evaluación de cada modelo y de cada experto. Luego, ponderar los resultados del vector promedio principal del paso 4 con los resultados de cada componente y así obtener el puntaje final para cada modelo, identificando la mejor opción para el objetivo planeado en el paso 1. La figura 3 presenta la estructura del modelo AHP planteado.

Figura 3. Estructura del AHP para evaluar las alternativas de modelos de negocio



Fuente: elaboración propia.

Tercera fase: diseño del modelo de localización

Se planteó un modelo de programación matemática inspirado en el problema clásico de máxima cobertura de conjuntos (Daskin y Dean, 2004) con una restricción adicional de presupuesto máximo. Lo anterior según el mejor modelo de negocio identificado en la fase 1 de investigación, y de acuerdo con sus elementos y componentes. Considera tres parámetros importantes: la demanda potencial ("Clientes") en la ciudad de Bogotá, los costos de los predios en la ciudad y las distancias de cada posible sitio. El resultado de la fase 2 de la metodología permite calcular estos parámetros, en esta fase. El modelo matemático usado es el siguiente:

Conjuntos:
I = conjunto de nodos de demanda potencial
J = conjunto de nodos candidatos a ubicar el nuevo centro médico

Parámetros:
h_i = demanda del nodo *i* [personas]
a_{ij} = distancia entre el nodo de demanda *i* y el nodo candidato de localización *j* [km]
c_j = costo de construcción en el nodo *j* [COP]
P = número de instalaciones que se van a localizar [centros médicos]
B = presupuesto disponible para comprar un lote en la localidad [COP]
Dmax = distancia máxima recorrida por un usuario [km]

Variables de decisión:
x_j = 1 *si el centro médico es localizado en el nodo j*, 0 *si no*
z_i = 1 *si la demanda de la localidad i es cubierta*, 0 *si no*

De acuerdo con estos parámetros y variables se formuló el modelo matemático de la siguiente forma:

$$Max \sum_{i \in I} h_j z_j \tag{1}$$

Sujeto a:

$$z_j \leq \sum_{j \in J \mid a_{ij} \leq Dmax} x_j, \qquad \forall i \in I, \forall j \in J \tag{2}$$

$$\sum_{j \in J} C_j * x_j \leq B \tag{3}$$

$$\sum_{j \in J} x_j = P \tag{4}$$

$$x_j \in \{0,1\} \qquad \forall j \in J \tag{5}$$

$$z_i \in \{0,1\} \qquad \forall j \in J \qquad (6)$$

El objetivo del modelo (1) es maximizar la cobertura de la demanda en la ciudad de Bogotá. La ecuación (2) indica que la demanda en el nodo *i* es cubierta si y solo si existe por, al menos, una instalación dentro del rango de cobertura, considerando la distancia máxima. La ecuación (3) considera los costos del predio *j* respecto a la localización *j,* según el tope máximo de presupuesto disponible. Luego, la ecuación (4) indica que se deben localizar exactamente *P* instalaciones en la ciudad. Las ecuaciones (5) y (6) son restricciones sobre la naturaleza binaria de las variables de decisión.

El modelo propuesto no considera una restricción de la capacidad máxima; sin embargo, dado que maximiza la población total que puede ser atendida por cada centro médico en el largo plazo, al conocer potencialmente el tamaño de la población beneficiada, el tomador de decisiones podrá diseñar las instalaciones de acuerdo con el requerimiento de capacidad.

Resultados

Los resultados generados con la metodología son los siguientes. Iniciamos con la elaboración de alternativas del modelo de negocio que se ajustan a las oportunidades del sector, según la opinión de los expertos entrevistados. Los rangos de edad de los clientes se estiman según la frecuencia de atención de los servicios, según la opinión de los expertos.

Modelo de negocio 1: este se centra en ofertar exclusivamente servicios médicos especializados, como neumología, cardiología, endocrinología, gastroenterología, para pacientes de 30 a 70 años de edad y de estrato socioeconómico 4, 5 y 6, con ingresos mayores a tres millones de pesos mensuales. Presta servicios a empresas de medicina prepagada, cobrando bonos y cuentas médicas a las aseguradoras; la promoción y distribución está centrada en negociaciones y licitaciones con estas empresas. Adicionalmente, se tienen puntos de contacto con el usuario por medio de un *contact center* y correo electrónico, que buscan hacer relaciones de confianza, calidez y seguridad. Las actividades clave se relacionan con la gestión humana, financiera, jurídica, tecnológica, administrativa, así como con procesos misionales como consulta externa y gestión de control y mejoramiento, planeación y dirección y servicio al cliente. Finalmente, la red de socios está compuesta por médicos, proveedores de insumos y empresas de servicios generales.

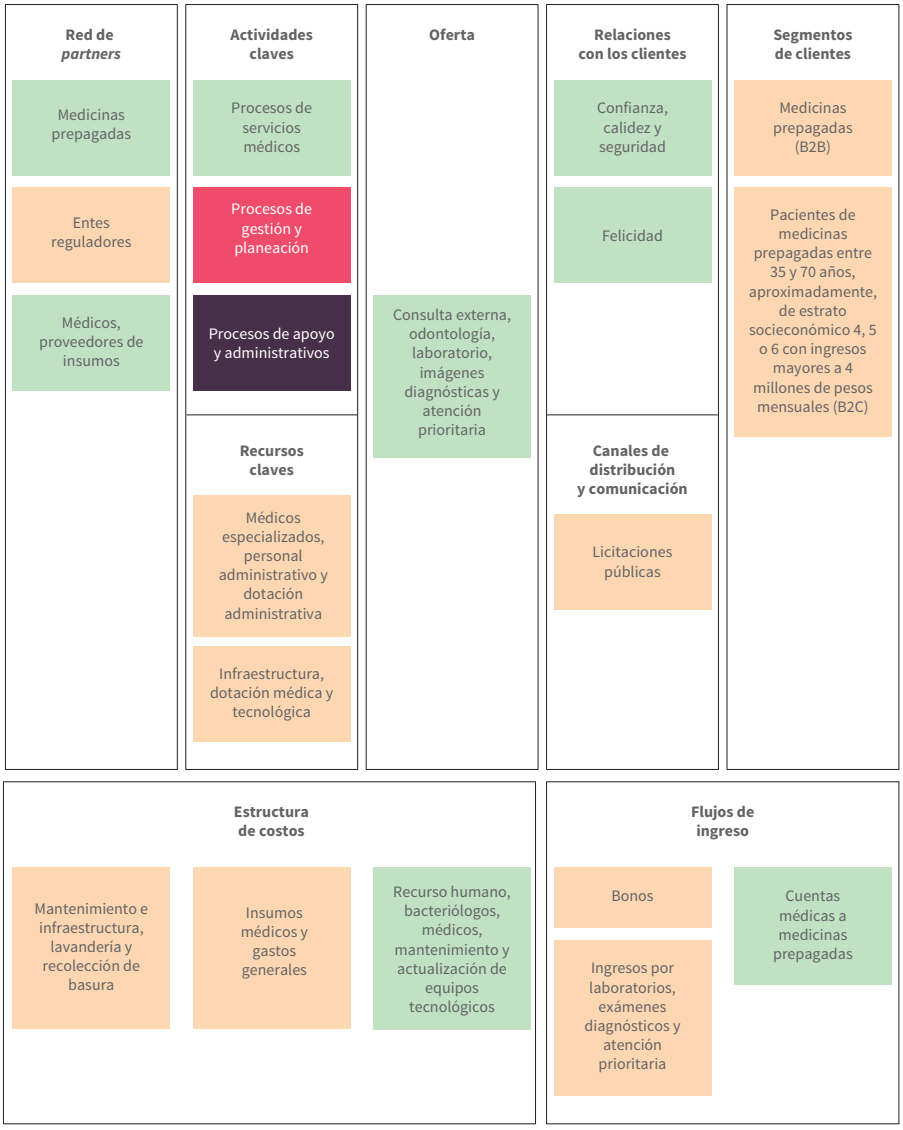
Modelo de negocio 2: este se centra en ofertar servicios médicos especializados por medio de unidades especializadas, especificados en la tabla 2, y utilizando la mejor tecnología disponible en el mercado para prestar tales servicios

(por ejemplo, exámenes de laboratorio con acceso a los resultados en línea). Solamente se centra en usuarios con planes corporativos con rango de edad entre 35 y 65 años y presta servicios a empresas de medicina prepagada, cobrando copagos directos y por servicios, laboratorios y exámenes diagnósticos. La promoción y distribución está centrada en la publicidad —anuncios en redes sociales, mercadeo en buscadores y correo electrónico— y busca mejorar a través del servicio al cliente. Las actividades clave del modelo se centran en actividades de prestación de servicio al cliente, gestión humana, gestión tecnológica, calidad, financiera, administrativa y mercadeo. Adicionalmente, la red de socios está compuesta por entes reguladores, medicinas prepagadas, médicos particulares, proveedores de tecnología y servicios generales.

Modelo de negocio 3: este se centra en ofertar exclusivamente servicios médicos especializados, especificados en la tabla 2, e integrarlos con servicios como exámenes de laboratorio, atención prioritaria y exámenes diagnósticos para pacientes entre 35 y 70 años de edad, con estrato socioeconómico de 4, 5 y 6 y con ingresos de cuatro millones de pesos mensuales. Presta servicios a usuarios de medicina prepagada, cobrando bonos por servicios de laboratorios, exámenes diagnósticos y atención prioritaria. La promoción y distribución está centrada en las licitaciones públicas, a fin de crear relaciones de confianza, calidez y seguridad. Las actividades clave se enfocan en los servicios médicos, los procesos de gestión y planeación y los administrativos y de apoyo. Al final, la red de socios está compuesta por entes reguladores, medicinas prepagadas, médicos particulares y proveedores de insumos.

En la tabla 4 se muestra un paralelo que resume las principales diferencias entre los tres modelos descritos con los aspectos más característicos de los modelos de negocio. El resultado de aplicar el AHP se puede resumir en la tabla 5. Los factores con mayor ponderación son los ingresos, costos, recursos clave, canales, relaciones y propuesta de valor. El factor de menor ponderación es la red de socios. De acuerdo con los resultados obtenidos en la aplicación del instrumento a los expertos del sector de la salud en Colombia y los cálculos realizados del método AHP, el mejor modelo de negocio que se ajusta al mercado de la medicina prepagada es el modelo 3, con un puntaje del 50,4 %. Cabe resaltar que el modelo 3 solo es superado en la calificación que los expertos hacen del factor cliente en el modelo 1. La figura 4 muestra el modelo Canvas.

Figura 4. Modelo de negocio 3



Fuente: elaboración propia.

Tabla 4. Principales diferencias entre los modelos de negocio planteados

Aspecto	Modelo 1	Modelo 2	Modelo 3
Servicios ofertados	Servicios médicos especializados, como neumología, cardiología, endocrinología y gastroenterología	Presta servicios médicos con diferenciador tecnológico	Presta servicios médicos especializados integrados con otros servicios como: laboratorio clínico, exámenes diagnósticos, entre otros
Clientes potenciales: edad	Pacientes afiliados entre 30 y 70 años de edad	Pacientes afiliados solamente corporativos entre 35 y 65 años de edad	Pacientes afiliados entre 35 y 70 años de edad
Ingresos de los clientes	Pacientes de estratos socioeconómicos 4, 5 y 6 con ingresos de 4 SMMLV	Exclusivo para pacientes corporativos	Pacientes de estratos socioeconómicos 4, 5 y 6 con ingresos de 5 SMMLV
Medios de contacto con el cliente	Distribución y promoción por medio de licitaciones privadas y negociaciones con medicinas prepagadas	Distribución y promoción por medio de publicidad, anuncios en redes sociales	Licitaciones públicas
Red de socios estratégicos	Está compuesta por medicina prepagada, médicos, proveedores de insumos y empresas de servicios generales y vigilancia	Está compuesta por entes reguladores, medicinas prepagadas y proveedores de tecnología	Está compuesta por entes reguladores, medicinas prepagadas, médicos y proveedores de insumos

Fuente: elaboración propia.

Tabla 5. Resumen de los resultados calculados con el AHP (en porcentajes)

	Cliente	Propuesta de valor	Relaciones	Canales	Actividades clave	Recursos clave	Red de socios	Costos	Ingresos	Total
Modelo 1	45	7	39	46	16	18	42	16	23	27,2
Modelo 2	18	30	17	26	29	18	22	37	12	22,4
Modelo 3	38	63	44	28	54	64	36	47	65	50,4
Ponderación	**20**	**14**	**10**	**7**	**10**	**10**	**4**	**10**	**14**	**100**

Fuente: elaboración propia.

Una vez definido el modelo de negocio, se propone usar el modelo matemático para definir la ubicación exacta del nuevo centro médico. El modelo matemático se implementó en el *software* GAMS-CPLEX V12.1 utilizando una estación de trabajo con un procesador Intel Core i7 a @2,6GHz y una memoria RAM de 8 Gb.

A fin de calcular los parámetros del modelo se utilizaron fuentes oficiales, como se describe a continuación. En primer lugar, se seleccionaron las posibles localidades candidatas en la ciudad de Bogotá para ubicar un centro médico de acuerdo con el segmento de cliente definido en el modelo seleccionado. Estas posibles localidades son Usaquén, Suba, Chapinero, Barrios Unidos, Teusaquillo y Fontibón. Se excluyeron las localidades sin una población significativa que esté dentro del sector socioeconómico definido para el mercado objetivo de las empresas de medicina prepagada.

De acuerdo con esta información y con estudios de proyección de población realizados por la Secretaría Distrital de Planeación, las posibles localidades se muestran en la tabla 6 junto con la proyección de población (Cardona Balanta, 2014). Es posible analizar que la Secretaría de Planeación proyecta para los próximos años un crecimiento lineal. Para la localidad de Suba, se espera un crecimiento más acelerado en la población que en las otras localidades. El mapa de la ciudad de Bogotá se muestra en la figura 5, con la distribución por estratos socioeconómicos de la población.

Tabla 6. Localidades y demanda para ubicar un centro médico, 2016-2020

Año	Suba	Usaquén	Chapinero	Barrios Unidos	Teusaquillo	Fontibón
2016	1 250 734	472 908	126 952	263 883	140 767	403 519
2017	1 282 978	474 186	126 591	267 106	140 473	413 734
2018	1 315 509	475 275	126 192	270 280	140 135	424 038
2019	1 348 372	476 184	125 750	273 396	139 776	434 446
2020	1 381 597	476 931	125 294	276 453	139 369	444 951

Fuente: Cardona Balanta (2014).

Respecto a los nodos de demanda y los nodos candidato para localizar el centro médico, se utiliza la metodología de kilómetro cuadrado, propuesta por Yoshizaki et al. (2018). En esta, generamos una cuadrícula virtual con celdas de 1 km × 1 km para caracterizar cada kilómetro cuadrado. Utilizando el mapa de la Secretaría Nacional de Planeación y aplicando la metodología, se identificaron 169 nodos en total, como se muestra en la figura 6. Cada celda de la cuadrícula es

un nodo que representa, entonces, un área de 1 km^2 de la ciudad y son candidatos simultáneamente nodos de demanda y nodos candidato para localizar un hospital. Cada nodo utiliza un código de identificación. Los nodos de la localidad de Suba tienen los códigos S00; Usaquén, U00; Chapinero, C00; Teusaquillo, T00; Barrios Unidos, B00, y Fontibón, F00. Estos nodos son candidatos para la localización del centro médico. De ahí se excluyen las zonas donde es infactible localizar la instalación, es decir, se excluyen parques, lagos, etc. En la cuadrícula se asignan unas coordenadas en los ejes *x* y *y* para geolocalizar los nodos.

En cuanto al cálculo de la demanda en cada nodo, se considera que cada uno (kilómetros cuadrados) tiene varios estratos en su interior. Para este cálculo se usa información oficial de la ciudad de Bogotá (Secretaría Distrital de Salud de Bogotá, 2017). Se estima el porcentaje de demanda asignado por cada kilómetro cuadrado, de acuerdo con la densidad poblacional de la localidad correspondiente y la concentración de estratos socioeconómicos, que se definan en el modelo de negocio para la empresa de medicina prepagada. Entonces, la demanda de un nodo será igual a la multiplicación de la densidad poblacional de cada localidad por el porcentaje de concentración del estrato socioeconómico del barrio correspondiente al mercado objetivo que se definió (estratos 4, 5, y 6). Adicionalmente, se calcula el costo promedio de un metro cuadrado para cada nodo. Aquí se utiliza un estimado de área de 600 m^2 para el centro médico propuesto, por lo cual el costo promedio se multiplica por 600 m^2.

Las distancias de cada sitio se calculan según la figura 6. En este mapa se asignan coordenadas para cada nodo, con coordenadas en los ejes *x* y *y*. Luego, se calculan las distancias euclidianas entre cada nodo. Existe *software* de georreferenciación que puede refinar el cálculo de las distancias entre puntos de la ciudad considerando la malla vial. Sin embargo, al desconocer el medio de transporte que usarán los pacientes para desplazarse al centro médico, consideramos que la distancia euclidiana es una aproximación adecuada.

Adicionalmente, se estima el tiempo máximo que una persona está dispuesta a viajar (*Dmax*) desde su casa a un centro médico u hospital. Para ello se utilizó una encuesta con muestreo por conveniencia a 27 personas (Otzen y Manterola, 2017). Los resultados mostraron que el 44 % de los encuestados están dispuestos a viajar entre 20 y 25 min, y el 37 %, entre 25 y 40 min. Por lo tanto, asumimos razonable que un cliente viaje hasta 20 min para llegar al centro médico. Esto entendiendo que la empresa quiere caracterizarse por su cercanía al paciente, en contraste con la competencia. Asumiendo una velocidad de viaje promedio de 24 km/h (Secretaría Distrital de Movilidad, 2018), el parámetro *Dmax* es aproximadamente 8 km.

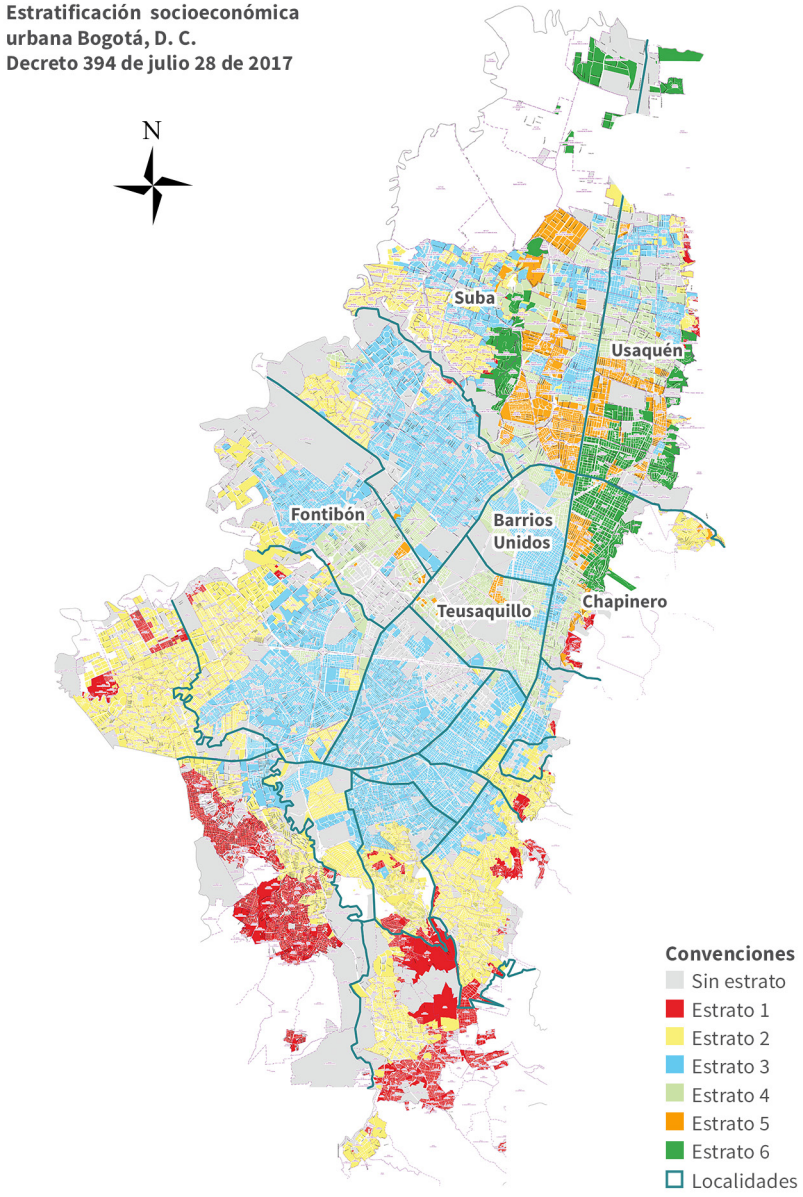
Con varios escenarios se analizó la sensibilidad del modelo ante los parámetros presupuestos. Estos análisis son interesantes ante cinco escenarios de decisión que consideraron diferentes presupuestos para la adquisición del predio

del centro médico. Estos rangos evalúan las alternativas de inversión en sectores socioeconómicos diferentes. Por ejemplo, el escenario 1 permite construir el centro médico en el estrato 1; el escenario 2, en los estratos socioeconómicos 1 o 2. Así, de forma análoga, con los escenarios 3 al 5. El presupuesto para cada escenario se muestra a continuación:

- Escenario 1: COP 800 000 000 (aprox. USD 250 000)
- Escenario 2: COP 1 378 000 000 (aprox. USD 250 000)
- Escenario 3: COP 2 400 000 000 (aprox. USD 430 000)
- Escenario 4: COP 3 100 000 000 (aprox. USD 950 000)
- Escenario 5: COP 5 200 000 000 (aprox. USD 1 610 000)

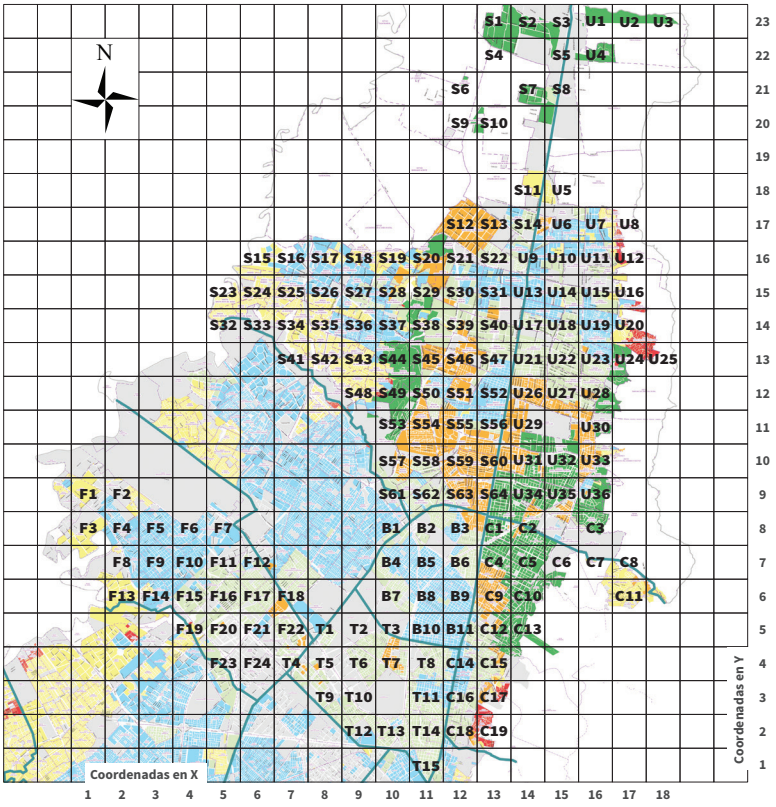
Después de esto, se realizó la optimización para cada escenario presupuestal. La ejecución del modelo nunca tarda más de un segundo en encontrar la solución óptima. Podemos ver que la decisión de localización de un centro médico se afecta fuertemente con variaciones en el presupuesto. El detalle de los resultados se muestra en la tabla 7. Por ejemplo, con un presupuesto de COP 800 000 000, el centro médico debe ubicarse en el punto U12, el cual queda en la localidad de Usaquén. Este punto cubre los kilómetros cuadrados sombreados en la figura 7, los cuales comprenden los kilómetros cuadrados de las localidades de Suba y Usaquén únicamente. La máxima cobertura es de 876 251 personas y se cubren un total de 61 puntos.

Figura 5. Mapa de Bogotá por localidades potenciales



Fuente: Secretaría de Planeación (2017).

Figura 6. Sitios posibles para ubicar un centro médico en la ciudad de Bogotá con coordenadas en *x-y*



Fuente: elaboración propia con base en el mapa de la Secretaría de Planeación (2017).

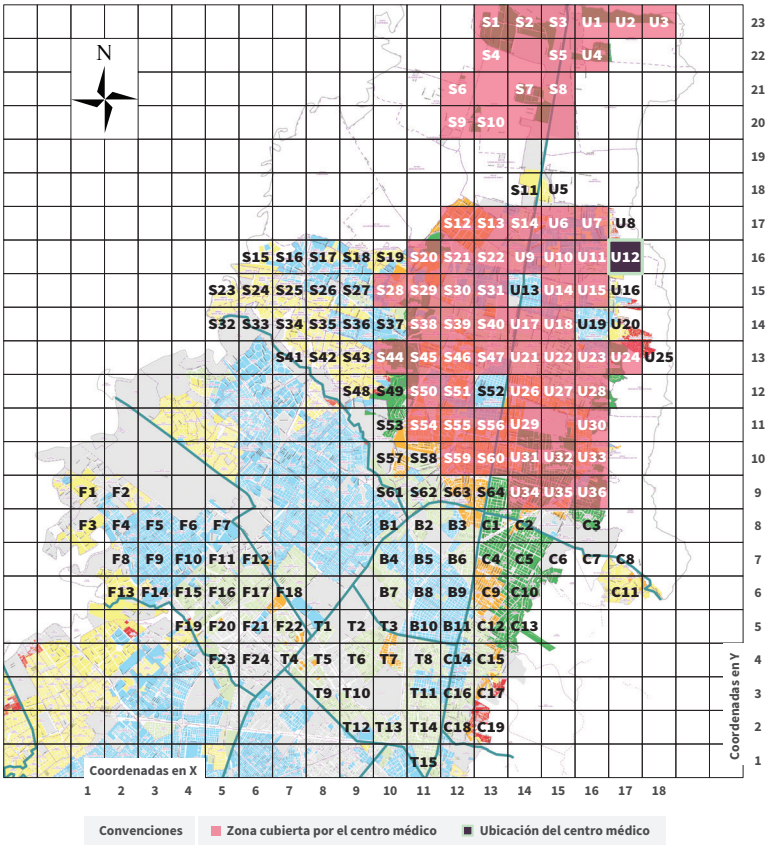
La tabla 7 sintetiza los resultados obtenidos por escenario presupuestal del modelo matemático. Vemos que la restricción presupuestal influye en la decisión de localización del centro médico y en el número de personas potencialmente beneficiarias de la nueva instalación, a medida que aumenta el presupuesto disponible. Sin embargo, por encima de COP 1 134 069, aumentar el presupuesto no tiene impacto. La figura 7 muestra la localización del hospital para los escenarios 4 y 5 y la cobertura lograda.

Tabla 7. Resumen de los resultados obtenidos del análisis de sensibilidad al parámetro de presupuesto disponible

Escenario	Kilómetros cuadrados cubiertos	Número de personas cubiertas	Localización
1	61	876 251	U12
2	82	1 043 640	S48
3	92	1 127 173	S58
4	97	1 134 069	S57
5	97	1 134 069	S57

Fuente: elaboración propia.

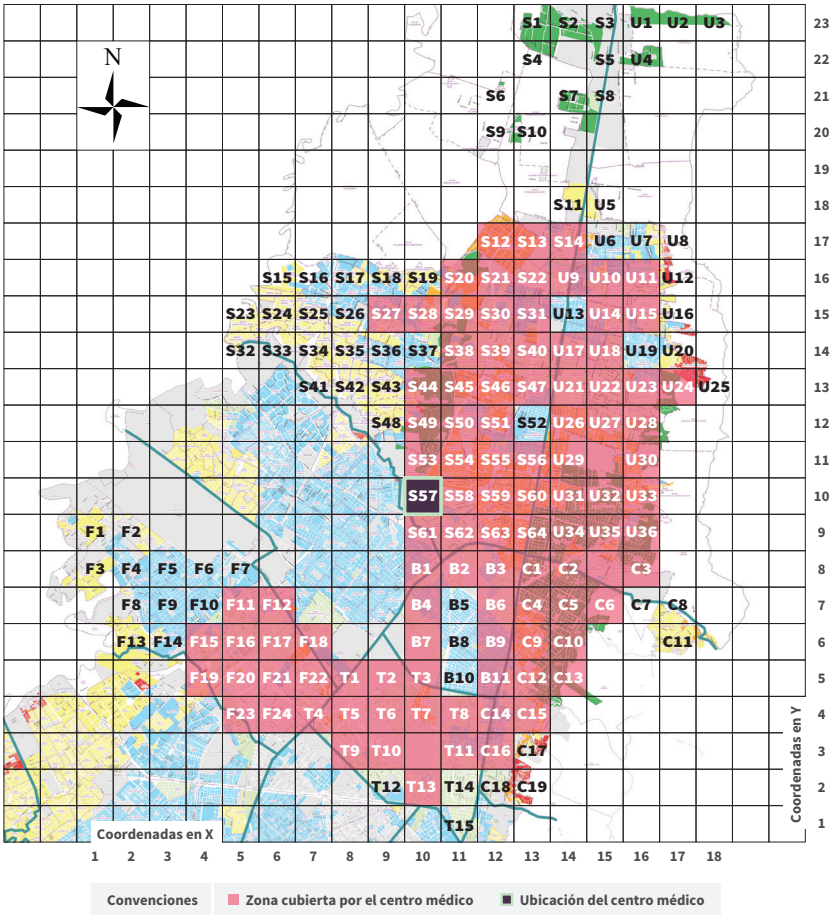
Figura 7. Modelo de localización de un centro médico en Bogotá en el escenario 1



Fuente: elaboración propia.

De acuerdo con el desarrollo de la propuesta metodológica planteada y con los resultados obtenidos, hacemos un análisis. En primer lugar, se identifica que la aproximación es innovadora para el sector de la salud y para el mercado de la medicina prepagada, ya que integra herramientas gerenciales y de ingeniería que no se han documentado en estudios previos en Colombia. Adicionalmente, al utilizar el modelo de negocio Canvas podemos comprender las diferentes ventajas competitivas de los distintos centros médicos que hay en el mercado de la medicina prepagada, y sus diferentes propuestas de valor. Al utilizar la metodología del AHP podemos comprender cuáles son los factores más importantes en el momento de diseñar un modelo de negocio de centros médicos, de acuerdo con la opinión de expertos en el sector (figura 8).

Figura 8. Modelo de localización de un centro médico en Bogotá en los escenarios 4 y 5



Fuente: elaboración propia.

Utilizando herramientas, como las entrevistas no estructuradas, es posible recolectar información cualitativa de forma más objetiva que permita medir la importancia de cada factor y así aplicar métodos como el AHP. Con el modelo matemático de cobertura de conjuntos se identifica la localización que logra la máxima cobertura de demanda en la ciudad considerando las demandas en cada zona de la ciudad, discretizada en kilómetros cuadrados y considerando la distancia máxima de viaje de la población. Adicionalmente, la metodología propuesta puede ser adaptada de manera fácil para ser desarrollada en otras ciudades y ayudar a las empresas de medicina prepagada y EPS a ubicar centros médicos especializados.

Como limitantes de la metodología propuesta, se identificó que se requiere tener en cuenta factores externos como restricciones legales, políticas o gubernamentales que regulan el sector, así como elementos ambientales o socioculturales. La metodología tampoco considera un modelo financiero que evalúe indicadores de rentabilidad como el valor presente, tasa interna de retorno, rentabilidad sobre los recursos, ni el punto de equilibrio del negocio. En la metodología propuesta no se contempla un análisis de competencia, pues esta es una metodología para el sector en general.

En futuras investigaciones, la metodología puede examinar herramientas como el PESTAL, análisis de las cinco fuerzas de Porter, modelos financieros de análisis de rentabilidad de negocio, punto de equilibrio y presupuestos, de acuerdo con las capacidades y recursos de la compañía. También se pueden desarrollar modelos de localización que consideren índices de prevalencia e incidencia de la población que se está estudiando para aproximarse mejor a las necesidades de salud de las comunidades en específico.

Conclusiones

En la presente investigación se planteó una aproximación metodológica práctica cuyo fin era localizar un centro médico en la ciudad de Bogotá para el mercado de la medicina prepagada, por medio de herramientas de ingeniería cualitativas y cuantitativas, que permiten apoyar el proceso de toma de decisiones, en cuanto a la expansión de servicios de una empresa y obtener el mejor resultado. Dentro de la investigación se diseñó un modelo de negocio de un centro médico exclusivo para el mercado de la medicina prepagada. También se aplicó un modelo de investigación de operaciones, llamado *modelo de cobertura de conjuntos*, que permite obtener el correcto posicionamiento de centros médicos utilizando herramientas de ingeniería, y se analizó la decisión en función de la capacidad presupuestal, la cantidad de población potencialmente beneficiaria de la nueva instalación y su posición geográfica.

La principal contribución de la investigación es la aproximación metodológica propuesta para la localización de un centro médico en la ciudad de Bogotá, pues en Colombia no hay un estudio documentado que presente una metodología

que permita ser desarrollada por fases y que pueda integrar herramientas cualitativas y cuantitativas para el sector de la salud y el mercado de la medicina prepagada. También podemos resaltar el modelo de negocio diseñado y priorizado con el método AHP, pues este es el ideal para un centro médico de acuerdo con el juicio de expertos en el sector. Finalmente, con el modelo matemático propuesto se obtuvo la mejor localización en la ciudad de Bogotá de acuerdo con el modelo de negocio diseñado y luego priorizado con el método AHP.

Con el modelo matemático presentado es posible considerar parámetros como: demanda potencial, costo de cada predio y distancias esperadas de viaje de los usuarios, que identifican la localización ideal de un centro médico en la ciudad de Bogotá. Los escenarios propuestos lograron mostrar una relación directa entre el presupuesto y la demanda, pues entre más presupuesto mayor será la cobertura, hasta alcanzar el tope máximo de demanda posible capaz de cubrir.

El impacto esperado en el sector es que la metodología propuesta sea de beneficio para la industria de la medicina prepagada, pues esta contiene herramientas que se pueden implementar en un contexto empresarial y que pueden utilizarse en la toma de decisiones. Los resultados son relevantes para el sector, pues están validados por expertos de este (por ejemplo, el modelo de negocio); además, el resultado del modelo matemático muestra la ubicación para cubrir la máxima demanda posible. Finalmente, la metodología se analiza en toda su integralidad, identificando las limitaciones, las fortalezas y las oportunidades de la metodología propuesta.

Las principales investigaciones futuras incluyen el desarrollo de esta metodología enriquecida con otras herramientas como PESTAL, análisis de las cinco fuerzas de Porter, modelos financieros que permitan análisis financieros, punto de equilibrios y presupuestos de acuerdo con las capacidades y recursos de la compañía. El modelo matemático puede considerar restricciones de capacidad del centro médico e índices de prevalencia e incidencia de la población y, así mismo, demandas como variables aleatorias. Un análisis de la robustez del modelo ante los cambios en el pronóstico de las demandas también se realizará en una investigación futura.

Adicionalmente, es posible que investigaciones futuras consideren otros elementos del modelo de negocio que ayuden a robustecer el modelo matemático: los canales de comunicación, las relaciones con el cliente, el flujo de ingresos, los recursos clave del centro médico, los principales proveedores de insumos y aliados clave y los costos asociados. Con la inclusión de estos elementos se analizarán las distancias entre los proveedores y los socios clave del centro médico; igualmente, costos como la tecnología, el recurso humano e insumos médicos, y costos fijos como los servicios públicos, los arriendos y los impuestos necesarios para la prestación del servicio y funcionamiento. Además, podemos analizar la localización del hospital a partir de restricciones de capacidad, lógica difusa para el parámetro de distancia máxima de viaje y la demanda. Otra posible modificación al trabajo es contemplar la apertura de varios centros médicos que diversifiquen el servicio, al integrar en el presupuesto otros aspectos como los costos fijos y los gastos de operación de acuerdo con el servicio que preste.

Referencias

Adunlin, G., Diaby, V. y Xiao, H. (2015). Application of multicriteria decision analysis in health care: A systematic review and bibliometric analysis. *Health Expectations*, *18*(6), 1894-1905. https://doi.org/10.1111/hex.12287

Ahmadi-Javid, A., Seyedi, P. y Syam, S. S. (2017). A survey of healthcare facility location. *Computers and Operations Research*, *79*, 223-263. https://doi.org/10.1016/j.cor.2016.05.018

Bhushan, N. y Rai, K. (2004). *Strategic decision making: Applying the analytic hierarchy process*. Springer. https://doi.org/10.1007/b97668

Cardona Balanta, M. I. (2014). *Proyecciones de población por localidades para Bogotá 2016-2020*. Secretaría Distrital de Planeación. http://www.sdp.gov.co/sites/default/files/boletin69.pdf

Chatterjee, D. y Mukherjee, B. (2013). Potential hospital location selection using AHP: A study in rural India. *International Journal of Computer Applications*, *71*(17), 1-7. https://doi.org/10.5120/12447-9144

Departamento Administrativo Nacional de Estadística (DANE). *Entidades de medicina prepagada, afiliados septiembre 2016* [circular única-archivo tipo 040].

Daskin, M. S. y Dean, L. K. (2004). Location of health care facilities. En M. L. Brandeau, F. Sainfort y W. P. Pierskalla (Eds.), *Operations research and health care: A handbook of methods and applications* (pp. 43-76). Springer. https://doi.org/10.1007/1-4020-8066-2_3

David, F. R. (2011). *Strategic management: Concepts and case* (13.ª ed.). Pearson. http://www.mim.ac.mw/books/David%27s%20Strategic%20Management%20Concepts%20and%20Cases%2013e%20.pdf

Griffin, P. M., Scherrer, C. R. y Swann, J. L. (2008). Optimization of community health center locations and service offerings with statistical need estimation. *IIE Transactions (Institute of Industrial Engineers)*, *40*(9), 880-892. https://doi.org/10.1080/07408170802165864

Guerrero, R., Gallego, A., Becerril-Montekio, V. y Vásquez, J. (2011). Sistema de salud en Colombia. *Salud Pública de México*, *53*(Suppl 2), s144-s155. https://www.redalyc.org/pdf/106/10619779010.pdf

Guerrero, W. J., Velasco, N. y Amaya, C. A. (2012). Multi-objective optimization for interfacility patient transfer. En *Production systems and supply chain management in emerging countries: Best practices* (pp. 81-95). Heidelberg. https://doi.org/10.1007/978-3-642-26004-9_5

Hwang, J. y Christensen, C. M. (2008). Disruptive innovation in health care delivery: A framework for business-model innovation. *Health Affairs*, *27*(5), 1329-1335. https://doi.org/10.1377/hlthaff.27.5.1329

Lin, C. T., Wu, C. R. y Chen, H. C. (2008). The study of construct key success factors for the Taiwanese hospitals of location selection by using the fuzzy AHP and sensi-

tivity analysis. *International Journal of Information and Management Sciences*, *19*(1), 175-200.

Meskarian, R., Penn, M. L., Williams, S. y Monks, T. (2017). A facility location model for analysis of current and future demand for sexual health services. *PloS One*, *12*(8), 1-18. https://doi.org/10.1371/journal.pone.0183942

Osterwalder, A. y Pigneur, Y. (2010). *Business model generation: A handbook for visionaries, game changers, and challengers*. John Wiley and Sons. https://profesores.virtual.uniandes.edu.co/~isis1404/dokuwiki/lib/exe/fetch.php?media=-bibliografia:9_business_model_generation.pdf

Otzen, T. y Manterola, C. (2017). Técnicas de muestreo sobre una población a estudio. *International Journal of Morphology*, *35*(1), 227-232. http://dx.doi.org/10.4067/S0717-95022017000100037

Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, *1*(1), 83. https://doi.org/10.1504/IJSSCI.2008.017590

Salehi, Z., Rajabi, M., Naghavi, M., Roshan, E. M. y Jamal, P. G. (2016). Using positioning city hospitals, multi-criteria analysis method. *Bulletin de la Société Royale des Sciences de Liège*, *85*, 620-630. https://popups.uliege.be/0037-9565/index.php?id=5530

Santos Ospina, A., García Molina, M. y Chicaíza Becerra, L. (2016). ¿Crisis financiera o de gestión? Evolución del sector salud desde la perspectiva de los entes territoriales. *Apuntes del CENES, 35*(61), 177-206. https://doi.org/10.19053/22565779.4151

Secretaría Distrital de Movilidad. (2018). *Datos abiertos.* https://www.movilidadbogota.gov.co/web/datos_abiertos

Secretaría Distrital de Planeación. (2017). *Estratificación socioeconómica*. Alcaldía Mayor de Bogotá. http://www.saludcapital.gov.co/Paginas2/Inicio.aspx

Secretaría Distrital de Salud de Bogotá. (2017). *Plan Territorial de Salud para Bogotá, D. C. 2016-2020*. Alcaldía Mayor de Bogotá. http://www.saludcapital.gov.co/Documents/Plan_Territorial_Salud_2016_2020.pdf

Yoshizaki, H. T. Y., Velázquez Martínez, J. C. y Mejía Argueta, C. (2018). *Supply chain management and logistics in Latin America: A multi-country perspective*. Emerald Publishing. https://www.emerald.com/insight/publication/doi/10.1108/9781787568037

Zanjirani Farahani, R., Hekmatfar, M., Fahimnia, B. y Kazemzadeh, N. (2014). Hierarchical facility location problem: Models, classifications, techniques, and applications. *Computers & Industrial Engineering*, *68*, 104-117. https://doi.org/10.1016/j.cie.2013.12.005

Anexo

Preguntas guía para el diseño de alternativas en el modelo Canvas (Osterwalder y Pigneur, 2010)

Para el segmento de clientes se utilizaron las siguientes preguntas guía: ¿para quienes creamos valor? y ¿quiénes son nuestros clientes más importantes?

Para la propuesta de valor se utilizaron las siguientes preguntas guía: ¿cuál es la propuesta de valor entregada al cliente?, ¿cuáles problemas del cliente estamos ayudando a solucionar?, ¿cuáles necesidades del cliente estamos satisfaciendo?, ¿cuál es el portafolio de productos o servicios que se está ofreciendo para cada segmento de clientes?

Para canales de comercialización se respondieron las siguientes preguntas: ¿a través de cual canal los clientes quieren ser alcanzados?, ¿cómo llegaremos a los usuarios?, ¿cómo se integran nuestros canales?, ¿cuál de ellos trabaja mejor?

Para definir cómo deben ser las relaciones con el cliente se respondió la siguiente pregunta clave: ¿qué tipo de relación con cada segmento se espera establecer y mantener? Para definir el flujo de ingresos del negocio se respondieron las siguientes preguntas: ¿por cuáles servicios nuestros clientes estarán dispuestos a pagar?, ¿por qué los pagarían?, ¿cómo los pagarían?, ¿cómo preferirían pagar?

Se diseñó el componente de recursos clave, respondiendo a la siguiente pregunta clave: ¿qué recursos clave requiere la propuesta de valor: nuestros canales, nuestras relaciones o nuestro flujo de ingresos?

Se identificó el componente de actividades clave con las siguientes preguntas: ¿quiénes son nuestros aliados clave?, ¿quiénes son nuestros proveedores clave?, ¿cuáles recursos clave serán adquiridos por medio de socios clave?, ¿cuáles actividades clave realizarán los socios clave?

Finalmente, se diseñó el componente de estructura de costos, respondiendo a las siguientes preguntas clave: ¿cuál es el costo más importante en el modelo de negocio?, ¿cuáles recursos clave son los más costosos?, ¿cuáles actividades clave son las más costosas?

Capítulo 3.
Evaluación de rutas de personal asistencial en el cuidado a la salud domiciliaria mediante programación matemática y simulación multiagente

Julián Alberto Espejo-Díaz,[1] Eduyn López-Santana[2] y William J. Guerrero[3]

[1] Facultad de Ingeniería
Universidad de La Sabana
Chía, Colombia
julianesdi@unisabana.edu.co

[2] Programa de Ingeniería Industrial
Universidad Distrital Francisco José de Caldas
Bogotá, Colombia
erlopezs@udistrital.edu.co

[3] Facultad de Ingeniería
Universidad de La Sabana
Chía, Colombia
william.guerrero1@unisabana.edu.co

Resumen

El cuidado a la salud domiciliaria es una estrategia costo-efectiva de prestación del servicio de salud en la cual personal capacitado atiende a los interesados en sus hogares. Debido al incremento y al envejecimiento de la población, las solicitudes del servicio han aumentado y los límites presupuestales en salud cada vez son más estrictos, por lo que se deben reducir costos utilizando eficientemente los recursos disponibles. Una decisión que influye en los costos operacionales es la secuencia en que el personal asistencial visita a los pacientes considerando sus ubicaciones y la urgencia en su atención. Este capítulo presenta un modelo matemático que determina las rutas óptimas minimizando tiempos de transporte del personal médico y de espera de los pacientes, al mismo tiempo. Además, la aparición o cancelación de solicitudes en forma dinámica y la necesidad de reoptimizar las rutas cuando estos eventos ocurren se evalúa mediante un modelo de simulación multiagente. Lo anterior se lleva a cabo utilizando un caso de estudio real de un proveedor del servicio de cuidado a la salud domiciliaria en la ciudad de Bogotá. Con esta propuesta de modelamiento computacional, se evalúan diferentes políticas operativas. Los resultados mostraron que optimizando las rutas con mayor frecuencia, se obtiene mejor servicio. Sin embargo, los costos operativos aumentan, por lo que es necesario considerar un balance entre ellos.

Palabras clave: cuidado a la salud domiciliaria, ruteo, optimización, simulación, sistemas multiagente

Introducción

En 2017, más de 650 millones de personas cumplieron 65 años o más (World Bank, 2018) y ello evidenció que cada día existen más individuos de la tercera edad que, por lo general, requieren atención médica frecuente. Por otra parte, el aumento de la expectativa de vida y el crecimiento de la población también están ocasionando que se requieran servicios médicos, específicamente servicios de cuidado en casa (Organisation for Economic Cooperation and Development, 2017).

El cuidado en casa o el cuidado domiciliario, sobre todo, consiste en cualquier atención médica o social que permita a los pacientes permanecer en sus viviendas mientras son atendidos por personal capacitado, el cual resuelve sus necesidades en materia de salud y les permite a los más vulnerables permanecer en sus hogares el mayor tiempo posible al tiempo que reciben ayuda profesional (Genet et al., 2012). Con ello se evita que los pacientes sean trasladados y admitidos en instalaciones médicas y se disminuyen costos en los sistemas de salud, por lo que esta modalidad resulta atractiva tanto para pacientes que prefieran recibir el cuidado en casa como para los proveedores del servicio médico (Rest et al., 2012).

En Colombia, la salud domiciliaria es manejada casi en su totalidad por instituciones prestadoras de servicios de salud, de carácter privado, creadas exclusivamente para ofrecer este servicio, con excepción de algunos programas públicos que incluyen esta modalidad en sus portafolios. Un ejemplo de lo anterior se evidencia en el hecho de que solo en la ciudad de Bogotá existen más de veinticinco instituciones privadas dedicadas al cuidado de la salud domiciliaria (Sánchez Cárdenas y Fuentes Bermúdez, 2016). Lo anterior obedece a que en Colombia esta modalidad se está convirtiendo en un sector cada vez más atractivo por su costo-efectividad y por la alta dependencia de la población afiliada. Por lo tanto, la competencia cada vez es mayor.

Con el fin de competir de manera exitosa en el mercado de la salud domiciliaria, que cada día es más demandado, se debe brindar un servicio de calidad y, a su vez, se deben disminuir los costos en su prestación. Para lograrlo es preciso contar con las herramientas adecuadas que permitan tomar decisiones que impacten positivamente en los costos en el momento de las visitas domiciliarias. Entre las decisiones cruciales se encuentra determinar las rutas del personal asistencial que realizará las visitas; en otras palabras, determinar la secuencia en la que el personal asistencial visitará a los pacientes asignados a ellos. Cuando la decisión se toma de forma manual, tal vez se desperdicien recursos y se empeoren los indi-

cadores de prestación, debido a las demoras por malos esquemas de ruteo (Fikar y Hirsch, 2017).

La determinación de las rutas de los equipos asistenciales hace parte de las decisiones operativas que se deben tomar según los requerimientos del servicio, es decir, a cada paciente que solicite un servicio se le debe asignar un equipo que se debe enrutar para llegar lo más pronto posible, de acuerdo con la urgencia en la atención, que se debe reflejar en la priorización de pacientes (Gutiérrez y Vidal, 2013). Es de resaltar que la urgencia en la atención es dada por la sintomatología, es decir, si los síntomas del individuo representan mayor peligro para su vida, mayor es la urgencia en su atención. Así, la decisión de enrutamiento se puede ver como un problema de enrutamiento de vehículos con restricciones adicionales propias del cuidado a la salud domiciliaria, lo cual ha sido ampliamente estudiado en la literatura sobre el tema, en la cual hay distintas aproximaciones que varían en objetivos, supuestos, consideraciones, restricciones y metodologías de solución.

Entre los criterios de optimización más frecuentes para el problema, está la minimización de los tiempos de viaje del personal asistencial, como se observa en el texto de Redjem et al. (2014), en el cual los autores buscan minimizar sus tiempos de viaje, así como sus tiempos de espera. Los objetivos de la minimización de costos también se han propuesto en trabajos previos. Por ejemplo, en Yuan et al. (2014) se tuvieron en cuenta el costo del viaje y la penalización con un costo monetario respecto a las demoras a las llegadas de los pacientes. Por otra parte, en Bard et al. (2014), los autores consideraron costos de horas extra y subcontratación del personal asistencial. La maximización de la satisfacción de los solicitantes es otro factor para considerar en el contexto de los sistemas de cuidado a la salud domiciliaria. En Carello y Lanzarone (2014), los autores buscaron que la compatibilidad de los pacientes y el personal asistencial sea la más alta y que, a su vez, exista continuidad en el cuidado, es decir, que en lo posible sea el mismo personal el encargado de visitar a determinados pacientes que requieren el servicio de manera constante. Por último, otro objetivo frecuente en la literatura es la minimización de la cantidad de personal asistencial necesario, como fue visto en Allaoua et al. (2013). Es de resaltar que, por lo general, las propuestas buscan cumplir más de un objetivo a la vez, como lo analizaron Braekers et al. (2016), donde se propuso un modelo de programación matemático multiobjetivo que pretende minimizar costos de operación y maximizar la satisfacción del cliente.

Otra característica que se debe tener en cuenta son las restricciones propias del cuidado a la salud domiciliaria, como lo son las ventanas de tiempo de los pacientes. En este caso, es mandatorio que se plantee dentro de los esquemas de ruteo que los pacientes no pueden ser atendidos en todo momento, sino que se debe respetar una franja horaria ya establecida (Redjem y Marcon, 2015). Otra condición es garantizar que el personal asistencial esté lo suficientemente califica-

do como para realizar las visitas domiciliarias a determinados pacientes, quienes, por la complejidad de sus síntomas, precisan una atención más especializada. En el texto de Rest y Hirsch (2016) se presentó un modelo matemático en el cual se toma la decisión de qué equipo de personal asistencial lleva a cabo la visita domiciliaria, si su habilidad es mayor o igual que el nivel de especialidad requerida por el paciente.

Del mismo modo que existen ventanas de tiempo para los pacientes, el tiempo de trabajo o el horario laboral del personal asistencial puede considerarse una ventana de tiempo de trabajo. Esta restricción se plantea tanto en Maya Duque et al. (2015) como en Yalçındağ et al. (2016), donde los autores establecen parámetros de tiempo de trabajo máximo que incluyen las rutas iniciales y finales del personal asistencial. Finalmente, dentro de las restricciones más comunes en sistemas de cuidado a la salud domiciliaria está el punto de partida y finalización, que puede estar en una instalación médica como punto de partida y de llegada (un punto) o los hogares de ellos (múltiples puntos) para el inicio y finalización de sus rutas.

Se han utilizado diferentes metodologías para solucionar el problema de ruteo de personal asistencial en el cuidado a la salud domiciliaria. Es de resaltar que el problema en cuestión es de tipo NP-completo, por lo que su tiempo de solución se da en un intervalo no polinómico o exponencial en función del tamaño del problema (Triki et al., 2014). Con base en lo anterior, los métodos exactos funcionan para instancias pequeñas, es decir, con un número reducido de pacientes y equipos de personal asistencial. Ejemplos de estos métodos son ramificación y corte (Rasmussen et al., 2012), generación de columnas (Liu et al., 2015) o programación lineal entera (Issaoui et al., 2015). Por otra parte, y con el fin de manejar instancias más grandes del problema, se han desarrollado distintos métodos de solución aproximados, como algoritmos de búsqueda (Braekers et al., 2016) o metaheurísticas (Hiermann et al., 2015), las cuales explotan las características del problema para encontrar buenas soluciones en un tiempo computacional menor. Finalmente, existen otras aproximaciones o paradigmas que se han utilizado para abordar el problema, incluidas la lógica difusa (Mutingi y Mbohwa, 2014), las cadenas de Markov (Koeleman et al., 2012) y los modelos de simulación multiagente (López-Santana et al., 2016). Los lectores interesados en el tema pueden referirsea Fikar y Hirsch (2017), donde se revisa de manera exhaustiva el estado del arte en ruteo y asignación de personal asistencial, identificando sus objetivos, restricciones, metodologías de solución y futuras direcciones en investigación.

Es de resaltar que el contexto del cuidado a la salud domiciliaria es dinámico, pues es posible que se den eventos que requieran modificar las rutas, como nuevas solicitudes, cancelación de solicitudes, cambios en las ventanas de tiempo de los pacientes o la indisponibilidad de algún equipo de personal asistencial. Lo anterior exige que la planeación del servicio se realice de forma dinámica o en

tiempo real y que el sistema responda adecuadamente (Gutiérrez y Vidal, 2013; Yuan y Fügenschuh, 2015). Este capítulo presenta un modelo de simulación multiagente que captura el dinamismo del cuidado de la salud domiciliaria permitiendo el ingreso de solicitudes y su cancelación. Las rutas se determinan usando un modelo de programación matemática que busca minimizar los tiempos de viaje de los equipos asistenciales y la demora en la atención a los pacientes. Mediante el modelo multiagente se evalúan distintas rutas del personal asistencial, que se obtienen variando la frecuencia de optimización, es decir, se evalúa el impacto de optimizar más frecuentemente, acumulando menos solicitudes, o de esperar a que se acumulen y determinar las rutas óptimas de los equipos asistenciales.

Este trabajo está organizado de la siguiente manera: la sección "Métodos" presenta la metodología empleada, donde se explica el funcionamiento del sistema multiagente, se define formalmente el problema, se plantea el modelo matemático y la configuración del modelo de simulación multiagente y se detalla la implementación computacional de lo propuesto. La siguiente sección presenta un caso de estudio inspirado en un proveedor del servicio de cuidado a la salud domiciliaria de la ciudad de Bogotá, el cual cuenta con información de localización de afiliados, frecuencias de solicitud por localidad y franja horaria. Así, se establecen tres frecuencias de optimización y se obtienen diferentes resultados, analizados en la sección "Análisis y discusión". Finalmente, se concluye y se dan direccionamientos para trabajos futuros en el área.

Métodos

El cuidado a la salud domiciliaria puede considerarse un sistema en el cual interactúan personas que solicitan o requieren el servicio de salud domiciliario. En algunos casos, hay personal asistencial que remite al paciente al servicio y debe haber un acuerdo entre el personal, el paciente y su familia para llevarlo a cabo. Adicionalmente, la entidad aseguradora del paciente debe autorizar el pago a la entidad prestadora del servicio. Una vez llevado a cabo este proceso, hay personal asistencial que realiza las visitas domiciliarias, personal encargado de registrar y actualizar la información del sistema y coordinadores que determinan las rutas de los equipos asistenciales.

De la interacción de las entidades descritas emerge el comportamiento del sistema, cuyo objetivo es la prestación exitosa del servicio de cuidado domiciliario. Un paradigma que puede emplearse para modelar este sistema es la aproximación multiagente, pues existe una clara diferenciación de las entidades que componen el sistema y su posición geográfica, se evidencian interacciones concretas dentro del sistema, hay un proceso de comunicación definido y no está centralizada la toma de decisiones, ya que, a pesar de que existe una entidad que determina y comunica las decisiones de ruteo de los equipos médicos, los pacien-

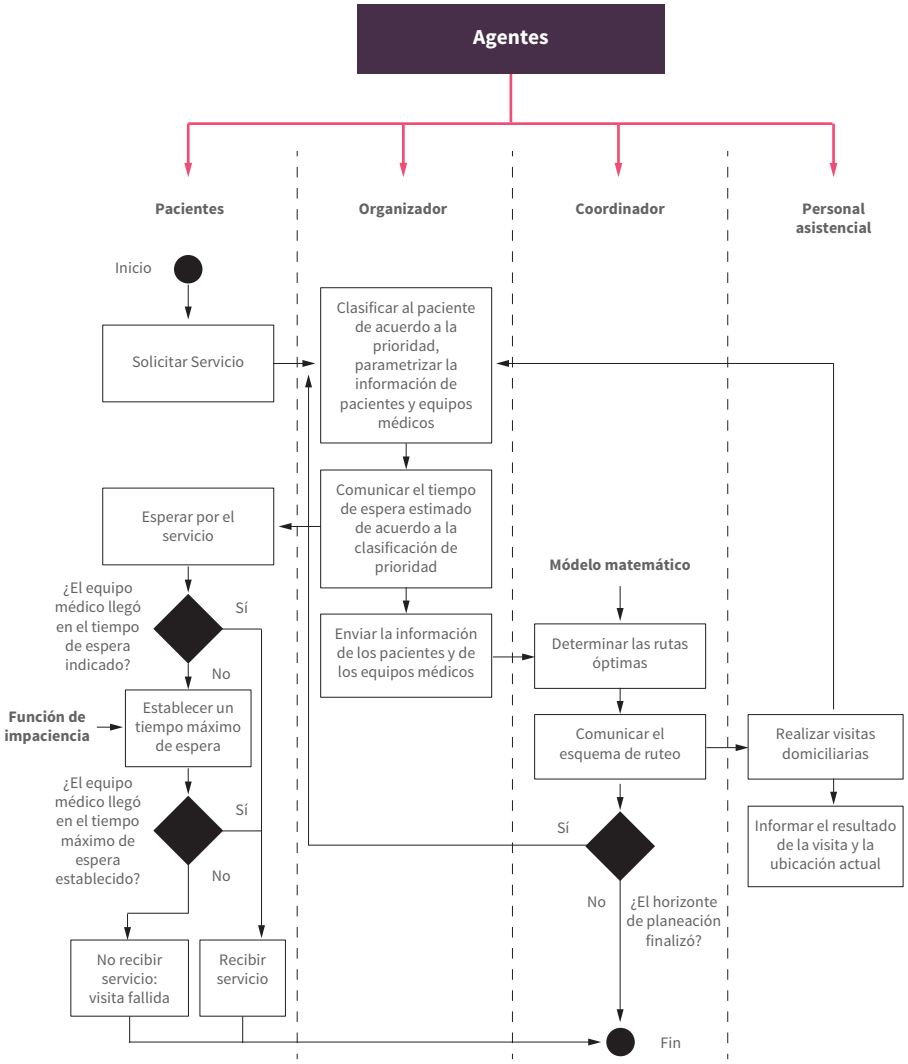
tes pueden decidir rechazar el servicio (por demoras, por ejemplo), lo que cambia el comportamiento del sistema y el resultado de las visitas domiciliarias.

Cabe destacar que el comportamiento descrito y en el que se basa este trabajo de investigación corresponde a un caso en particular de un proveedor que brinda el servicio de cuidado a la salud domiciliaria en Bogotá. Este proveedor tiene adscritos a más de 20 000 usuarios, distribuidos en 23 localidades de la ciudad; cuenta con datos históricos del comportamiento de las solicitudes y del proceso de respuesta que se les da. Usando esta información histórica, se realizaron análisis estadísticos. Finalmente, dentro de su portafolio de cuidado a la salud domiciliaria existen varios planes de atención. En esta investigación solo se tomaron en cuenta las visitas médicas solicitadas por los pacientes de manera telefónica para sintomatologías que pueden esperar el desplazamiento del personal asistencial. En caso de que los síntomas descritos por el paciente pongan en peligro su vida, se le indicará dirigirse a un centro asistencial lo más pronto posible y no se presta atención domiciliaria.

El inicio del sistema de cuidado a la salud domiciliaria bajo estudio se da cuando una persona solicita el servicio. Un organizador registra a información, clasifica al paciente de acuerdo con su prioridad y le indica un tiempo de espera (basado en el *triage*), en el cual el personal asistencial asignado debe llegar. Es de resaltar que el organizador es un funcionario del proveedor del servicio, quien, con base en sus conocimientos y habilidades en salud, determina la urgencia en la atención partiendo de los síntomas descritos. Por otra parte, el *triage* es una clasificación hospitalaria que establece los tiempos de respuesta según la prioridad de los pacientes (Martín García, 2013).

Después, el organizador recoge solicitudes durante determinado intervalo de tiempo, el cual incidirá directamente en las decisiones de ruteo (por ejemplo, de 30 a 60 minutos) y las envía al coordinador, que en el escenario deseado utiliza un modelo de programación matemática que establece las rutas óptimas del personal asistencial, buscando minimizar los tiempos de viaje, así como el tiempo de espera de los solicitantes. El tiempo de recolección de solicitudes es una decisión de diseño del sistema que se debe optimizar. Recolectar solicitudes durante mucho tiempo hará que los pacientes deban esperar más, pero permitirá distribuir mejor los equipos asistenciales. Recolectar las solicitudes durante muy poco tiempo disminuye el tiempo de espera, pero puede hacer que las rutas de viaje sean más largas (menos eficientes). El esquema de ruteo es comunicado a los equipos de personal asistencial que realizan las visitas domiciliarias. Ellos informan el resultado de las visitas y su ubicación actual al organizador del sistema, el cual ya ha recogido nuevas solicitudes en un nuevo periodo hasta la finalización del horizonte de planeación establecido. Este periodo puede corresponder a un turno o día de planeación según los requerimientos del proveedor. Lo anterior es ilustrado con detalle en la figura 1, en forma de diagrama de flujo.

Figura 1. Configuración del sistema de cuidado a la salud domiciliaria del caso de estudio

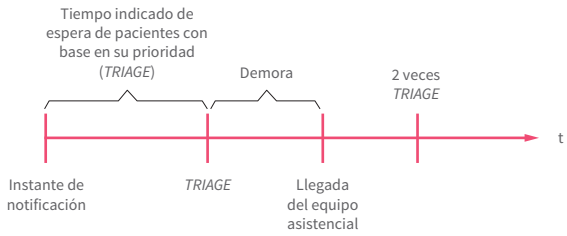


Fuente: elaboración propia.

Por otra parte, en caso de que se presenten demoras en el servicio, es decir, los equipos de personal asistencial no lleguen en el tiempo de espera indicado al paciente, este último puede decidir no aceptar la visita y buscar otras alternativas. Por ejemplo, optar por ir a un centro de salud cercano, y ello resulta en una visita domiciliaria fallida. En la figura 2 se presenta la línea de tiempo de una solicitud del servicio en la cual existe una demora generada porque se excedió el tiempo de

llegada del equipo de personal asistencial al hogar del paciente luego del tiempo de espera indicado. Es de aclarar que la secuencia de actividades presentada en la figura 1 no contempla el proceso administrativo previo, como la inscripción y autorización del servicio al paciente, sino desde que el paciente ya se encuentra inscrito y requiere utilizar el servicio de cuidado a la salud domiciliaria. El sistema considerado entra en función cuando la solicitud ha sido aprobada por la entidad aseguradora y llega al coordinador del servicio.

Figura 2. Línea de tiempo de solicitud de servicio de cuidado a la salud domiciliaria

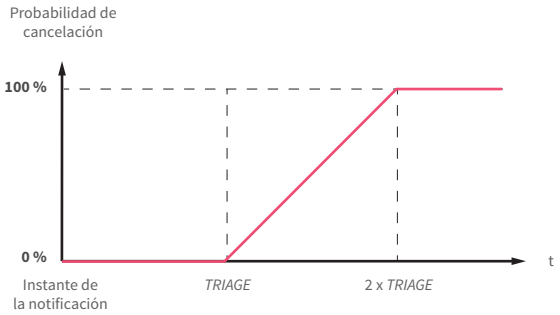


Fuente: elaboración propia.

Para el caso de estudio, se realizó el supuesto de que la impaciencia y posterior cancelación del servicio se puede modelar de acuerdo con una distribución uniforme continua con valores entre 0 y 1. Entre menor sea este valor, el paciente esperará menos tiempo, en caso de que personal asistencial incurra en demoras. Lo anterior es comparado con el porcentaje de demora respecto al tiempo de espera prometido. Así, entre mayor sea la demora, existe mayor probabilidad de que el paciente cancele el servicio, lo que corresponde a la realidad. Es de resaltar que este supuesto se dio en ausencia de metodologías o estudios que aborden la impaciencia de los usuarios de servicios de cuidado a la salud domiciliaria en la literatura disponible. La investigación futura incluye afinar el modelamiento probabilístico del comportamiento de las cancelaciones de los solicitantes.

A manera de ejemplo, considere a un paciente que tiene un valor de impaciencia de 0,2 o de un 20 % (generado aleatoriamente con la distribución uniforme continua). Si el personal asistencial llega con más del 20 % de demora respecto al tiempo indicado (basado en su *triage*), el paciente buscará otra alternativa y la visita será fallida. La figura 3 muestra lo descrito. Finalmente, para pacientes que describan sintomatologías que ponen en riesgo su vida, se le indicará solicitar asistencia al servicio de ambulancias de la ciudad o dirigirse al centro de salud más cercano lo más pronto posible. Por lo anterior, en el sistema no se aceptan pacientes con nivel de *triage* 1, que no tiene tiempo de espera.

Figura 3. Probabilidad de cancelación de la visita domiciliaria cuando hay demoras



Fuente: elaboración propia.

En la siguiente subsección se define formalmente el problema de ruteo de equipos de personal asistencial, estableciendo los objetivos, las consideraciones y las características del problema. Luego, en la subsección "Modelo de programación matemática del problema" se determinan las rutas óptimas de los equipos que son utilizadas en el modelo de simulación multiagente mostrado en detalle en la siguiente subsección. Finalmente, se cierra la sección describiendo la implementación de la propuesta en un *software* especializado en agentes y en modelos de programación matemática. De esa manera, se expone la propuesta de simulación-optimización empleada para evaluar las rutas de equipos o personal asistencial en el cuidado de la salud domiciliaria.

Contexto del problema

Con el fin de determinar las rutas óptimas del personal asistencial que visitará y atenderá a los pacientes, se considera un conjunto de pacientes $P = \{1, 2, \ldots, N\}$ que solicitan el servicio y están geográficamente distribuidos en un área determinada. Se asume que cada paciente representa una visita. Si requiere más visitas, estas se tratarán como solicitudes diferentes. Con el fin de visitar a los pacientes y brindar el servicio de cuidado a la salud domiciliaria, el proveedor cuenta con un conjunto de equipos de personal asistencial $K = \{1, 2, \ldots, M\}$, los cuales tienen diferentes habilidades. Es decir, son heterogéneos en cuanto a su preparación y capacidad de atención. Es importante resaltar que los equipos están distribuidos geográficamente cuando inician y finalizan su jornada laboral; en otras palabras, parten desde sus hogares, cumplen las visitas domiciliarias que se les asignan y regresan de nuevo a sus hogares.

Considerando los conjuntos de pacientes y de personal asistencial descritos, el problema de ruteo en el cuidado a la salud domiciliaria puede definirse mediante un grafo $G = (V, A)$, el cual está conformado por un conjunto de nodos $V =$

$KI \cup P \cup KF$, donde KI representa los nodos de inicio del personal asistencial (sus hogares), P representa los hogares de los pacientes que se van a visitar y KF es un conjunto de nodos ficticios que equivalen al destino de los equipos asistenciales. Como se resaltó, los equipos salen y regresan a sus hogares; por lo tanto, $KI = KF$ $\forall\ k \in K$. Por otra parte, los arcos correspondientes al grafo están definidos por el conjunto $A = \{(i, j): i, j \in V, i \neq j\}$, donde cada arco (i, j) representa las conexiones entre los nodos que, a su vez, tienen asociado un parámetro que indica el tiempo de viaje t_{ij} entre ellos. Cuando el paciente solicita el servicio en el instante a_i, es clasificado según su prioridad p_i en una escala de 2 a 5, siendo 2 la categoría más urgente dentro del sistema estudiado, en el cual la sintomatología permite espera del personal asistencial al hogar del solicitante. No se toma en cuenta el nivel de *triage* 1, debido a que es urgente. Por otra parte, a cada visita domiciliaria le corresponde una duración s_i y es realizada por un equipo de personal asistencial con la habilidad o experticia q_k, la cual debe ser igual o superior a la habilidad o experticia requerida por los pacientes h_i. La habilidad requerida por los pacientes está dada en función de la atención solicitada: entre más compleja sea la visita domiciliaria, mayor será este valor. Los tiempos de espera se determinan de acuerdo con su prioridad p_i, que equivale a la clasificación de prioridad hospitalaria *triage*. En la tabla 1 se presentan los tiempos caracterizados por Martín García (2013), los cuales son utilizados en este trabajo, y el tiempo de espera máximo tolerado por los pacientes según la propuesta realizada.

Tabla 1. Tiempos de atención basados en el nivel de clasificación *triage*

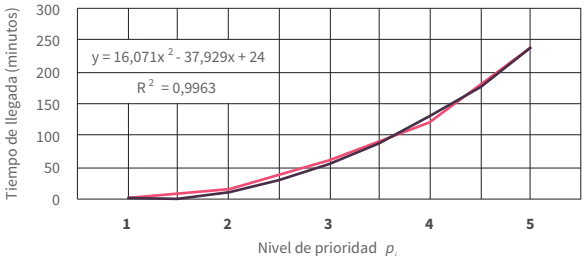
Nivel de *triage*	Tiempo de espera para atención (en minutos)	Tiempo de espera máximo tolerado por los pacientes (en minutos)
1	0	Nivel no considerado
2	15	30
3	60	120
4	120	240
5	240	480

Fuente: adaptado de Martín García (2013).

Partiendo de los tiempos de espera presentados, se elaboró una función cuadrática para obtener los tiempos indicados de espera de los pacientes con base en su prioridad. Los coeficientes de la función cuadrática son $\gamma = 16{,}071$; $\delta = -37{,}929$ y $\varepsilon = 24$, los cuales se obtuvieron con una regresión cuadrática de los tiempos de la tabla 1. En la figura 4 se representa la función cuadrática y se mues-

tra la cota inferior del tiempo de atención a los pacientes. Como se determinó en un inicio, el objetivo de la propuesta es minimizar los tiempos de transporte del personal asistencial y las demoras en el servicio; por lo tanto, en una situación ideal los tiempos de espera de los pacientes son los representados por la función, es decir, sin demoras.

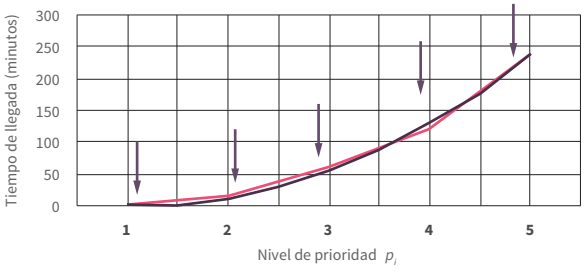
Figura 4. Función cuadrática del tiempo de atención a los pacientes con base en su prioridad



Fuente: elaboración propia.

Para lograr el objetivo de minimizar el tiempo de transporte y la demora, los tiempos de llegada tienen que acercarse a los tiempos de atención determinados por el *triage*, lo que es representado en la figura 5.

Figura 5. Minimización de demoras



Fuente: elaboración propia.

Por otra parte, se deben considerar las restricciones propias de los sistemas de cuidado domiciliario, que son:

- Es obligatorio que se respeten las habilidades requeridas por el procedimiento o la visita médica que se va a realizar a los pacientes. Esto es, no

se puede enviar personal asistencial que no esté suficientemente calificado para llevar a cabo procedimientos complejos.
- Existen horarios definidos de los equipos de personal asistencial.
- Existen puntos geográficos de inicio y finalización de las rutas de los equipos de personal asistencial, por ejemplo, sus hogares.
- Se deben respetar las duraciones mínimas y máximas de las visitas domiciliarias.

Lo anterior, junto con los objetivos de la propuesta, se detalla en la siguiente sección, que es la formulación matemática del problema de ruteo de equipos asistenciales.

Modelo de programación matemática del problema

Con el objetivo de determinar las rutas óptimas del personal asistencial, se propone un modelo de programación entero mixto. En la tabla 2 se encuentra la notación utilizada en el modelo matemático correspondiente a los conjuntos, parámetros y variables de decisión.

Tabla 2. Notación del modelo matemático

Conjuntos
P: conjunto de pacientes
K: conjunto de operadores de salud
KI: conjunto de orígenes de los operadores de salud
KF: conjunto de destinos de los operadores de salud
V: conjunto de vértices ($V = KI + P + KF$)
A: conjunto de arcos
$V1$: conjunto de vértices 1 ($V1 = KI + P$)
$V2$: conjunto de vértices 2 ($V2 = P + KF$)
Parámetros
t_{ij}: tiempo de viaje desde el paciente o personal asistencial i al paciente o personal asistencial j [en minutos]
a_i: momento en el cual el paciente i comunica la necesidad del servicio [en minutos]
s_i: duración de la visita domiciliaria al paciente i [en minutos]
h_i: habilidad mínima requerida por el paciente i [1: nivel de requerimiento más bajo, 5: nivel de requerimiento más alto]
p_i: prioridad en la atención del paciente i en el sistema [2: prioridad máxima, 5: prioridad mínima]
q_k: habilidad del operador de salud k [1: nivel de experticia más bajo, 5: nivel de experticia más alto]
l_k: comienzo de la jornada laboral del operador de salud k [en minutos]
u_k: finalización de la jornada laboral del operador de salud k [en minutos]
M: valor numérico alto para restricciones lógicas
α: ponderación para la primera función objetivo
β: ponderación para la segunda función objetivo
G: tamaño del conjunto P más el tamaño del conjunto KI
$\gamma, \delta, \varepsilon$: coeficientes de la función *triage*

Variables
X_{ijk}: 1, si el operador de salud *k* cubre la ruta (*i* - *j*) y 0, en otro caso
Y_{ik}: momento de llegada del operador de salud *k* donde el paciente *i* [en minutos]
Z_i: demora en la llegada donde el paciente *i* [en minutos]

Fuente: elaboración propia.

La función objetivo corresponde a la minimización de los tiempos de transporte del personal asistencial y la minimización de la demora a los pacientes ponderados utilizando unos pesos α y β, como se muestra en las ecuaciones (1) a (3).

$$Minimizar\, FO = \alpha\, FO_1 = \beta\, FO_2 \tag{1}$$

$$FO_1 = \sum_{i \in V} \sum_{j \in V} \sum_{k \in K} X_{ijk}\, t_{ij} \tag{2}$$

$$FO_2 = \sum_{i \in P} Z_i \tag{3}$$

Las restricciones del modelo de programación entero mixto corresponden a las ecuaciones (4) a (19).

$$\sum_{k \in K} \sum_{i \in V1} X_{ijk} = 1 \qquad \forall j \in P \tag{4}$$

$$\sum_{j \in V1} X_{jek} = \sum_{j \in V2} X_{ejk} \qquad \forall k \in K,\, e \in P \tag{5}$$

$$\sum_{j \in V2} X_{ijk \mid i=k} = 1 \qquad \forall k \in K,\, i \in KI \tag{6}$$

$$\sum_{i \in KI} \sum_{j \in V2} X_{ijk} = 0 \qquad \forall k \in K \tag{7}$$

$$\sum_{i \in V1} X_{ik+Gk} = 1 \qquad \forall k \in K \tag{8}$$

$$\sum_{i \in V1} \sum_{j \in KF \mid j \neq K+G} X_{ijk} = 0 \qquad \forall k \in K \tag{9}$$

$$Y_{ij} + t_{ij} + s_i - M(1 - X_{ijk}) \leq Y_{jk} \qquad \forall k \in K, (i,j) \in P \tag{10}$$

$$Y_{ik} + t_{ij} - M(1 - X_{ijk}) \leq Y_{jk} \qquad \forall k \in K, i \in KI, j \in V2 \tag{11}$$

$$Y_{ik} + t_{ij} - M(1 - X_{ijk}) \leq Y_{jk} \qquad \forall k \in K, i \in V1, j \in KF \tag{12}$$

$$Y_{ik} - (\gamma p_i^2 - \delta p_i + \varepsilon + \alpha_i) \leq Z_i \qquad \forall i \in P, k \in K \tag{13}$$

$$h_j - M(1 - X_{ijk}) \leq q_k \qquad \forall k \in K, i \in V1, j \in P \tag{14}$$

$$Y_{ik} \geq l_i \qquad \forall k \in K, i \in KI \tag{15}$$

$$Y_{jk} \leq u_j \qquad \forall k \in K, j \in KF \tag{16}$$

$$X_{ijk} \in \{0, 1\} \qquad \forall (i,j) \in V | i \neq j, k \in K \tag{17}$$

$$Y_{ik} \geq 0 \qquad \forall i \in V, k \in K \tag{18}$$

$$Z_i \geq 0 \qquad \forall i \in P \tag{19}$$

La restricción (4) asegura que todos los pacientes sean visitados y atendidos. En la (5) se determina la conservación del flujo, es decir, si un equipo de personal asistencial visita a un paciente, este debe irse al finalizar la visita domiciliaria. Las restricciones (6), (7), (8) y (9) garantizan que cada equipo asistencial salga y regrese a su propio hogar. Para los equipos asistenciales el tiempo de llegada donde los pacientes, el tiempo de salida y los regresos a sus hogares es calculado en (10), (11) y (12). El cálculo de la demora en la llegada de los equipos de personal asistencial es determinado en (13), siendo esta la diferencia entre la hora

de llegada y el tiempo de atención determinado por el *triage* con base en la prioridad. En (14), la habilidad del equipo asistencial para realizar la visita domiciliaria es observada. En (15) y en (16) se contempla la jornada laboral de los equipos de personal asistencial. En (17) se asegura que las variables X_{ijk} sean binarias, y en (18) y (19), que el tiempo de llegada y la promesa de servicio sean no negativas.

Modelo de simulación multiagente

Con el fin de capturar la naturaleza dinámica del cuidado a la salud domiciliaria, en la que nuevas solicitudes del servicio aparecen en el tiempo y existen cancelaciones, debido a las demoras, se construyó un modelo de simulación basado en agentes. Por otra parte, con el modelo de simulación también es posible contemplar la comunicación y las múltiples interacciones entre las entidades del sistema acercando la propuesta a la realidad. En esta subsección se detalla la definición de las entidades del sistema, sus funciones, procesos de interacción y comunicación para finalmente mostrar cómo se puede configurar el sistema para evaluar distintas estrategias de reoptimización de las rutas médicas obtenidas con el modelo matemático presentado.

Como se indicó, el sistema de cuidado a la salud domiciliaria está compuesto por cuatro entidades principales, las cuales son:

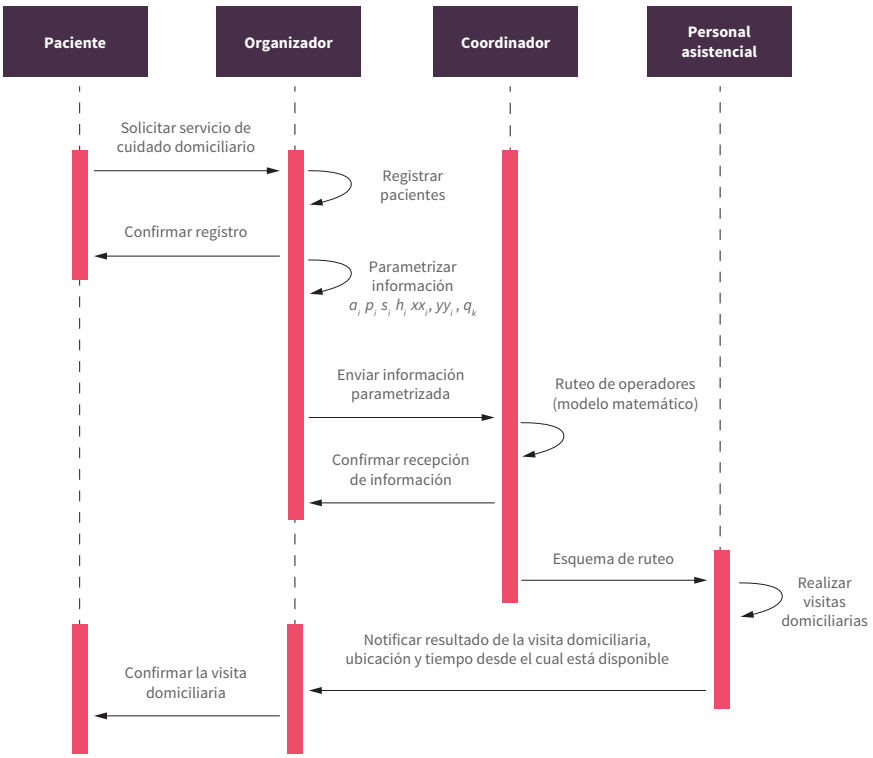
- Paciente: entidad que requiere atención médica en casa; por lo tanto, solicita el servicio de cuidado a la salud domiciliaria a su proveedor de salud. Cuando se contacta, debe suministrar su información personal para conocer su ubicación exacta y describir sus síntomas, con el fin de categorizarlos según su prioridad y que le sea brindado un tiempo de espera estimado con base en su urgencia de atención.
- Organizador: entidad especializada en clasificación *triage*. Su función principal es clasificar a los pacientes de acuerdo con su prioridad e indicarles el tiempo de espera estimado en el cual el equipo del personal asistencial asignado llegará a visitarlos. Por otra parte, la entidad también debe registrar la información concerniente al paciente, como ubicación actual, momento de solicitud del servicio y nivel de experticia requerido por el paciente que el equipo de personal asistencial debe cumplir para atenderlo. Finalmente, esta entidad también se encarga de actualizar las ubicaciones de los equipos de personal asistencial y los resultados de las visitas domiciliarias. Toda la información anterior es comunicada al coordinador que tomará las decisiones de ruteo en el sistema.
- Coordinador: entidad cuya función principal es determinar las rutas óptimas de los equipos asistenciales que atenderán a los pacientes. Para lograrlo utiliza la información del sistema recolectada por la entidad organizadora y, mediante un modelo matemático, establece el esquema de ruteo que minimice los tiempos de viaje de los equipos asistenciales y

la demora en el servicio, como se evidenció en el modelo matemático presentado. Por último, comunica los pacientes asignados y la secuencia de visitas al personal asistencial.

- Personal asistencial: entidad encargada de atender a los pacientes en sus hogares de acuerdo con la secuencia comunicada por el coordinador. Cuando visita a cada paciente, informa al organizador el resultado de la visita (fallida o exitosa) y su ubicación actual, con el fin de determinar las rutas del siguiente esquema de ruteo.

Los procesos de interacción y de comunicación de las entidades del sistema están representados en la figura 6.

Figura 6. Secuencia de comunicación del sistema de cuidado a la salud domiciliaria

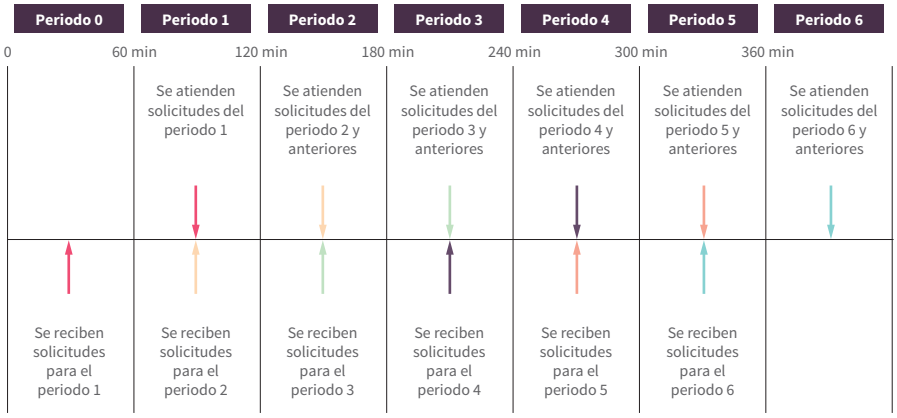


Fuente: elaboración propia.

Por último, las solicitudes en el sistema multiagente se recolectan en un intervalo de tiempo determinado y se atienden después de cerrarlo, como se muestra en figura 7. A manera de ejemplo, imagine una configuración de recepción de solicitudes cada 40 minutos: en el periodo 1 (0-40 minutos) solo se reciben

solicitudes que se atienden después de finalizar el periodo 1 y, simultáneamente, en el periodo 2 (40-80 minutos) se recolectan las solicitudes que serán atendidas después de finalizar el periodo 2. Por lo tanto, un paciente que solicite el servicio en un determinado instante del periodo 1 debe esperar a que se termine el periodo de recolección de solicitudes para que se le asigne un equipo asistencial, el cual lo visitará idealmente en el siguiente periodo o después de este, teniendo en cuenta que el paciente ya esperó determinado tiempo de cierre del periodo de recolección, y con base en su prioridad lo visitará de primero o no en la secuencia. Por lo tanto, un problema relevante para la operación es determinar cuál es la longitud óptima de cada periodo, en ella se hará la recolección de las solicitudes. *A priori* se puede plantear la hipótesis que hacer este periodo muy pequeño implicará que la asignación de personal asistencial no considera suficientes solicitudes y se puede estar realizando una asignación que es operativamente muy costosa, por no considerar que están llegando solicitudes de manera dinámica. Por el contrario, hacer los periodos muy largos obliga a esperar de modo significativo, lo que deteriora el nivel de servicio.

Figura 7. Recolección de solicitudes por periodos



Fuente: elaboración propia.

Es importante resaltar que en esta investigación se evaluaron las rutas de los equipos asistenciales en función de la longitud del periodo de recolección de las solicitudes, es decir, cómo es el comportamiento del sistema si se optimiza en 40, 60 o 90 minutos evaluando el nivel de servicio, así como los costos generales del sistema. En intentos preliminares se probó con optimizar las decisiones cada 20 minutos, pero en muchas ocasiones no hubo requerimientos nuevos en este intervalo de tiempo. Optimizar en intervalos mayores a 90 minutos es operativamente infactible, porque los pacientes no toleran una espera tan larga.

Implementación de la propuesta

El modelo de simulación multiagente y el modelo de programación matemática se implementaron en el lenguaje de programación Java, utilizando las siguientes bibliotecas de programas e interfaces:

- Jade (Bellifemine et al., 2005): biblioteca de programas especializada en construcción de sistemas multiagente que permite la ejecución simultánea de las distintas clases de agentes en diferentes hilos de ejecución (*threads*). Por lo tanto, no existe y no se especifica un método principal (*main*), sino que se crean y ejecutan los tipos de agentes simultáneamente. Cada tipo de agente fue construido en una clase distinta, teniendo cuatro clases que representan las cuatro entidades del sistema. Las distintas entidades se comunican mediante mensajes que, al ser recibidos, modifican los comportamientos de los agentes.
- Xpress-MP BLC (Builder Component Library): interfaz que comunica a Java con el *software* especializado en optimización Xpress-MP, el cual permite crear y ejecutar modelos de programación matemática directamente en Java. Por lo tanto, el *software* de optimización es utilizado como una biblioteca de programas que cuenta con distintos objetos para la creación de variables, restricciones, funciones objetivo, entre otros.
- JSC (Java Statistical Classes): biblioteca de programas utilizada para generar la variable aleatoria que representa la impaciencia en la demora del servicio para cada paciente. Con la fijación de la semilla, se pueden realizar diferentes experimentaciones sin cambiar las probabilidades para cada paciente, lo cual facilita el manejo estadístico en la aproximación multiagente.
- JAVA IO: biblioteca de programas utilizada para crear, escribir, leer y almacenar información en archivos de texto plano que se utilizan frecuentemente en la aproximación multiagente.

Lo anterior se codificó en el entorno de desarrollo integrado Eclipse, en un equipo con una memoria RAM de 6 GB y 2,5 GHz de procesador con el *software* licenciado Xpress-MP.

Resultados

Se cuenta con información real de un proveedor de cuidado a la salud domiciliaria en Bogotá, con más de 20 000 usuarios inscritos y georreferenciados, divididos en 23 localidades y que solicitan el servicio con una frecuencia por localidad que se muestra en la tabla 3, la cual corresponde a tres meses de datos históricos detallados de solicitudes.

Tabla 3. Usuarios inscritos y frecuencia de solicitudes por localidad

ID	Localidad	Usuarios Inscritos	Promedio de solicitudes mensuales	Frecuencia de solicitudes mensuales (en porcentaje)
1	Suba	4881	169	19,29
2	Engativá	3008	157	17,92
3	Usaquén	2917	98	11,19
4	Kennedy	2195	123	14,04
5	Fontibón	1950	90	10,27
6	Teusaquillo	1355	45	5,14
7	Barrios Unidos	1013	41	4,68
8	Chapinero	886	32	3,65
9	Puente Aranda	748	22	2,51
10	Santafé	402	11	1,26
11	Los Mártires	283	12	1,37
12	Antonio Nariño	263	9	1,03
13	Bosa	232	15	1,71
14	Rafael Uribe	228	12	1,37
15	San Cristóbal	220	12	1,37
16	La Candelaria	200	5	0,57
17	Tunjuelito	174	7	0,80
18	Ciudad Bolívar	156	9	1,03
19	Usme	73	2	0,23
20	Comuna 5	51	2	0,23
21	Comuna 1	34	1	0,11
22	Comuna 2	32	1	0,11
23	Comuna 6	18	1	0,11
Total	-	21319	876	100

Fuente: datos históricos de un proveedor del servicio de cuidado a la salud domiciliaria (nombre confidencial).

Se desea conocer cuál es el periodo de optimización que mejore los indicadores del servicio de cuidado a la salud domiciliaria a un menor costo, representado en un balance entre los tiempos de viaje de los equipos de personal asistencial y los tiempos de demora en la atención a los pacientes. Con base en el modelo matemático, se establece una ponderación o importancia denotada con la letra griega α para la función objetivo de minimización de tiempos de viaje de los equipos de personal asistencial y ponderación β para la función objetivo de tiempo de demora de atención a los pacientes. Naturalmente, los pesos deben sumar 1; por lo tanto, $\alpha+\beta=1$. Se experimentó con un horizonte de planeación de 6 horas (tiempo de un turno de un equipo asistencial) con 48 pacientes (solicitudes promedio en un turno de servicio) escogidos aleatoriamente, según la frecuencia de solicitudes por localidad. Es de resaltar que los tiempos de viaje entre ubicaciones, ya sea entre los hogares de los pacientes o entre los hogares de los equipos asistenciales y los pacientes, se calcularon teniendo en cuenta la velocidad promedio de viaje en Bogotá, que es de 24,98 km/h (Secretaría Distrital de Movilidad, 2015), y se utilizó la distancia euclidiana (20) entre 2 puntos para determinar la distancia de recorrido total, que es multiplicada por la velocidad para obtener el tiempo de desplazamiento:

$$d_{P1,\,P2} = \sqrt{(Latitud_2 - Latitud_1)^2 + (Longitud_2 - Longitud_1)^2} \qquad (20)$$

Configuración de la experimentación

Se supone un proveedor del servicio de cuidado a la salud domiciliaria que cuenta con cuatro equipos de personal asistencial, ubicados en distintos lugares de la ciudad y con distintas habilidades, como se muestra en la tabla 4. El nivel de habilidad 5 representa al equipo más experimentado; mientras que el nivel de habilidad 3 indica al equipo de personal asistencial más novato para este caso de estudio.

Tabla 4. Georreferenciación y habilidades de los equipos asistenciales

Equipo asistencial	Latitud	Longitud	Habilidad
1	4,702898	−74,105946	5
2	4,672673	−74,047434	4
3	4,609888	−74,148019	4
4	4,58058	−74,105962	3

Fuente: elaboración propia.

Por otra parte, se escogieron aleatoriamente 48 pacientes con base en la cantidad de usuarios y las frecuencias de solicitudes por localidades de la tabla 3. La georreferenciación de los usuarios elegidos para la experimentación se encuentra en el anexo del capítulo. Se resalta que el tiempo de la simulación se cuenta en milisegundos, es decir, cada segundo equivale a 1000 unidades de tiempo en la simulación. Los valores de los parámetros de prioridad, habilidad requerida, momento de notificación y duración de la visita se establecieron con variables de tipo uniforme discretas, a excepción de los tiempos entre notificaciones, que se dejaron constantes cada 7,5 minutos (tiempo promedio entre arribos de solicitudes de pacientes según la base de datos histórica), como se muestra en la tabla 5. La escala de habilidades requeridas para cada visita domiciliaria se construyó con base en la sintomatología descrita y posteriores procedimientos que se tenían que realizar. Por ejemplo, en la escala 1, el personal asistencial valora al paciente y le suministra medicamentos; mientras que en un nivel superior, como el 2 o 3, el personal asistencial debe suministrar medicamentos como antibióticos vía intravenosa, limpiar heridas o tomar decisiones de mayor complejidad, que requieren más experticia y habilidad. En la tabla 6 se presenta el resultado de la parametrización de la información para los pacientes.

Tabla 5. Estimación de los parámetros de los pacientes

Parámetro	Valor	Observación
Prioridad	Uniforme discreta [2,5]	2 mayor prioridad
Habilidad requerida	Uniforme discreta [1,4]	4 mayor habilidad
Duración de la visita domiciliaria	Uniforme discreta [20 000, 40 000]	20-40 minutos
Tiempos entre notificaciones	7500	Cada 7,5 minutos

Fuente: elaboración propia.

Tabla 6. Parámetros de los pacientes

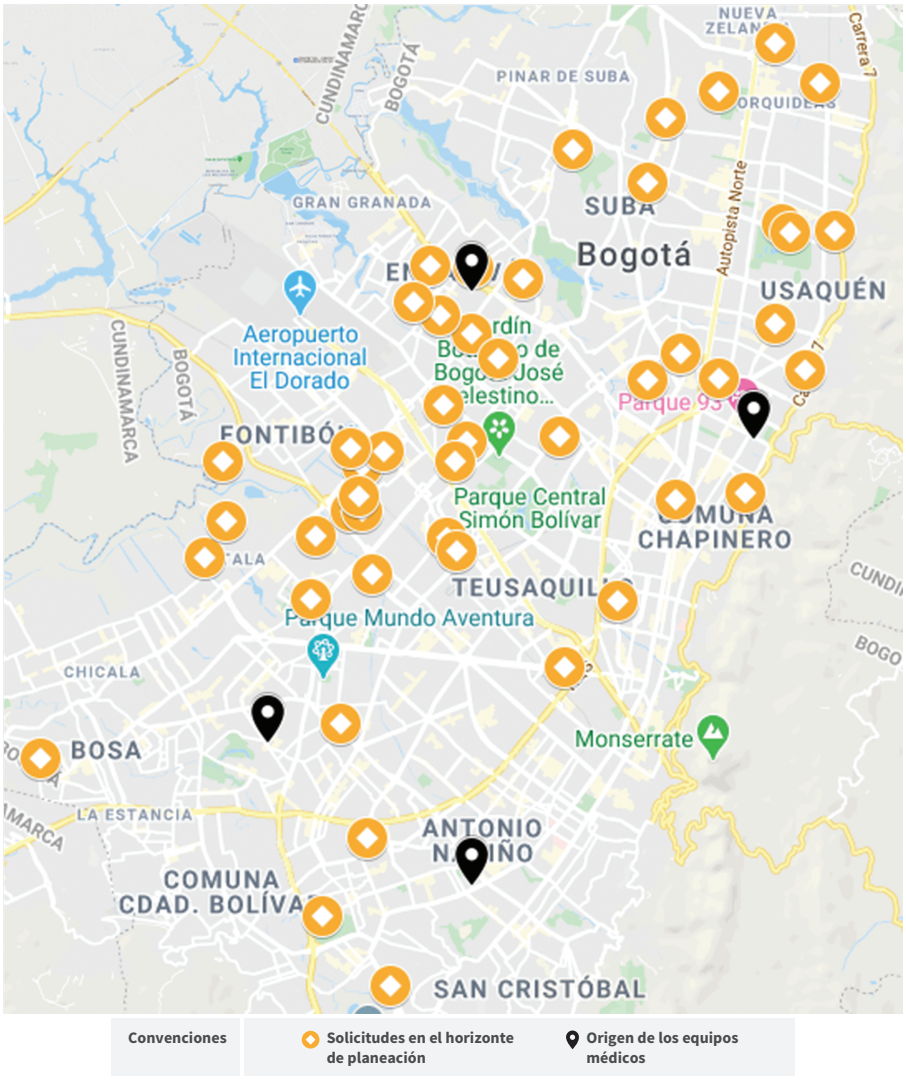
ID	Localidad	p_i	h_i	Tiempo del servicio (en milisegundos)	ID	Localidad	p_i	h_i	Tiempo servicio (en milisegundos)
1	Suba	5	2	27 360	25	Suba	5	3	34 632
2	Kennedy	3	1	39 394	26	Kennedy	4	4	32 718
3	Engativá	2	3	21 166	27	Usaquén	5	1	24 331
4	Engativá	5	4	25 422	28	Usaquén	5	2	24 118

ID	Localidad	p_i	h_i	Tiempo del servicio (en milisegundos)	ID	Localidad	p_i	h_i	Tiempo servicio (en milisegundos)
5	Chapinero	5	3	29 165	29	Fontibón	3	2	28 919
6	Usaquén	5	4	38 656	30	Barrios Unidos	2	4	28 733
7	Bosa	3	4	37 868	31	Fontibón	3	1	33 203
8	Engativá	2	3	21 527	32	Fontibón	3	2	21 566
9	Usaquén	2	3	37 621	33	Kennedy	5	4	25 029
10	Usaquén	4	3	37 194	34	Kennedy	2	3	38 390
11	Fontibón	5	2	21 407	35	Engativá	4	1	30 874
12	Barrios Unidos	2	2	31 423	36	Engativá	5	4	21 466
13	Tunjuelito	2	3	35 634	37	Engativá	4	4	22 295
14	Rafael Uribe	5	2	34 563	38	Engativá	5	4	26 085
15	Engativá	2	3	20 298	39	Engativá	3	3	23 203
16	Kennedy	4	3	21 580	40	Chapinero	4	1	28 945
17	Usaquén	3	2	21 448	41	Fontibón	2	4	28 325
18	Teusaquillo	4	4	24 559	42	Usaquén	3	1	20 757
19	Fontibón	5	2	26 078	43	Fontibón	2	4	30 855
20	Fontibón	3	2	38 261	44	Kennedy	4	4	32 078
21	Suba	2	3	22 780	45	Rafael Uribe	5	3	34 384
22	Barrios Unidos	3	2	35 323	46	Engativá	4	3	36 779
23	Fontibón	4	2	27 887	47	Teusaquillo	3	1	29 074
24	Suba	5	3	26 290	48	Suba	2	1	23 620

Fuente: elaboración propia.

En la figura 8 se observa la ubicación inicial de los equipos de personal asistencial (alfileres) y la de los 48 pacientes (puntos) atendidos en el horizonte de planeación con un tiempo entre las solicitudes de 7,5 minutos.

Figura 8. Ubicación de los pacientes en el mapa de Bogotá

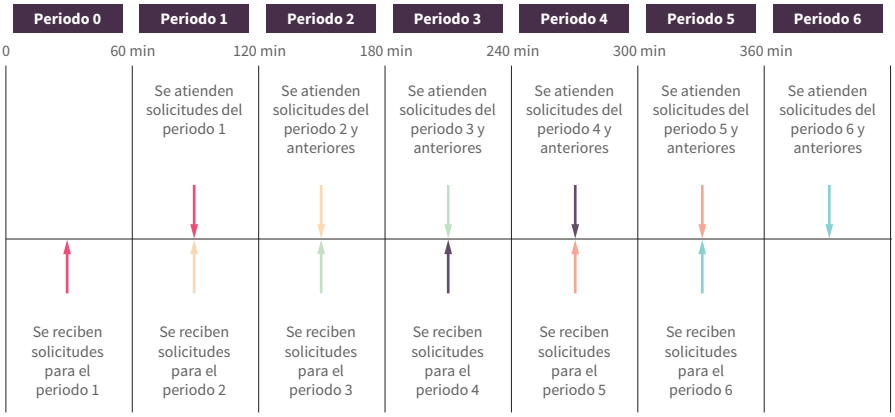


Fuente: elaboración propia con Google Maps.

A manera de otro ejemplo: considere un periodo de optimización de 60 minutos, en el cual en el periodo 1 ingresan 8 solicitudes para las cuales se organiza

un esquema de ruteo óptimo que es comunicado a los equipos asistenciales para atender a esos pacientes en el periodo 2 (60-120 minutos) o después; mientras tanto, otras 8 solicitudes están ingresando al sistema, que serán atendidas después de finalizar el periodo 2, así hasta finalizar el horizonte de planeación. Lo anterior se representa en la figura 9.

Figura 9. Ejemplo de configuración del periodo de simulación de 60 minutos



Fuente: elaboración propia.

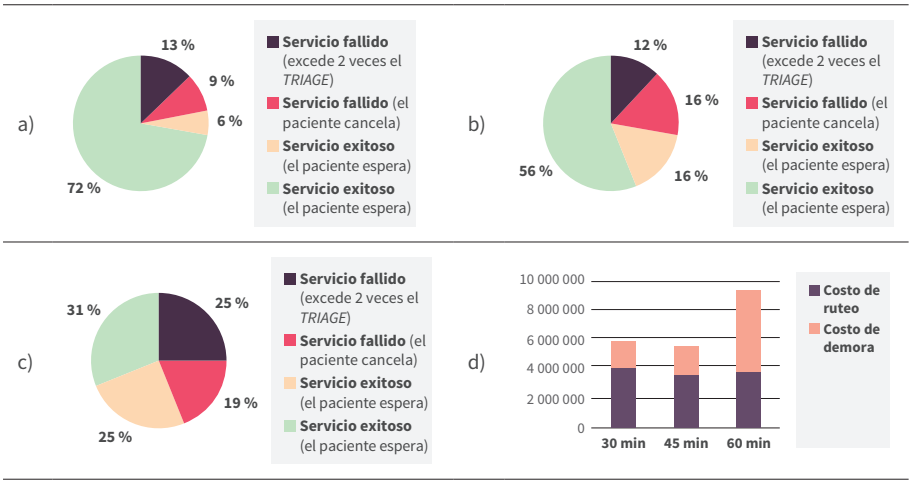
Se experimentó con periodos de optimización de 30, 45 y 60 minutos con un α (importancia del tiempo de viaje en la función objetivo) y una β (importancia de la demora en la función objetivo) de 0,5. Es importante resaltar que los parámetros de los pacientes (prioridades, localizaciones, habilidades requeridas, tiempo de servicio y momento de notificación) y los parámetros de los equipos de personal asistencial (habilidades y puntos de partida) fueron los mismos para los distintos periodos de optimización evaluados. Finalmente, y con el fin de alcanzar el estado estable del sistema, se utiliza un tiempo de calentamiento para las simulaciones.

Resultados de la experimentación

En la figura 10 se muestran los resultados correspondientes a los indicadores de servicio y los costos incurridos con las diferentes estrategias de optimización utilizando unos pesos de $\alpha=\beta=0{,}5$ que establecen la importancia del tiempo de viaje de los equipos asistenciales y las demoras a los pacientes, respectivamente. El nivel de servicio está clasificado en cuatro categorías, según el resultado de las visitas domiciliarias: 1) visitas fallidas con motivo de que el equipo asistencial llegó después del doble del tiempo de espera indicado al paciente (dos veces el *triage*);

2) visita fallida porque el paciente cancela la visita, pues, aunque el equipo asistencial llegó antes del doble del tiempo de espera indicado, la función de espera del paciente le indicó cancelar la visita; 3) visitas exitosas, porque el paciente decide esperar, es decir, la demora no superó el desespero del paciente, y 4) visitas exitosas dentro del tiempo de espera indicado.

Figura 10. Resultados de la experimentación

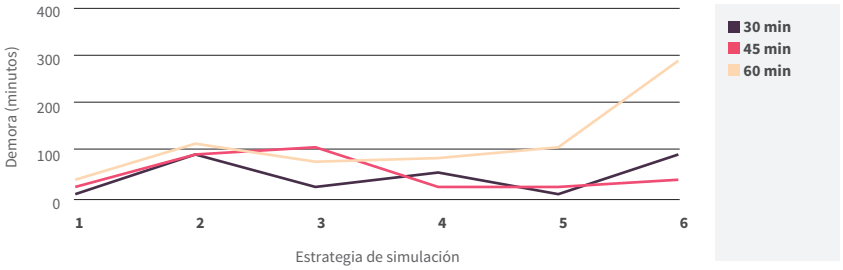


Nota: a) $\alpha = \beta = 0{,}5$ para periodos de 30 minutos; b) 45 minutos; c) 60 minutos; d) costos generales de los periodos de optimización.

Fuente: elaboración propia.

Los resultados mostrados en la figura 10 no incluyen ni los costos ni los indicadores de servicio de las 2 primeras horas de simulación (tiempo de calentamiento). Lo anterior, debido a que el sistema inicia vacío (sin solicitudes) y llega a un estado estable respecto a las demoras para llegar donde los pacientes después de las 2 horas de simulación, como se muestra en la figura 11. Además, la figura 11 permite ver que para una estrategia de optimización cada 60 minutos, luego de 6 horas de simulación, la demora promedio se incrementa drásticamente (hasta 300 minutos), por lo que se puede descartar por incumplir los requerimientos técnicos de operación del servicio, dado que los pacientes no tolerarían esperas de este tipo.

Figura 11. Tiempos de demora en las estrategias de optimización



Fuente: elaboración propia.

Análisis y discusión

Como ya se indicó, las estrategias de optimización están evaluadas con base en los niveles de servicio y los costos en que se incurre al aplicar la estrategia. En la figura 10a se muestra que para un periodo de optimización de 30 minutos (optimizar la decisión de ruteo cada 30 minutos) los indicadores de servicio son los mejores. Es decir, con un 72 % de las visitas (23 pacientes), se llegó en el tiempo de espera indicado; en un 6 % de las visitas (2 pacientes) existió demora, pero los pacientes esperaron, por lo que también se consideran como visitas exitosas. Por otra parte, en un 9 % (3 pacientes) de las visitas existió demora y los pacientes no esperaron. Y en un 13 % (4 pacientes) el tiempo de llegada excedió el doble del tiempo de espera indicado, por lo que fueron visitas fallidas. Por lo tanto, para el periodo de optimización se puede concluir que hubo un 78 % de visitas exitosas (25 pacientes) contra un 12 % (7 pacientes) de visitas fallidas.

Al realizar el mismo análisis para un periodo de optimización de 45 minutos, se halló que un 72 % de las visitas fueron exitosas (23 pacientes) contra un 28 % (9 pacientes) de visitas fallidas. Estas dos estrategias no distan mucho en su nivel de desempeño y se observa en la figura 10d que el costo de la estrategia de optimizar cada 30 minutos es mayor al de optimizar cada 45 minutos. Queda a criterio del proveedor de cuidado a la salud domiciliaria decidir qué estrategia seguir con base en el nivel de servicio deseado y los costos que está dispuesto a asumir. Finalmente, en la figura 10c se observa que con un periodo de optimización de 60 minutos los indicadores de servicio disminuyen de manera considerable, lo que resulta en solo un 56 % (18 pacientes) de visitas exitosas y en un muy elevado costo operacional. Lo anterior se debe a que los pacientes que ingresan al inicio del periodo de optimización tienen que esperar más de 30 minutos a que se les sea asignado un equipo de personal asistencial, lo que, sumado al tiempo de viaje, supera el doble del tiempo de espera indicado por el proveedor. Por lo tanto, que-

da demostrado que un proceso de optimización de asignación de pacientes y rutas más frecuente lleva a mejores indicadores de servicio a un alto costo. Lo ideal es encontrar el balance entre el periodo de optimización y los costos del sistema que, como se observó en este caso, se logró con una optimización de las decisiones cada 45 minutos.

Conclusiones

El cuidado a la salud domiciliaria es una estrategia de prestación del servicio de salud que ha demostrado ser costo-efectiva para atender las necesidades médicas de los pacientes en sus hogares. Esta estrategia evita desplazamientos innecesarios en pacientes que pueden esperar que un médico los visite en sus hogares y disminuye costos en los sistemas de salud, debido a que el solicitante no está usando los recursos con que cuentan las instalaciones médicas. Por otra parte, también se ha evidenciado que la demanda del servicio de cuidado a la salud domiciliaria se ha incrementado en los últimos años, debido al crecimiento y al envejecimiento de la población, así como al aumento de la expectativa de vida de las personas.

Prestar un servicio de calidad y a un bajo costo significa ser competitivo en el cada vez más concurrido mercado del cuidado en casa. Para lograr lo anterior, se debe contar con herramientas acertadas de toma de decisiones en los aspectos fundamentales. Este trabajo presentó una herramienta para determinar las rutas de los equipos de personal asistencial que visitarán a los pacientes de acuerdo con la prioridad en la atención, junto con un modelo de programación matemática, cuyo fin es definir las rutas de los equipos, minimizando sus tiempos de viaje y la demora en el servicio respecto al tiempo de espera indicado, y que se calcula con base en el *triage* del paciente. Por otra parte, también se presentó un modelamiento computacional con el cual simular las llegadas y las cancelaciones de las solicitudes por el servicio, y, de esta manera, evaluar diferentes frecuencias para los periodos de optimización o de realización de los esquemas de ruteo. Mediante un caso de estudio inspirado en un proveedor del servicio de cuidado a la salud domiciliaria, se encontró que los mejores indicadores de atención se dan con una optimización cada 30 minutos, es decir, el mayor número de visitas domiciliarias exitosas se logró recolectando solicitudes de atención domiciliaria y optimizando las rutas cada media hora. Por otra parte, también se encontró que, para este caso particular, con ese periodo de optimización los costos se incrementan y, desde el punto de vista económico, el mejor periodo de optimización es de 45 minutos, sin disminuir considerablemente los indicadores de servicio, comparados con los obtenidos en un periodo de optimización de 30 minutos.

Es importante resaltar que en el caso de estudio se emplearon cuatro equipos de personal asistencial con habilidades preestablecidas. Con la herramienta computacional desarrollada también se modificó el número de equipos asisten-

ciales y sus habilidades, así como el número de pacientes que ingresarán a cada periodo de optimización. La metodología puede llevarse a cabo en otras ciudades y periodos, para determinar el impacto de cambiar las frecuencias de recolección de solicitudes y optimización de las decisiones.

Una limitación importante en este trabajo fue el tamaño del problema. Con el fin de evaluar periodos de optimización amplios, se deben ejecutar modelos con muchos más pacientes, lo que no es posible con la aproximación exacta que se desarrolló en este trabajo. Esto requiere métodos de optimización más sofisticados, como las metaheurísticas. También es de resaltar que este trabajo se planteó con distancias en línea recta entre los hogares de los pacientes, es decir, no consideró la malla vial ni las congestiones frecuentes en algunas zonas de la ciudad.

Trabajos futuros pueden apuntar a incluir sistemas de información geográfica para acercar más la propuesta a la realidad; del mismo modo, también pueden incorporar la posibilidad del ausentismo de equipos de personal asistencial o contemplar los inventarios de los equipos asistenciales, que pueden ser insumos médicos que se van gastando conforme se visitan pacientes. Por otra parte, trabajos futuros también pueden apuntar a implementar metodologías de solución eficientes para abordar problemas de mayor tamaño con más pacientes y equipos asistenciales. Finalmente, se resalta que es posible combinar en otros campos metodologías de simulación basadas en agentes con modelos de optimización para tomar decisiones en la simulación, como la logística humanitaria, en la que, utilizando modelos matemáticos, se pueden tomar decisiones de ruteo de inventario y, con el modelo de simulación, considerar diferentes eventos de la logística humanitaria.

Agradecimientos

Los autores agradecen a la Universidad de La Sabana, por la financiación parcial de este proyecto. Se agradece también a Fair Isaac Corporation, por proveer las licencias de uso académico de Xpress-MP bajo el programa académico Academic Partner Program, suscrito con la Universidad Distrital Francisco José de Caldas.

Referencias

Allaoua, H., Borne, S., Létocart, L. y Wolfler Calvo, R. (2013). A matheuristic approach for solving a home health care problem. *Electronic Notes in Discrete Mathematics*, *41*, 471-478. https://doi.org/10.1016/j.endm.2013.05.127

Bard, J. F., Shao Y., Qi, X. y Jarrah, A. I. (2014). The traveling therapist scheduling problem. *IIE Transactions*, *46*(7), 683-706. https://doi.org/10.1080/0740817X.2013.851434

Bellifemine, F., Bergenti, F., Caire, G. y Poggi, A. (2005). Jade-A Java agent development framework. En *Multi-agent programming: Languages, platforms and applications* (pp. 125-147). Springer. https://doi.org/10.1007/0-387-26350-0_5

Braekers, K., Hartl, R. F., Parragh, S. N. y Tricoire, F. (2016). A bi-objective home care scheduling problem: Analyzing the trade-off between costs and client inconvenience. *European Journal of Operational Research*, *248*(2), 428-443. https://doi.org/10.1016/j.ejor.2015.07.028

Carello, G. y Lanzarone, E. (2014). A cardinality-constrained robust model for the assignment problem in home care services. *European Journal of Operational Research*, *236*(2), 748-762. https://doi.org/10.1016/j.ejor.2014.01.009

Fikar, C. y Hirsch, P. (2017). Home health care routing and scheduling: A review. *Computers & Operations Research*, *77*, 86-95. https://doi.org/10.1016/j.cor.2016.07.019

Genet, N., Boerma, W., Kroneman, M., Hutchinson, A. y Saltman, R. B. (Eds.). (2012). *Home care across Europe: Current structure and future challenges*. World Health Organization. https://www.euro.who.int/en/publications/abstracts/home-care-across-europe.-current-structure-and-future-challenges-2012

Gutiérrez, E. V. y Vidal, C. J. (2013). Home health care logistics management problems: Framework and research perspectives. *International Journal of Industrial Engineering and Management*, *4*(3), 173-182.

Hiermann, G., Prandtstetter, M., Rendl, A., Puchinger, J. y Raidl, G. R. (2015). Metaheuristics for solving a multimodal home-healthcare scheduling problem. *Central European Journal of Operations Research*, *23*(1), 89-113. https://doi.org/10.1007/s10100-013-0305-8

Issaoui, B., Zidi, I., Marcon, E. y Ghedira, K. (2015). New multi-objective approach for the home care service problem based on scheduling algorithms and variable neighborhood descent. *Electronic Notes in Discrete Mathematics*, *47*, 181-188. https://doi.org/10.1016/j.endm.2014.11.024

Koeleman, P. M., Bhulai, S. y Van Meersbergen, M. (2012). Optimal patient and personnel scheduling policies for care-at-home service facilities. *European Journal of Operational Research*, *219*(3), 557-563. https://doi.org/10.1016/j.ejor.2011.10.046

Liu, M., Xu, L. y Su, Q. (2015). Mathematical model for pattern-based medical team planning problem. En *12th International Conference on Service Systems and Service Management (ICSSSM)*, Guangzhou (pp. 1-6). https://doi.org/10.1109/ICSSSM.2015.7170273

López-Santana, E. R., Espejo-Díaz, J. A. y Méndez-Giraldo, G. A. (2016). Multi-agent approach for solving the dynamic home health care routing problem. En *Applied Computer Sciences in Engineering. WEA 2016. Communications in Computer and Information Science* (vol. 657). Springer, Cham. https://doi.org/10.1007/978-3-319-50880-1_17

Martín García, M. (2013). Estudio del triaje en un servicio de urgencias hospitalario. *Revista Enfermería C y L*, *5*(1), 42-49. http://www.revistaenfermeriacyl.com/index.php/revistaenfermeriacyl/article/view/91

Maya Duque, P. A., Castro, M., Sörensen, K. y Goos, P. (2015). Home care service planning. The case of Landelijke Thuiszorg. *European Journal of Operational Research*, *243*(1), 292-301. https://doi.org/10.1016/j.ejor.2014.11.008

Mutingi, M. y Mbohwa, C. (2014). A fuzzy-based particle swarm optimization approach for task assignment in home health care. *South African Journal of Industrial Engineering*, *25*(November), 84-95. https://ieeexplore.ieee.org/document/6962576

Organisation for Economic Cooperation and Development. (2017). *Health at a glance 2017*. OECD. https://doi.org/10.1787/health_glance-2017-en

Rasmussen, M. S., Justesen, T., Dohn, A. y Larsen, J. (2012). The home care crew scheduling problem: Preference-based visit clustering and temporal dependencies. *European Journal of Operational Research*, *219*(3), 598-610. https://doi.org/10.1016/j.ejor.2011.10.048

Redjem, R. y Marcon, E. (2015). Operations management in the home care services: A heuristic for the caregivers' routing problem. *Flexible Services and Manufacturing Journal*, (28), 280-303. https://doi.org/10.1007/s10696-015-9220-8

Redjem, R., Marcon, E. y Xie, X. (2014). Management of the care activities in home health care services: The routing and scheduling of caregivers level. En *IEEE International Conference on Industrial Engineering and Engineering Management*, Bandar Sunway (pp. 712-718). https://doi.org/10.1109/IEEM.2014.7058731

Rest, K.-D. y Hirsch, P. (2016). Daily scheduling of home health care services using time-dependent public transport. *Flexible Services and Manufacturing Journal*, (28), 405-525. https://doi.org/10.1007/s10696-015-9227-1

Rest, K. -D., Trautsamwieser, A. y Hirsch, P. (2012). Trends and risks in home health care. *Journal of Humanitarian Logistics and Supply Chain Management*, *2*(1), 34-53. https://doi.org/10.1108/20426741211225993

Sánchez Cárdenas, M. A. y Fuentes Bermúdez, G. P. (2016). Gestión clínica de programas de cuidado domiciliario. *CES Salud Pública*, *7*(2). https://doi.org/10.21615/cessp.7.2.5

Secretaría Distrital de Movilidad. (2015). *Movilidad en cifras 2015*. https://www.movilidadbogota.gov.co/web/SIMUR/ARCHIVOS/Movilidad_Cifras_2015_V4_marzo2017.pdf

Triki, N., Garaix, T. y Xie, X. (2014). A two-phase approach for periodic home health care planning. En *IEEE International Conference on Automation Science and Engineering (CASE)*, Taipei (pp. 518-523). https://doi.org/10.1109/CoASE.2014.6899375

World Bank. (2018). *Population ages 65 and above*. https://data.worldbank.org/indicator/sp.pop.65up.to

Yalçındağ, S., Cappanera, P., Grazia Scutellà, M., Şahin, E. y Matta, A. (2016). Pattern-based decompositions for human resource planning in home health care services. *Computers & Operations Research*, *73*, 12-26. https://doi.org/10.1016/j.cor.2016.02.011

Yuan, B., Liu, R. y Jiang, Z. (2014). Home health care crew scheduling and routing problem with stochastic service times. En *IEEE International Conference on Automation Science and Engineering (CASE)*, Taipei (pp. 564-569). https://doi.org/10.1109/CoASE.2014.6899382

Yuan, Z. y Fügenschuh, A. (2015). Home health care scheduling: A case study. *AMOS*, (22), 1-18. https://epub.sub.uni-hamburg.de/epub/volltexte/2020/100615/pdf/828597456.pdf

Anexo. Georreferenciación de los pacientes de la experimentación

ID	Latitud	Longitud	ID	Latitud	Longitud
1	4,74459537	−74,0548	25	4,6905	−74,0626
2	4,63986682	−74,139	26	4,6448	−74,1266
3	4,6944407	−74,106	27	4,7156	−74,0306
4	4,70566425	−74,0953	28	4,7543	−74,0431
5	4,66158575	−74,0492	29	4,6579	−74,1286
6	4,71721353	−74,0418	30	4,6732	−74,0876
7	4,60686319	−74,1952	31	4,6709	−74,131
8	4,67231817	−74,1068	32	4,6524	−74,111
9	4,6869224	−74,0368	33	4,614	−74,1332
10	4,69657758	−74,043	34	4,6482	−74,1612
11	4,66721309	−74,1284	35	4,6983	−74,1125
12	4,66035655	−74,0635	36	4,7083	−74,1057
13	4,57439499	−74,1369	37	4,7084	−74,1146
14	4,59043899	−74,1277	38	4,7008	−74,1179
15	4,66804827	−74,1092	39	4,6892	−74,1005
16	4,65270777	−74,138	40	4,6852	−74,0548
17	4,71531367	−74,0398	41	4,6611	−74,1293
18	4,62576026	−74,0866	42	4,7462	−74,0338
19	4,65803219	−74,1308	43	4,6498	−74,1089
20	4,66827921	−74,1573	44	4,6557	−74,1567
21	4,73258921	−74,085	45	4,5598	−74,1228
22	4,6848137	−74,0696	46	4,6797	−74,1118
23	4,67016586	−74,1241	47	4,6394	−74,0756
24	4,7255073	−74,0693	48	4,739	−74,0658

Fuente: elaboración propia.

Capítulo 4.
Mejora de los tiempos de espera en la sala de urgencias de un hospital público, por medio de la simulación de eventos discretos

Álvaro Junior Caidedo-Rolón,[1]
Deisy Paola Leal Hernández,[2]
Liceth Paola Ortiz López[3] *y*
Leonardo Rivera Cadavid[4]

[1] Departamento de Procesos Industriales
Universidad Francisco de Paula Santander
Cúcuta, Colombia
alvarojuniorcr@ufps.edu.co

[2] Programa de Ingeniería Industrial
Universidad Francisco de Paula Santander
Cúcuta, Colombia
deisypaolalh@ufps.edu.co

[3] Programa de Ingeniería Industrial
Universidad Francisco de Paula Santander
Cúcuta, Colombia
licethpaolaol@ufps.edu.co

[4] Escuela de Ingeniería Industrial
Universidad del Valle
Cali, Colombia
leonardo.rivera.c@correounivalle.edu.co

Resumen

La sobreocupación en las salas de urgencias es un problema que afecta tanto el desempeño del personal médico y asistencial como la calidad y oportunidad en la prestación de los servicios de salud. La investigación propuso mejoras en la prestación de los servicios de urgencias de un hospital público que reduzcan el tiempo promedio de espera para la consulta médica y la duración promedio de la estancia de los pacientes, por medio de un modelo de simulación de eventos discretos. Eliminar el proceso de preadmisiones permitirá: determinar el indicador de pacientes que salen del sistema sin ser vistos por un médico; reducir el tiempo promedio de espera para la consulta médica y la duración de la estancia en cerca de 25 y 30 minutos, respectivamente; disminuir la ocupación de la auxiliar de enfermería, y aumentar el egreso de pacientes en promedio un 34,47 %. Adicionar un médico para la consulta general reduce el tiempo promedio de espera para recibir la consulta en 10,65 minutos y la duración de la estancia en 20,32 minutos. El modelo de simulación se constituye en una herramienta flexible y valiosa, a fin de apoyar la toma de decisiones estratégicas, tácticas y operacionales por parte de los directivos del hospital.

Palabras clave: duración de la estancia, gestión de salud, investigación de operaciones, logística hospitalaria, sobreocupación

Introducción

Las salas de urgencias son una parte crítica del sistema de salud, con una alta complejidad por sus diferentes flujos multiservicio para diversas categorías de pacientes, y son una de las áreas más importantes en la búsqueda por cumplir la función misional de las instituciones prestadoras de servicios de salud (IPS) (Ajami et al., 2012; Guo et al., 2017; Tsai et al., 2010; Velásquez-Restrepo et al., 2011). Esta salas desempeñan un papel vital en la atención primaria a las llegadas no programadas de pacientes y son reconocidas por la contribución que le proporcionan a la sociedad, por su gran responsabilidad en proveer un servicio de cuidado crítico a pacientes con lesiones accidentales o enfermedades diversas, de una manera rápida (Arunachalam et al., 2017; Huang y Chu, 2011), y por reducir la tasa de mortalidad y morbilidad (Aringhieri et al., 2016).

Durante el 2015, las salas de urgencias de los Estados Unidos fueron visitadas por cerca de 137 millones de personas, una cifra equivalente al 43,3 % de la población total (Rui y Kang, 2015). En Taiwán, el número total de personas que asistieron a estos servicios aumentó un 34 %, entre 1998 y 2011 (Chen y Wang, 2016). Se estima que más de 46 millones de colombianos se encuentran asegurados a servicios de salud (Ministerio de Salud y Protección Social [MSPS], 2015a). En el 2014, el 19,38 % de los afiliados al régimen subsidiado y el 33,06 % de los afiliados al régimen contributivo usaron los servicios de urgencias y se registraron, en promedio, cerca de 19 millones de pacientes cada año en el periodo 2010-2014 (Asociación Colombiana de Empresas de Medicina Integral, 2017).

Respecto a la oferta de servicios de urgencias de baja, mediana y alta complejidad, en Colombia, en el 2016, se presentó una oferta de 1115 IPS públicas habilitadas y 503 IPS privadas, del total de 10 366 IPS disponibles en el país. La razón de camas por cada 1000 habitantes es de 1,61 camas, y teniendo en cuenta que la norma internacional es de 2,5-4 camas por cada 1000 habitantes, según la Organización Mundial de la Salud, Colombia no alcanza dicho parámetro (Rozo Lesmes y Acosta Hernández, 2016).

Las salas de urgencias —principal punto de entrada de pacientes en los hospitales modernos— han presentado una frecuente sobreocupación, por lo que los gerentes de los hospitales y los líderes de los departamentos de urgencias prestan cada vez más atención a ello, con el objeto de mejorar la eficiencia operativa y brindar un mejor servicio (Afilal et al., 2016; Sinreich et al., 2012). La sobreocupación es un problema de salud pública (Zhao et al., 2015) que afecta el desempeño del personal médico (Sinreich et al., 2012) y de la salas de urgencias (Chan et al., 2012; Ghanes et al., 2016), el acceso a la atención médica oportuna

y la calidad del servicio médico de urgencias (Azadeh et al., 2013; Khalifa, 2015; Khalifa y Zabani, 2016; Zhang et al., 2013).

Con frecuencia, hay sobreocupación en las salas de urgencias, debido a la poca eficiencia de la colocación de los recursos o del flujo deficiente del paciente, debido a operaciones ineficientes en su interior (Bedoya-Valencia y Kirac, 2016; Tsai et al., 2010; Zhang et al., 2013; Zhao et al., 2015); a pacientes no considerados urgencias (Arunachalam et al., 2017; Hoot y Aronsky, 2008); a recursos limitados para la creciente y variable demanda de atención (Ashour y Okudan Kremer, 2016; Azadeh et al., 2013; Cocke et al., 2016; Daldoul et al., 2016; Zhao et al., 2015); al crecimiento y envejecimiento de la población (Kuo et al., 2016; Luscombe y Kozan, 2016); al bloqueo de acceso desde el servicio de urgencias hacia la hospitalización, por la escasez de camas (Abo-Hamad y Arisha, 2012; Hoot y Aronsky, 2008; Lee, 2014); a la falta de coordinación entre los departamentos de los hospitales, que produce una utilización desequilibrada de sus recursos y afecta todo el proceso de atención (Abo-Hamad y Arisha, 2012), o a pacientes con reingreso y a la temporada de influenza (Hoot y Aronsky, 2008).

El tiempo de espera es uno de los factores importantes en la gestión de la atención de la salud, que determina la capacidad de respuesta y, por ende, el desempeño del servicio de salud (Ajami et al., 2012; Ashour y Okudan Kremer, 2016) y el rendimiento de las salas de urgencias (Luscombe y Kozan, 2016). Además, ha sido uno de los principales criterios considerados por los investigadores que impacta directamente en la calidad y la oportunidad del servicio (Sepehri et al., 2015; Velásquez-Restrepo et al., 2011). En general, los sistemas de salud —y en particular las salas de urgencias alrededor del mundo— están enfrentando enormes desafíos para cumplir los objetivos cada vez más conflictivos de proporcionar una amplia accesibilidad y eficiencia, a la vez que deben brindar servicios rápidos y de alta calidad (Abo-Hamad y Arisha, 2012). Las operaciones inefectivas pueden llevar a graves consecuencias, como la demora de tratamientos o incluso la muerte de pacientes críticos (Kuo et al., 2016; Sepehri et al., 2015).

En las salas de urgencias de Irán, la demora en la atención no solo aumenta la insatisfacción de los servicios de salud, sino que también retrasa la admisión de nuevos pacientes, y ello incrementa la mortalidad de los accidentes (Ajami et al., 2012; Sepehri et al., 2015). Los largos tiempos de espera son un motivo para que los pacientes abandonen la sala de urgencias sin ser atendidos y muchos de ellos buscarán otro tipo de atención médica (Khalifa, 2015). Acceder a una sala de urgencias con un tiempo de espera más corto puede significar la vida o la muerte para alguien víctima de un accidente cerebrovascular e infarto de miocardio, por ejemplo (Househ y Yunus, 2014).

En los Estados Unidos, de los cerca de 137 millones de personas que visitaron las salas de urgencias en el 2015, el 35,4 % tuvo que esperar menos de 15 minutos para recibir atención médica y el 32,1 % entre 15 y 59 minutos (Rai y Kang,

2015). En Colombia, la meta establecida para el 2018 para la persona clasificada como *triage* II (aquel cuya condición clínica puede evolucionar hacia un rápido deterioro o a su muerte) fue de 20 minutos (MSPS, 2015a), y la Resolución 5596 del 2015 lo reglamentó en 30 minutos; sin embargo, durante los periodos 2017-II, 2017-I, 2016-II y 2016-I, los resultados fueron de 94,36, 90,3, 111,02 y 119,01 minutos, respectivamente (Observatorio Nacional de Calidad en Salud, 2019). En el ámbito internacional estos tiempos deben ser: Hong Kong y Malasia, 15 minutos; Australia, 10 minutos. Esto permite evidenciar el bajo nivel de calidad del servicio de urgencias en el país, el cual fue calificado por los usuarios con un nivel de satisfacción cercano al 65 %, durante 2014, 2016 y 2017, siendo el peor servicio de salud (MSPS, 2015b, 2017).

La simulación digital es una técnica que imita en un ordenador el comportamiento de un sistema real o hipotético, según ciertas condiciones particulares de operación. Los sistemas orientados a eventos discretos se caracterizan por que las propiedades de interés del sistema cambian únicamente en una secuencia de instantes de tiempo, y se considera que permanecen constantes el resto de tiempo. La secuencia de instantes en los cuales el sistema puede presentar un cambio obedece a un patrón aleatorio (Guasch et al., 2005).

Las siguientes investigaciones han utilizado la simulación de eventos discretos en la mejora de la prestación de los servicios de urgencias: Bedoya-Valencia y Kirac (2016) analizaron la asignación de recursos en pro de reducir la estancia de los pacientes y el tiempo antes de ser vistos por un médico mientras se nivela la utilización de recursos; Masmoudi et al. (2014) desarrollaron un modelo de simulación que evalúa el impacto del cambio del número de médicos en las áreas de traumatología y medicina en la duración de la estancia del paciente en un hospital en Francia; Ajami et al. (2012) propusieron escenarios en un modelo de simulación para reducir el tiempo de espera en un hospital en Irán, y Weng et al. (2011) crearon un modelo de simulación para encontrar la asignación óptima de recursos en la sala de urgencias que aumente su rendimiento, el número de pacientes tratados y el grado de satisfacción.

Las siguientes investigaciones han utilizado la simulación de eventos discretos y la optimización en la mejora de la prestación de los servicios de urgencias: Ibrahim et al. (2017) construyeron un modelo de optimización-simulación capaz de identificar los cuellos de botella y de utilizar la función OptQuest de Arena para obtener la posible combinación de recursos que reduzca el tiempo medio de espera de los pacientes y aumente el número de atendidos en un hospital público de Malasia; Kuo et al. (2016) propusieron un enfoque de simulación-optimización (integración de la simulación con metaheurística), para prever el impacto en los tiempos de espera de los pacientes y la utilización de los médicos, al ajustar los niveles de personal o los tiempos de los turnos en un hospital en Hong Kong; Guo y Tang (2015) integraron las técnicas de simulación y optimización

para diseñar la configuración óptima de recursos, con el objetivo de minimizar la duración de la estancia de los pacientes y los costos generales, y Sinreich et al. (2012) desarrollaron dos algoritmos heurísticos iterativos que combinan modelos de simulación y optimización lineal para la programación de turnos de trabajo del personal (médicos, enfermeras y radiólogos), a fin de reducir el tiempo de espera de los pacientes y nivelar la utilización de los recursos.

El hospital público objeto de estudio es una entidad de referencia en un departamento fronterizo de Colombia, clasificado en cuarto grado de complejidad, que imparte servicios ambulatorios de urgencias con personal especializado en todas las áreas, para toda la población del departamento y su área de influencia. El hospital cuenta con dos salas de urgencias (pediatría y adultos) que están disponibles las 24 horas del día, los 7 días de la semana. En el 2017, el hospital atendió a 39 714 pacientes en el servicio de urgencias de adultos, de los cuales el 7,34 % fue población inmigrante venezolana.

El área de urgencias del hospital caso de estudio presenta una sobreocupación en sus instalaciones que ha alcanzado entre el 130 % y 165 % (Leal y Ortiz, 2018), lo cual genera insatisfacción en el usuario, por los largos tiempos de espera (en promedio 78,11 minutos) y por la duración de la estancia (mínima de 148 minutos, promedio de 3,71 días y máxima de 10,46 días); además de la sobrecarga laboral que debe experimentar el personal de urgencias. Esta problemática se debe a varios factores, como el desarrollo de operaciones ineficientes en el interior de la sala de urgencias; los recursos físicos y humanos limitados para la creciente demanda, con un incremento de los inmigrantes venezolanos; la falta de disponibilidad de camas en hospitalización; entidades promotoras de salud (EPS) sin convenio vigente con IPS o red de prestadores de servicios de salud, que dejan a los usuarios sin servicios de salud, quienes deben acudir al hospital público; la ineficiencia en la red de atención primaria pública de urgencias, y las limitaciones presupuestales para la ampliación de la capacidad.

Por lo anterior, el propósito de la investigación fue proponer mejoras en la prestación de los servicios de urgencias en un hospital público ubicado en una ciudad fronteriza de Colombia que reduzcan el tiempo promedio de espera para la consulta médica y la duración de la estancia de los pacientes y que mejoren la calidad y la oportunidad en los servicios de salud. El capítulo presenta las siguientes secciones: la primera expone los métodos y herramientas empleados en la investigación; la segunda aborda la aplicación de la metodología de la simulación junto con los resultados obtenidos con el modelo real y el modelo simulado con dos escenarios de mejora; la tercera plantea una discusión con otras investigaciones, y, finalmente, la cuarta presenta las conclusiones de la investigación y trabajos futuros.

Métodos

La investigación utilizó la técnica de la simulación de eventos discretos para estudiar y plantear mejoras en el caso de estudio, por las siguientes razones: 1) la pertinencia metodológica; 2) la revisión de antecedentes, que muestra que los autores la han aplicado con resultados favorables para tratar el mismo problema de largos tiempos de espera y estancia; 3) la disponibilidad de *software* gratuito con capacidad de resolución de problemas de tamaño medio, y 4) la practicidad, sencillez y rapidez a la hora de plantear escenarios de mejora. Este apartado involucra las etapas de un proyecto de simulación consideradas en la investigación y el respectivo análisis estadístico aplicado, a efectos de determinar la confiabilidad de los datos y la validez de la simulación.

Etapas de un proyecto de simulación

La investigación aplicó las etapas de un proyecto de simulación propuestas por Guasch et al. (2005), que se presentan en la tabla 1.

Tabla 1. Etapas de un proyecto de simulación

Etapa	Descripción
Formulación del problema	Define el problema que se pretende estudiar e incluye los objetivos por escrito.
Diseño del modelo conceptual	Especificación del modelo a partir de las características de los elementos del sistema que se quiere estudiar y sus interacciones teniendo en cuenta los objetivos del problema.
Recogida de datos	Identificar, recoger y analizar los datos necesarios para el estudio.
Construcción del modelo	Construcción del modelo de simulación partiendo del modelo conceptual y de los datos.
Verificación y validación	Comprobar que el modelo se comporta como es de esperar y que existe la correspondencia adecuada entre el sistema real y el modelo.
Análisis	Analizar resultados de la simulación con la finalidad de detectar problemas y recomendar mejoras o soluciones.
Documentación	Proporcionar documentación sobre el trabajo efectuado.
Implementación	Poner en práctica las decisiones efectuadas con el apoyo del estudio de simulación.

Fuente: Guasch et al. (2005, p. 358).

Análisis estadístico

En la etapa de recolección de datos (tabla 1) se recurrió al departamento de estadística de la institución de salud para obtener registros históricos de los tiempos de los procesos; sin embargo, parte de estos no estaban registrados, por lo que fue necesario tomar los tiempos, empleando el cronómetro y el método de regresos a ceros. A fin de determinar la confiabilidad de los datos recolectados e identificados de cada una de las etapas del proceso de atención de la sala de urgencias, se determinó el intervalo de confianza y el error típico de los datos o precisión de estimación (e), como se aprecia en las ecuaciones (1) y (2), respectivamente (Mathur y Solow, 1996).

$$x \pm e \tag{1}$$

Siendo e,

$$e = t_{(n\text{-}1,\frac{\alpha}{2})} * \frac{s}{\sqrt{n}} \tag{2}$$

A efectos de aumentar la confiabilidad de los datos, se disminuyó su error típico, y con ello se redujo el intervalo de confianza. Con el nuevo valor de (e*) se determinó el número de observaciones necesarias *(n*)* para cada proceso y el número de corridas requeridas para la simulación, mediante la ecuación (3) (Mathur y Solow, 1996).

$$n^* = \left(\frac{t_{((n^*\text{-}1),\frac{\alpha}{2})} * s}{e}\right) \tag{3}$$

Para la simulación se empleó el *software* Arena 14.5®, y para las pruebas de bondad de ajuste, la herramienta Input Analyzer de este *software*, el cual arrojó las distribuciones de probabilidad que representan a cada proceso de atención en la sala de urgencias.

Con el fin de validar el modelo de simulación se utilizó el procedimiento estadístico de prueba de hipótesis, prueba *t* de Student, definiendo una hipótesis nula y una alternativa y calculando el estadístico *t* como se expresa en la ecuación (4) (Banks y Carson, 1984).

$$t_0 = \frac{d - \mu_d}{S_d / \sqrt{n}} \tag{4}$$

Donde:

$t_{(n\text{-}1,\frac{\alpha}{2})}$ = valor de la distribución *t* con $(n-1)$ grados de libertad, de modo que el área que está a la derecha de este valor es $(\alpha/2)$.

α = nivel de significancia

n = tamaño de la muestra

s = desviación estándar de la muestra
$\bar{x}$ = promedio de la muestra
t_o = estadístico t
d = promedio de la diferencia de las muestras
μ_d = diferencia de medias de las muestras
s_d = desviación estándar de la diferencia de las muestras

Entonces, determinar el número de médicos requeridos para el proceso de consulta en el periodo de mayor congestión del sistema de urgencias, conllevó aplicar la teoría de colas bajo el sistema de colas (*M/M/c*), por medio del *software* Winqsb version 2.0, cuyo dato se utilizó como parámetro de entrada para el modelo de simulación, en el escenario 2.

Resultados

En esta sección se detallan los resultados de la investigación, una vez se aplicaron los métodos y las herramientas presentados en el apartado anterior.

Etapas de un proyecto de simulación

A continuación, se presentan los resultados de la aplicación de cada una de las etapas de un proyecto de simulación en el caso de estudio, que condujo a las propuestas de mejora en el proceso de atención de los servicios de urgencias del hospital.

Diseño del modelo conceptual

En esta etapa se describen las características de operación del sistema de estudio, los recursos disponibles y los procesos e interacciones entre los actores involucrados. En la tabla 2 se muestran cada una de las áreas de la sala de urgencias con el tipo y cantidad de personal y los recursos físicos disponibles para cada turno.

Tabla 2. Personal y recursos físicos disponibles en la sala de urgencias

Área	Personal	Personal por turno			Recurso	Cantidad
		1	2	3		
Preadmisiones	Auxiliar de enfermería	1	1	1	Sala de espera	1
	Vigilante	1	1	1		
Triage	Enfermera jefe	1	1	1	Consultorio	1
	Auxiliar de enfermería	1	1	1	Camilla	1
Admisiones	Auxiliar administrativo	1	1	1	Computador	1
Consulta médica	Médico general	3	3	2	Consultorios	4
					Camilla	3
					Sala de espera	1
Observación	Médico general	1	1	1	Salas	9
	Auxiliar de enfermería	16	16	16	Camillas	75
	Jefe de enfermería	4	4	4		
Laboratorio	Bacteriólogos	5	5	2	Equipos	2
Sala de procedimientos	Médico general	1	1	1	Camillas	5
	Auxiliar de enfermería	2	2	2		
UCIM	Auxiliar de enfermería	2	2	2	Camillas	7
Facturación-egreso	Auxiliar administrativo	2	2	2	N/A	N/A

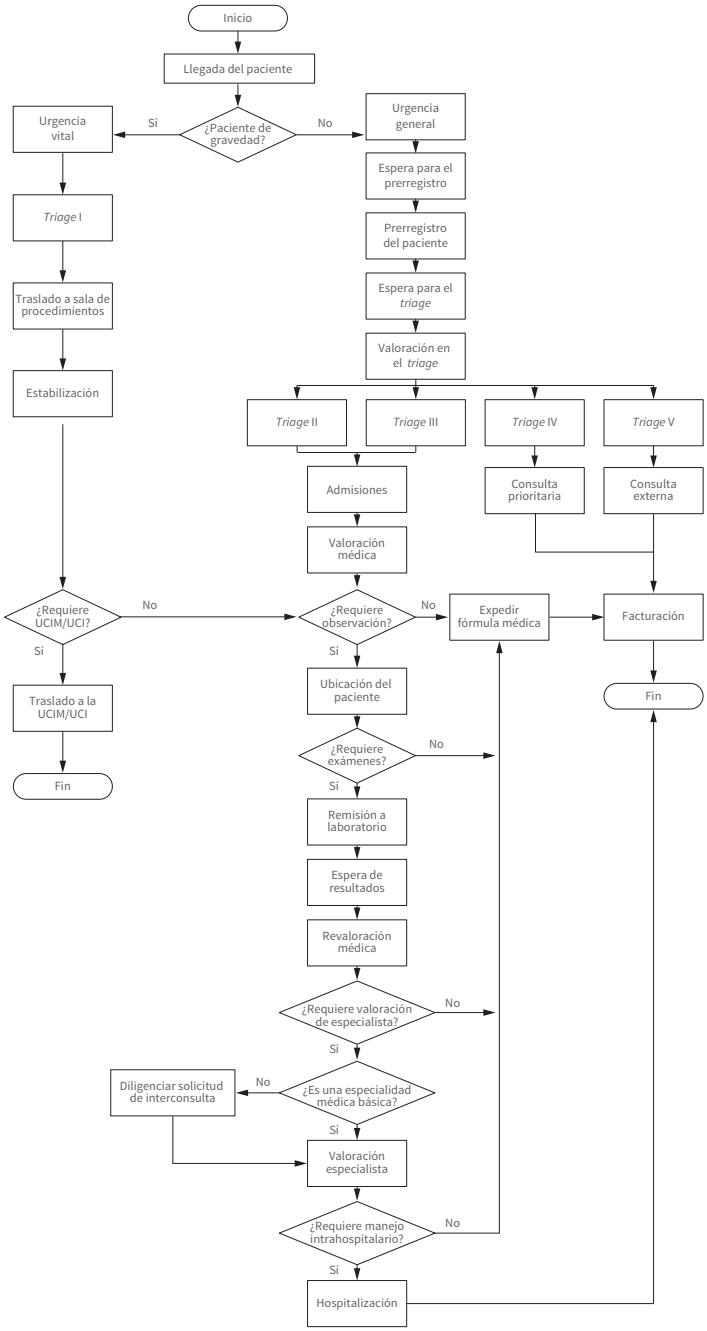
Fuente: elaboración propia.

La sala de urgencias labora en tres turnos: el de la mañana y el de la tarde, de seis horas cada uno, y el de la noche, de doce horas. El proceso desarrollado allí se realiza en diez etapas: 1) *preadmisión*, en la cual la auxiliar de enfermería o el vigilante toman los datos personales del paciente y este espera a ser atendido; 2) *triage*, en la cual se hace una valoración inicial al paciente, quien de acuerdo con sus síntomas se clasifica en una escala del *triage* para determinar la prioridad de

su atención; 3) *admisión*, en la cual el paciente procede a registrarse en el sistema y es identificado con una manilla de acuerdo con su prioridad; 4) *consulta médica*, en la cual el médico general valora al paciente, quien podría requerir exámenes de laboratorio, rayos X, entre otros; 5) *observación*, en la cual se ubican los pacientes para monitorear la evolución de su condición médica. Algunos se ubican en camillas en los pasillos, debido a que en observación no hay disponibilidad de espacio; 6) *exámenes de laboratorio*, en la cual se toman las muestras de sangre y se envían al laboratorio; 7) *revaloración*, en la cual, según los resultados de los exámenes, el médico determina el diagnóstico y el procedimiento, ya sea la hospitalización o la salida; 8) *sala de procedimientos (quirófano)*, en la cual se ubica el espacio dedicado a la atención de las personas clasificadas en el *triage* I, quienes requieren reanimación o pequeñas cirugías; 9) *unidad de cuidados intermedios (UCIM)*, en la cual se ubica a la persona en un espacio especializado para atender a pacientes críticos que luego son trasladados a la unidad de cuidados intensivos (UCI) del hospital, y 10) *facturación*, en la cual, según la historia clínica del paciente, se procede a facturar los servicios prestados. No se consideró la interconsulta con especialistas, debido a que no se disponía de registros históricos de los tiempos de atención.

En la figura 1 se esquematiza el proceso de atención de pacientes por prioridad, donde el *triage* II tiene prioridad en la atención sobre el *triage* III.

Figura 1. Proceso de atención del servicio de urgencias



Fuente: elaboración propia.

Teniendo en cuenta la información estadística suministrada por el hospital, y de acuerdo con la nomenclatura médica, se determinaron los porcentajes de flujo interno de pacientes (posterior al ingreso y la valoración médica inicial), que son: evaluación HC, que hace referencia a los pacientes revalorados (73,43 %); epicrisis, que se emplea para denominar a las personas que van egresando del sistema (19,38 %); procedimientos quirúrgicos-cirugía (0,53 %), y consulta externa para los pacientes clasificados como *triage* IV (2,48 %). De los registros históricos se identificaron los tipos de especialidades médicas ofertadas por la institución y de las cuales hacen uso los pacientes: medicina general (91,569 %), ortopedia y traumatología (2,239 %), medicina interna (2,173 %), cirugía de mano (0,942 %), cirugía general (0,881 %), neurocirugía (0,462 %), urología (0,284 %), cirugía de tórax (0,247 %), trabajo social (0,245 %) y toxicología (0,220 %). De acuerdo con los datos históricos, el 5 % de los pacientes que ingresan a urgencias son clasificados en el *triage* I.

Recolección de datos

En esta etapa se recurrió al departamento de estadística de la institución para obtener registros históricos de los tiempos de los procesos; sin embargo, parte de estos no estaban registrados, por lo que fue preciso tomar los tiempos, empleando el cronómetro y el método de regresos a ceros.

Partiendo de la recolección e identificación de 25 datos de los tiempos de cada proceso (n), en la tabla 3 se determinó la desviación estándar (s), la media ($\bar{x}$) y el error típico de los datos o precisión de estimación (e), en minutos, considerando un nivel de significancia del 5 % y la tabla de distribución t de Student. Para aumentar la confiabilidad de los datos, se disminuyó su error típico, reduciendo de esta manera el intervalo de confianza. Con el nuevo valor de (e^*) se determinó el número de observaciones necesarias (n^*) para cada estación, que se aprecia en la tabla 4.

Tabla 3. Datos de muestras preliminares

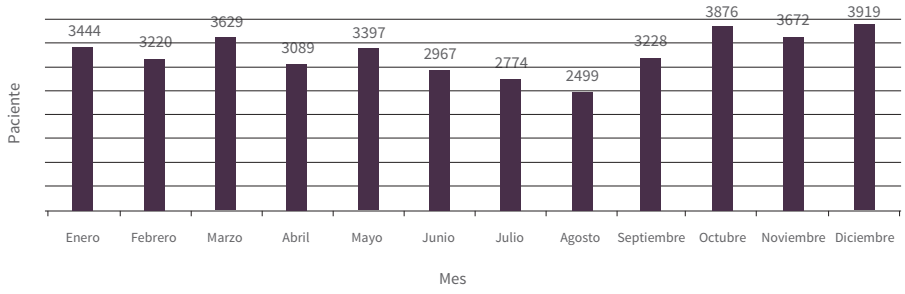
n.°	Estación	*n*	$\bar{x}$	*s*	*e*
1	Tiempo entre llegadas al sistema	25	3,75	3,52	1,45
2	Tiempo de espera para *triage*	25	27,08	15,74	6,50
3	Tiempo entre llegadas para *triage*	25	4,25	1,63	0,67
4	Tiempo de consulta en el *triage*	25	10,27	4,66	1,92
5	Tiempo entre llegadas a admisión	25	5,15	2,87	1,19
6	Tiempo de admisión	25	5,43	1,96	0,19
7	Tiempo en la sala de procedimientos	25	31,45	26,45	0,22
8	Tiempo en la UCIM	25	480,76	179,12	98,14
9	Tiempo de espera para registro	25	20,18	6,84	2,82
10	Tiempo entre llegadas a registro-consulta	25	8,25	7,27	3,00
11	Tiempo de consulta	25	18,43	11,78	4,86
12	Tiempo de espera para asignar cama	25	1656,75	3319,73	1370,38
13	Tiempo en observación	25	3421,00	5167,50	2133,14
14	Tiempo entre llegadas a laboratorio	25	12,95	12,66	5,23
15	Tiempo entre llegadas a consulta-revaloración	25	14,90	14,89	6,15
16	Tiempo de consulta revaloración	25	12,10	9,25	3,82
17	Tiempo entre llegadas a salida	25	19,15	14,90	6,15
18	Tiempo de facturación	25	7,50	3,46	2,20

Fuente: elaboración propia.

Se analizaron los registros históricos de la demanda de los pacientes en la sala de urgencias durante el 2017, cuyo comportamiento se registró en las figuras de la 2 a la 6. En la figura 2 se aprecia que la mayor demanda se presentó al final del año, en diciembre, debido al incremento de enfermedades respiratorias por la época invernal. En la figura 3 se presentan las personas atendidas en promedio por día, durante cada mes del 2017. En la figura 4 se muestra que los días lunes, martes y miércoles son los de mayor demanda; mientras que los fines de semana esta disminuye. En la figura 5 se observa que la prioridad de atención de mayor demanda fue el *triage* III y, en menor proporción, el *triage* I. Y en la figura 6 se evidencia la demanda de pacientes por hora del día, donde esta empieza a crecer a partir de las 6:00 a. m. con el pico más alto de 10:00 a 11:00 a. m., teniendo de

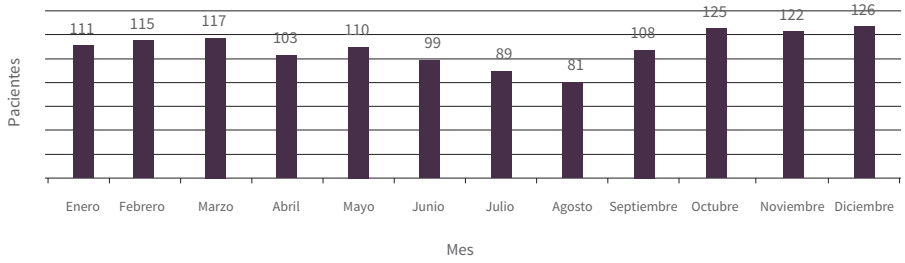
nuevo una ligera alza a las 3:00 p. m. Con esta figura se pudo establecer un patrón de demanda, en el cual se identificó que en el rango horario de 8:00 a. m. a 12:00 m. se presentó la mayor demanda de pacientes, con un total de 14 733, y el que tiene menos registro es el rango de la medianoche hasta las 4:00 a. m., por lo que el turno de la mañana de 7:00 a. m. a 1:00 p. m. es el que presenta mayor congestión.

Figura 2. Demanda de pacientes por mes



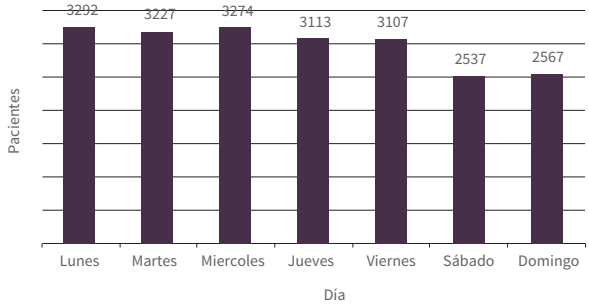
Fuente: elaboración propia.

Figura 3. Demanda diaria promedio por mes



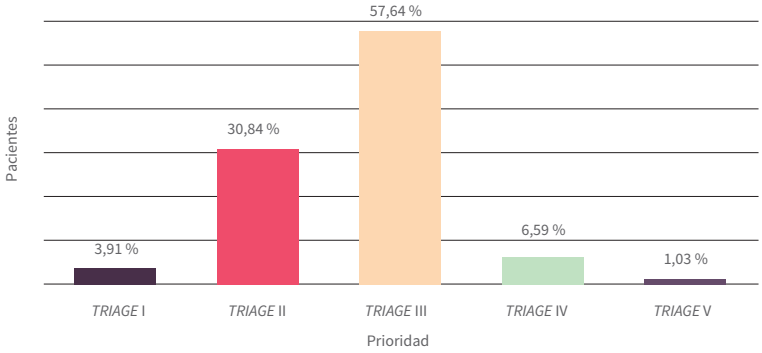
Fuente: elaboración propia.

Figura 4. Demanda por día de la semana



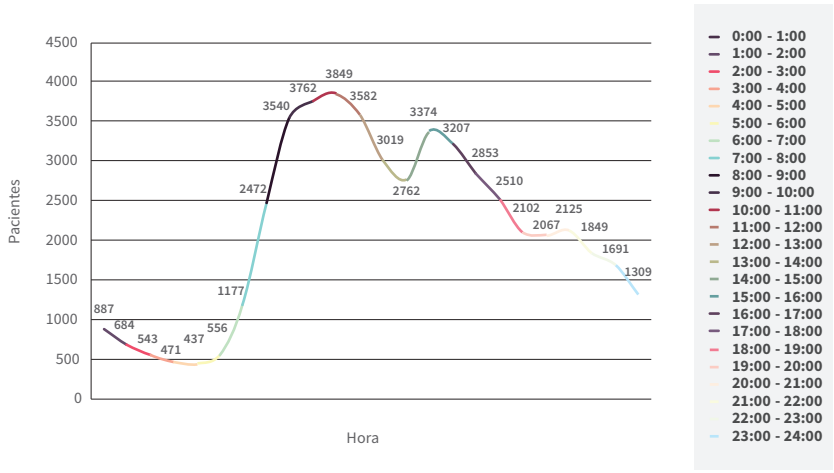
Fuente: elaboración propia.

Figura 5. Demanda por clasificación del *triage*



Fuente: elaboración propia.

Figura 6. Demanda por hora del día



Fuente: elaboración propia.

Construcción del modelo

Con los datos identificados y recolectados de la etapa anterior, se realizaron las pruebas de bondad de ajuste por medio de la herramienta Input Analyzer del *software* Arena, la cual arrojó las distribuciones de probabilidad que representan cada proceso de atención en la sala de urgencias, que se muestran en la tabla 4.

Tabla 4. Distribuciones de probabilidad por cada etapa del proceso

n.°	Estación	e*	*n**	$\bar{x}$	*s*	Tipo de distribución
1	Tiempo entre llegadas al sistema	0,93	62	4,46	4,08	EXPO(4,46)
2	Tiempo de espera para *triage*	4,77	47	26,20	15,20	1 + 53 * BETA(1,09, 1,15)
3	Tiempo entre llegadas para *triage*	0,55	38	4,15	1,51	1,03 + LOGN(3,13, 1,6)
4	Tiempo de consulta en el *triage*	1,48	42	13,30	7,20	1 + GAMM(4,84, 2,54)
5	Tiempo entre llegadas a admisión	0,84	50	4,81	2,12	NORM(4,81, 2,1)

n.°	Estación	e*	n*	$\bar{x}$	s	Tipo de distribución
6	Tiempo de admisión	0,15	45	6,38	2,73	TRIA(3, 7, 15)
7	Tiempo de sala de procedimientos	0,15	42	37,50	32,60	TRIA(5, 30, 120)
8	Tiempo en la UCIM	103,95	48	540,60	197,80	TRIA(240, 660, 1048)
9	Tiempo de espera para registro	2,21	41	22,80	13,00	3 + WEIB(21,9, 1,52)
10	Tiempo entre llegadas a registro-consulta	1,84	67	9,19	8,80	EXPO(9,19)
11	Tiempo de consulta	3,47	49	17,70	8,48	6 + 34 * BETA(0,911, 1,73)
12	Tiempo de espera para asignar cama	736,55	87	918,00	979,00	-28 + GAMM(807, 1,17)
13	Tiempo en observación	1245,24	74	3610,00	5230,00	1 + WEIB(2,07e+003, 0,502)
14	Tiempo entre llegadas a laboratorio	3,62	53	15,50	15,40	GAMM(20,6, 0,756)
15	Tiempo entre llegadas a consulta-revaloración	4,24	53	14,10	13,70	1 + GAMM(15,4, 0,85)
16	Tiempo de consulta revaloración	3,03	40	11,30	4,35	5 + LOGN(6,86, 7,61)
17	Tiempo entre llegadas a salida	4,69	44	20,20	14,90	GAMM(11,3, 1,79)
18	Tiempo de facturación	2,95	45	8,21	3,76	TRIA(4, 7, 10)

Fuente: elaboración propia.

Con la información de la tabla 4, se procedió a crear el modelo de simulación en el *software* Arena 14.5®. Inicialmente, se realizaron 15 corridas de la simulación con un horizonte de tiempo de 18 horas, que representa el turno de la mañana de 7:00 a. m. a 1:00 p. m. de los días críticos de la semana (lunes, martes y miércoles). Después, para la variable "número de pacientes que llegaron al sistema", se calculó la media ($\bar{x}$), la desviación estándar (s), la precisión de estimación (e) y los intervalos de confianza, como se observa en la tabla 5, considerando un nivel de significancia del 5 % y la tabla de distribución t de Student. Para generar

datos confiables se ajustó la precisión de estimación (e^*) a un valor de 5, y se determinó que el número de corridas de la simulación (n^*) requeridas era de 20.

Tabla 5. Determinación del número de corridas de la simulación

Número de corridas	Pacientes que llegaron al sistema simulado
1	107
2	113
3	128
4	110
5	141
6	112
7	113
8	128
9	115
10	141
11	122
12	114
13	118
14	124
15	117
$\bar{x}$	*120,2*
s	*10,47*
e	*5,798*
Límite superior	*125,99*
Límite inferior	*114,40*
e^*	*5*
n^*	*20,17*

Fuente: elaboración propia.

Verificación y validación

Para llevar a cabo la validación del modelo se aplicó el procedimiento estadístico de la prueba de hipótesis *t*, mediante el cual se quiere determinar si existe diferencia de las medias entre el sistema simulado y el sistema real, considerando la variable "número de pacientes que llegaron al sistema". En la tabla 6 se muestran los resultados.

Tabla 6. Comparación del modelo simulado con el modelo real

Número de corridas	Pacientes que llegaron al sistema simulado	Pacientes que llegaron al sistema real	Diferencia (*d*)
1	107	118	−11
2	113	124	−11
3	128	112	16
4	110	128	−18
5	141	114	27
6	112	109	3
7	113	115	−2
8	128	123	5
9	115	111	4
10	141	125	16
11	122	127	−5
12	114	106	8
13	118	113	5
14	124	126	−2
15	117	105	12
16	123	119	4
17	112	115	−3
18	133	129	4
19	126	109	17
20	110	123	−13
$\bar{x}$	*120,35*	*117,55*	*2,8*
s	*10,03*	*7,67*	*11,39*

Número de corridas	Pacientes que llegaron al sistema simulado	Pacientes que llegaron al sistema real	Diferencia (*d*)
e	*4,69*	-	-
Límite superior	*125,04*	-	-
Límite inferior	*115,66*	-	-

Fuente: elaboración propia.

Para la prueba de hipótesis se consideró un nivel de significancia del 5 %, 20 datos y un valor *t* (distribución *t* de Student) de 2,093. El planteamiento de la hipótesis fue la siguiente:

Hipótesis nula, H_o = la diferencia de las medias entre el sistema simulado y el real es igual a cero (H_o: $\mu_d = 0$).

Hipótesis nula, H_1 = la diferencia de las medias entre el sistema simulado y el real es mayor que cero (H_1: $\mu_d > 0$).

Luego de aplicar la ecuación (4), se obtuvo un valor t_o = 1,093, el cual es menor que el valor crítico de $t = 2{,}093$; por lo tanto, se acepta la hipótesis nula y se concluye que no hay diferencia significativa entre las medias del modelo del sistema simulado y el sistema real. De esta manera, se valida el modelo de simulación.

Análisis

Luego de las corridas respectivas de la simulación, se obtuvieron los resultados de los tiempos promedio de atención en cada etapa del proceso, los tiempos promedio de espera de cada paciente (cola) y los porcentajes de utilización de cada uno de los recursos considerados, los cuales se presentan en las tablas 7 y 8.

Tabla 7. Tiempos promedio por proceso y en espera (cola) por paciente

Proceso	Tiempo promedio (en horas)	Cola	Tiempo promedio (en horas)
Preadmisiones	0,0185	Preadmisiones	0,4124
Triage	0,2146	*Triage*	0,2834
Admisiones	0,0872	Admisiones	0,0223
Sala de procedimientos	1,0590	Sala de procedimientos	0,0112
UCIM	10,9887	UCIM	0
Consulta médica	0,3199	Consulta médica	0,2635
Laboratorio	1,2144	Laboratorio	0,2612
Observación	11,9216	Observación	3,2125
Revaloración	0,2576	Revaloración	0,2563
Facturación	0,1152	Facturación	0,0952

Fuente: elaboración propia.

Tabla 8. Porcentaje de utilización de cada recurso

Recurso	Utilización promedio (en porcentaje)
Auxiliar de *triage*	96,40
Vigilante	95,33
Auxiliar administrativo de admisión	72,50
Auxiliar administrativo de facturación	56,41
Jefe de enfermera de *triage*	45,40
Jefe de enfermera-observación 1	52,16
Jefe de enfermera-observación 2	54,28
Jefe de enfermera-observación 3	47,65
Jefe de enfermera-observación 4	55,11
Bacteriólogos	21,25
Médico sala de procedimientos	26,13
Médico consultorio 1	85,41

Recurso	Utilización promedio (en porcentaje)
Médico consultorio 2	79,82
Médico consultorio 3	82,12
Camillas de observación	99,22
Camillas en el espacio libre (observación)	92,66
Camillas en los pasillos (observación)	91,82
Auxiliares de enfermería	85,91

Fuente: elaboración propia.

Respecto al tiempo promedio en proceso, en la tabla 7 se aprecia que el tiempo que el paciente permanece en observación y en la UCIM son los mayores, con 11,9216 h y 10,9887 h, respectivamente, y respecto al tiempo en cola, los pacientes tienen que esperar el mayor tiempo para que les asignen una camilla en observación (3,2125 h), seguido de la espera para preadmisiones, *triage* y consulta médica. En la tabla 8 se observa que el vigilante y la auxiliar del *triage* son los que presentan mayor ocupación, con 96,40 % y 95,33 %, respectivamente. Los resultados de las tablas 7 y 8 fueron validados por el coordinador de urgencias del hospital.

A fin de mejorar el desempeño actual de la prestación del servicio de urgencias en función de la oportunidad y calidad, se plantearon dos escenarios de propuestas para implementar en el corto y el mediano plazo. El primer escenario surgió al considerar la etapa de preadmisiones como un exceso de proceso, ya que algunas clínicas no lo consideran, y el segundo escenario nace del requerimiento del coordinador de urgencias, para argumentar la necesidad de más personal médico en horarios de alta demanda. Dichos escenarios fueron corridos en el *software* de simulación y se describen a continuación:

Escenario 1. Eliminación del proceso de preadmisiones. Al llegar los pacientes directamente al proceso de admisiones quedarán registrados desde ese momento y así se podrá calcular el indicador de pacientes que salen del sistema sin ser vistos por un médico, ya que este se desconoce. Al eliminar esta etapa, también se eliminaría su tiempo promedio de proceso y de espera (25,854 minutos) y las actividades que realiza la auxiliar de enfermería, lo que disminuye su porcentaje de ocupación, para que pueda desarrollar otras tareas; además, aumentaría el egreso de pacientes en promedio un 34,47 %, equivalente a 6 personas adicionales (tabla 9). Ello mejoraría la eficiencia del sistema, la prestación del servicio y, por ende, la satisfacción de los pacientes. Además, esta propuesta reduce en al menos 30 minutos la duración promedio de la estancia de los pacientes

dentro del sistema y no requiere ninguna inversión para su implementación. En la tabla 9 se muestra el número personas atendidas con el sistema real y el sistema simulado en este primer escenario.

Tabla 9. Número de personas atendidas en el escenario 1 y en el sistema real

Número de corrida	Pacientes atendidos en el escenario 1	Pacientes atendidos en el sistema real	Diferencia	Diferencia (en porcentaje)
1	28	22	6	27,27
2	23	17	6	35,29
3	22	16	6	37,50
4	19	11	8	72,73
5	28	24	4	16,67
6	26	22	4	18,18
7	26	20	6	30,00
8	29	21	8	38,10
$\bar{x}$	-	-	*6*	*34,47*

Fuente: elaboración propia.

Escenario 2. Incrementar el número de médicos generales. Para determinar el número de médicos requeridos en el proceso de consulta en el periodo de mayor congestión del sistema de urgencias, se aplicó la teoría analítica de colas bajo el sistema de colas (*M/M/c*), donde se considera lo siguiente: 1) una población de pacientes infinita; 2) un proceso de llegadas de pacientes de acuerdo con un proceso Poisson, con una tasa promedio λ de pacientes por unidad de tiempo; 3) un proceso de colas de una sola fila de espera con capacidad infinita, con una disciplina de colas de primero en entrar, primero en salir; 4) un proceso de consulta con *c* médicos idénticos, con una distribución exponencial con una cantidad promedio μ de pacientes por unidad de tiempo (Mathur y Solow, 1996). El número de médicos se determinó por medio del *software* Winqsb, el cual determina las medidas de desempeño para cada número de médicos definido, considerando los siguientes parámetros de operación (tabla 10):

λ = 10 pacientes por hora
μ = 4 pacientes por hora
c = de 2 a 6 médicos
Cs = COP 45 000/hora (costo por hora por médico)
Cw = COP 44 520/hora (costo por hora por paciente en el sistema)

Tabla 10. Determinación del número de médicos por teoría de colas

Medida de desempeño	Número de médicos (*c*)				
	2	3	4	5	6
Utilización (*U*) (en porcentaje)	Sistema inestable	83,33	62,50	50	41,67
Pacientes promedio en espera (*Lq*)		3,5112	0,5331	0,1304	0,0339
Pacientes promedio en el sistema (*L*)		6,0112	3,0331	2,6304	2,5339
Probabilidad de que un paciente tenga que esperar (*Pw*)		0,7022	0,3199	0,1304	0,0474
Tiempo promedio de espera (*Wq*) (en horas)		0,3511	0,0533	0,013	0,0034
Tiempo promedio en el sistema (*W*) (en horas)		0,6011	0,3033	0,263	0,2534
Costo total médicos (*Cs*c*) (en pesos colombianos)		135 000	180 000	225 000	270 000
Costo total pacientes en el sistema (*Cw*L*) (en pesos colombianos)		267 619	135 034	117 105	112 809
Costo total (en pesos colombianos)		402 619	315 034	342 105	382 809

Fuente: elaboración propia.

Con base en los resultados arrojados por el *software* Winqsb de la tabla 10, se determinó que el número de médicos generales que presenta mejores medidas de desempeño en términos de tiempos de espera y costos equivale a cuatro, tres que ya están en los consultorios (modelo actual) y uno adicional para el periodo de mayor congestión de los días lunes, martes y miércoles. Posteriormente, se realizó la simulación considerando los cuatro médicos para la consulta, cuyos resultados se observan en la tabla 11. Se destaca la reducción del tiempo promedio de espera para recibir la consulta médica, que pasó de 15,81 a 5,16 minutos, con una reducción del 67,36 %, y de la espera para la revaloración, que pasó de 15,38 a 4,71 minutos, con una reducción del 69,37 %. Ello mejoró la oportunidad en la atención de los pacientes, la duración de la estancia y, por ende, la sobreocupación de la sala de urgencias.

Tabla 11. Tiempos promedio de espera en el escenario 2 y en el modelo actual

Proceso	Escenario 2 (en horas)	Modelo actual (en horas)
Preadmisiones	0,4112	0,4124
Triage	0,2826	0,2834
Admisiones	0,0231	0,0223
Sala de procedimientos	0,0117	0,0112
UCIM	0,0000	0,0000
Consulta médica	0,0860	0,2635
Laboratorio	0,2418	0,2612
Observación	3,0820	3,2125
Revaloración	0,0785	0,2563
Facturación	0,0520	0,0952

Fuente: elaboración propia.

Documentación e implementación

El estudio fue presentado a las directivas del hospital, con el fin de que sirviera de soporte técnico para la toma de decisiones operacionales y que mejorara la prestación de los servicios de urgencias.

Análisis y discusión

En el 2017, el hospital atendió 39 714 pacientes en el servicio de urgencias de adultos, de los cuales el 7,34 % eran miembros de la población inmigrante venezolana, quienes presentaron una tendencia creciente durante el año, desde el 0,8 % de los pacientes atendidos en enero hasta el 13,8 % en diciembre, debido a la crisis económica, política y social de Venezuela. Con la aprobación del Decreto 1288 del 2018 por parte de la Presidencia de la República de Colombia, que establece que "teniendo en cuenta la Constitución Política de 1991, las leyes 1751 de 2015, 100 de 1993, entre otras, la atención de urgencias debe ser prestada a toda persona nacional o extranjera, sin ninguna exigencia ni discriminación", el problema de la sobreocupación podría empeorar si se mantienen los recursos disponibles actualmente. En Italia, las salas de urgencias han tenido que enfrentar un conjunto de demandas que deberían ser manejadas por las unidades de atención primaria y, recientemente, el incremento de los inmigrantes, con frecuencia ilegales (Leo et al., 2015).

Con respecto al tiempo promedio que el paciente permanece en observación (11,9216 h), ocurre, por la falta de disponibilidad de camas en hospitalización, que no permite transferir pacientes a esta siguiente etapa. De forma similar lo argumentan Abo-Hamad y Arisha (2012), Hoot y Aronsky (2008) y Lee (2014), porque se presenta un bloqueo del acceso desde el servicio de urgencias hacia la hospitalización por la escasez de camas. Otra razón es porque algunos pacientes pertenecen a una EPS que tiene una deuda con el hospital y no tiene un convenio vigente, por lo cual deben recibir el tratamiento en observación y se impide la entrada de un nuevo paciente a esta área del servicio de urgencias. Lo anterior resulta afectando los tiempos en cola para que le asignen al paciente una camilla en observación (3,2125 h), debido a que estas generalmente están ocupadas, con un porcentaje de utilización del 99,22 %. Lo mismo sucede a la del tiempo promedio que el paciente permanece en la UCIM (10,9887 h), pues faltan camas disponibles en la UCI del hospital.

Eliminar el proceso de preadmisiones le permitirá al hospital determinar el indicador de pacientes que salen del sistema sin ser vistos por un médico, ya que los usuarios quedarían registrados en admisión desde el momento que arriban a la sala de urgencias. En Khalifa (2015) se identificó que los pacientes se van sin ser atendidos, debido a los largos tiempos de espera, y muchos de ellos buscarán otro tipo de atención médica. Los resultados de la simulación del estado actual del sistema indican que los pacientes, en promedio, deben esperar 1,3019 horas o 78,114 minutos para recibir atención por parte de un médico, por lo cual es posible que los pacientes abandonen el sistema. Con este primer escenario propuesto, el tiempo promedio de espera se disminuye a 52,26 minutos, con una reducción del 33,1 %, que muestra un comportamiento cercano a las estadísticas presentadas por Rai y Kang (2015), donde en los Estados Unidos de los cerca de 137 millones de personas que visitaron las salas de urgencias en el 2015, el 32,1 % tuvo que esperar entre 15 y 59 minutos para recibir atención médica. Además, se incrementa el egreso de pacientes en promedio un 34,47 %, situación similar a la del estudio de Ibrahim et al. (2017), donde se incrementó el número total de pacientes atendidos en un 25 %.

Los servicios de urgencias requieren una adecuada asignación de recursos humanos y materiales para aumentar su eficacia y eficiencia y reducir al máximo el tiempo de espera de los pacientes (Masmoudi et al., 2014). En la presente investigación, la propuesta de incluir un nuevo médico general reduce el tiempo de espera en 10,65 minutos (67,36 %) y la espera para la revaloración en 10,67 minutos (69,37 %), lo que disminuye la duración promedio de la estancia en al menos 20,32 minutos. De manera similar, la investigación de Bedoya-Valencia y Kirac (2016) evidenció cómo disminuye el tiempo de espera para que un paciente sea visto por un médico y la estancia un 16 % (11 minutos) y un 14 % (20 minutos), respectivamente, al incluir un médico adicional en el personal. Por su parte, Sin-

reich et al. (2012), al ajustar los turnos de trabajo del personal médico, enfermeras y radiólogos, reducen la media del tiempo total de espera de los pacientes entre el 20 % y el 64 %, y reducen entre el 11 % y el 29 % de la estancia media combinada de los pacientes.

El estudio de Ibrahim et al. (2017) muestra que, al añadir más recursos en el proceso de cuello de botella, el tiempo medio de espera de los pacientes puede reducirse en 148,32 minutos (2,47 horas), lo que supone una reducción del 32 %. En Masmoudi et al. (2014), de acuerdo con los resultados, modificar el número de médicos puede tener un fuerte efecto en la duración de la estancia de los pacientes y en el número de pacientes que abandonan el servicio. El análisis de los resultados del estudio de Weng et al. (2011) señaló que el rendimiento global de los servicios de urgencias puede incrementarse en un 8 %, gracias a la nueva asignación de médicos y enfermeras estudiada. Finalmente, Bedoya-Valencia y Kirac (2016) concluyeron que el modelo de simulación puede ser utilizado como una herramienta para ayudar a los responsables de la toma de decisiones en la sala de urgencias del hospital en la asignación de recursos.

Conclusiones

Esta investigación propuso escenarios de mejoras en la prestación de los servicios de urgencias de un hospital público ubicado en una ciudad fronteriza de Colombia, con la finalidad de reducir el tiempo promedio de espera para la consulta médica y la duración promedio de la estancia de los pacientes, por medio del diseño de un modelo de simulación de eventos discretos. El escenario 1 elimina el proceso de preadmisiones, que traerá como beneficios que la institución de salud: mida el número de pacientes que salen del sistema sin ser vistos por un médico; reduzca el tiempo promedio de espera para la consulta en 25,854 minutos; elimine las actividades que son realizadas por la auxiliar de enfermería, lo que disminuye su nivel de ocupación; incremente la salida de pacientes en promedio en un 34,47 %, y reduzca en al menos 30 minutos la duración promedio de la estancia de los pacientes dentro del sistema. Este escenario no requiere inversión alguna para su implementación y mejora la eficiencia del sistema, la prestación del servicio y, por ende, la satisfacción de los pacientes.

En el escenario 2 se determinó el número de médicos generales adicionales para el proceso de consulta por medio de la teoría de colas, para el periodo de mayor congestión en la sala de urgencias. Al incrementar en uno la capacidad de este proceso, los resultados de la simulación muestran que se reduce el tiempo promedio de espera para recibir la consulta en 10,65 minutos, equivalente a una mejora del 67,36 %, y la espera para la revaloración en 10,67 minutos, equivalente a una mejora del 69,37 %, lo que se traduce en una disminución de la duración

promedio de la estancia en al menos 20,32 minutos. Esto mejora la oportunidad en la atención de los pacientes y, por ende, la sobreocupación de la sala de urgencias.

Se identificó el incremento en la demanda de pacientes en los servicios de urgencias durante el 2017, con un crecimiento en la participación de la población inmigrante venezolana, del 0,8 % del total de pacientes atendidos en enero al 13,8 % en diciembre. Por tal razón, inicialmente las directivas del hospital deben tomar decisiones operacionales y tácticas, en el corto y mediano plazo, apoyándose en los resultados de esta investigación, obtenidos por medio del modelo de simulación diseñado, que permite enfrentar el problema de la sobreocupación de la sala de urgencias objeto de estudio y mejorar la calidad y oportunidad en la prestación de los servicios de urgencias.

La metodología de la simulación aplicada en esta investigación permitió conocer el funcionamiento del sistema de servicios de urgencias y, desde allí, proponer escenarios de mejora, cuyas decisiones tácticas u operacionales analizadas y evaluadas respondan a las medidas de desempeño clave en el funcionamiento del sistema; así como tener viabilidad para su implementación en el corto o mediano plazo. De este modo, el proceso desarrollado en la investigación para el alcance del objetivo se presenta como un referente a la hora de plantear mejoras en cualquier sistema de servicios de urgencias de clínicas y hospitales que presenten el problema de la sobreocupación.

Como trabajo futuro se pretende evaluar escenarios del orden estratégico por medio del modelo de simulación diseñado, que considere incrementos en la demanda de pacientes y en la capacidad del sistema de servicios, que sirva para respaldar el proceso de toma de decisiones a largo plazo como la inversión en infraestructura, equipos y personal médico y asistencial, con el objetivo de mejorar la prestación de los servicios de salud en la sala de urgencias del hospital público. Finalmente, el modelo de simulación diseñado se constituye en una herramienta flexible y valiosa, para apoyar la toma de decisiones estratégicas, tácticas y operacionales por parte de los directivos de la institución de salud.

Agradecimientos

Al personal médico, asistencial y administrativo del hospital, por el apoyo recibido en la realización de la investigación.

Referencias

Abo-Hamad, W. y Arisha, A. (2012). Multi-criteria framework for emergency department in Irish hospital. En E. Laroque, J. Himmelspach, R. Pasupathy, O. Rose y A. M. Uhrmacher (Eds.), *Proceedings of the 2012 Winter Simulation Conference C.* (pp. 1-12). IEEE. http://doi.org/10.1109/WSC.2012.6465229

Afilal, M., Yalaoui, F., Dugardin, F., Amodeo, L., Laplanche, D. y Blua, P. (2016). Forecasting the emergency department patients flow. *Journal of Medical Systems*, *40*(7). http://doi.org/10.1007/s10916-016-0527-0

Ajami, S., Ketabi, S., Yarmohammadian, M. H. y Bagherian, H. (2012). Wait time in emergency department (ED) processes. *Medical Archives*, *66*(1), 53-57. http://doi.org/10.5455/medarh.2012.66.53-57

Aringhieri, R., Bruni, M. E., Khodaparasti, S. y Essen, J. T. van. (2016). Emergency medical services and beyond: Addressing new challenges through a wide literature review. *Computers & Operations Research*, *78*, 349-368. http://doi.org/10.1016/j.cor.2016.09.016

Arunachalam, S. P., Sir, M., Marisamy, G., Sadosty, A., Nestler, D., Hellmich, T. y Pasupathy, K. (2017). Identifying factors influencing patient alone time at the emergency department using RFID data: What is next? En *2017 IEEE EMBS International Conference on Biomedical and Health Informatics, BHI 2017* (pp. 225-228). IEEE. http://doi.org/10.1109/BHI.2017.7897246

Ashour, O. M. y Okudan Kremer, G. E. (2016). Dynamic patient grouping and prioritization: A new approach to emergency department flow improvement. *Health Care Management Science*, *19*(2), 192-205. http://doi.org/10.1007/s10729-014-9311-1

Asociación Colombiana de Empresas de Medicina Integral. (2017). *Cifras e indicadores del sistema de salud*. https://www.acemi.org.co/index.php/acemi-defecto/20-actualidad/noticias/983-cifras-e-indicadores-del-sistema-de-salud-publicacion-2017

Azadeh, A., Sepahi, M. y Haghighi, S. M. (2013). An integrated simulation-DEA approach to improve quality care of medical centres. *International Journal of Process Management and Benchmarking*, *3*(3), 352. http://doi.org/10.1504/IJPMB.2013.058160

Banks, J. y Carson, J. (1984). *Discrete-event system simulation*. Prentice Hall.

Bedoya-Valencia, L. y Kirac, E. (2016). Evaluating alternative resource allocation in an emergency department using discrete event simulation. *Simulation*, *92*(12), 1041-1051. http://doi.org/10.1177/0037549716673150

Chan, C. L., Huang, H. T. y You, H. J. (2012). Intelligence modeling for coping strategies to reduce emergency department overcrowding in hospitals. *Journal of Intelligent Manufacturing*, *23*(6), 2307-2318. http://doi.org/10.1007/s10845-011-0574-9

Chen, T. L. y Wang, C. C. (2016). Multi-objective simulation optimization for medical capacity allocation in emergency department. *Journal of Simulation*, *10*(1), 50-68. http://doi.org/10.1057/jos.2014.39

Cocke, S., Guinn, D., MacBlane, E., Walshak, S., Willenbrock, N., White, K. P., … Kang, H. (2016). UVA emergency department patient flow simulation and analy-

sis. En *2016 IEEE Systems and Information Engineering Design Symposium (SIEDS)* (pp. 118-123). IEEE. http://doi.org/10.1109/SIEDS.2016.7489282

Daldoul, D., Nouaouri, I., Bouchriha, H. y Allaoui, H. (2016). Optimization on human and material resources in Emergency Department. In *Proceedings of 2015 International Conference on Industrial Engineering and Systems Management, IEEE IESM 2015* (pp. 633-638). IEEE. http://doi.org/10.1109/IESM.2015.7380224

Decreto 1288 (2018, 25 de julio). Por el cual se adoptan medidas para garantizar el acceso de las personas inscritas en el Registro Administrativo de Migrantes Venezolanos a la oferta institucional y se dictan otras medidas sobre el retorno de colombianos. Departamento Administrativo de la Presidencia de la República. http://es.presidencia.gov.co/normativa/normativa/DECRETO 1288 DEL 25 DE JULIO DE 2018.pdf

Ghanes, K., Jouini, O., Wargon, M. y Jemai, Z. (2016). Modeling and analysis of triage nurse ordering in emergency departments. En *Proceedings of 2015 International Conference on Industrial Engineering and Systems Management, IEEE IESM 2015* (pp. 228-235). IEEE. http://doi.org/10.1109/IESM.2015.7380163

Guasch, A., Piera, M. Á., Casanovas, J. y Figueras, J. (2005). *Modelado y simulación: Aplicación a procesos logísticos de fabricación y servicios*. Alfaomega.

Guo, H., Gao, S., Tsui, K. L. y Niu, T. (2017). Simulation optimization for medical staff configuration at emergency department in Hong Kong. *IEEE Transactions on Automation Science and Engineering*, *14*(4), 1655-1665. http://doi.org/10.1109/TASE.2017.2697899

Guo, H. y Tang, J. (2015). Integrating simulation with optimization in emergency department management. En P. W. Tse, J. Mathew, K. Wong, R. Lam y C. N. Ko (Eds.), *Engineering asset management: Systems, professional practices and certification, lecture notes in mechanical engineering* (pp. 1483-1496). Springer, Cham. http://doi.org/10.1007/978-3-319-09507-3

Hoot, N. y Aronsky, D. (2008). Systematic review of emergency department crowding: Causes, effects, and solutions. *Annals of Emergency Medicine*, *52*(2), 126-136. http://doi.org/10.1016/j.annemergmed.2008.03.014

Househ, M. y Yunus, F. (2014). Emergency department waiting times (EDWaT): A patient flow management and quality of care rating mHealth application. *Studies in Health Technology and Informatics*, *202*, 229-232. http://doi.org/10.3233/978-1-61499-423-7-229

Huang, Y.-C. y Chu, C.-P. (2011). RFID applications in hospitals: A case study for emergency department. *Journal of Communication and Computer*, *8*(7), 1548-7709.

Ibrahim, I. M., Liong, C.-Y., Bakar, S. A., Ahmad, N. y Najmuddin, A. F. (2017). Minimizing patient waiting time in emergency department of public hospital using simulation optimization approach. *AIP Conference Proceedings, 1830*(1). http://doi.org/10.1063/1.4980949

Khalifa, M. (2015). Developing an emergency physician productivity index using descriptive health analytics. En J. Mantas, A. Hasman y M. S. Househ (Eds.), *Enabling health informatics applications* (p. 320). IOS Press. http://doi.org/10.3233/978-1-61499-538-8-167

Khalifa, M. y Zabani, I. (2016). Reducing emergency department crowding: Evidence based strategies. En J. Mantas, A. Hasman, G. Gallos, A. Kolokathi y M. S. Househ (Eds.), *Unifying the applications and foundations of biomedical and health informatics* (pp. 67-70). IOS Press. http://doi.org/10.3233/978-1-61499-664-4-67

Kuo, Y. H., Rado, O., Lupia, B., Leung, J. M. Y. y Graham, C. A. (2016). Improving the efficiency of a hospital emergency department: A simulation study with indirectly imputed service-time distributions. *Flexible Services and Manufacturing Journal*, *28*(1-2), 120-147. http://doi.org/10.1007/s10696-014-9198-7

Leal, D. y Ortiz, L. (2018). *Propuesta de mejora en el proceso de atención del área de urgencias en el E. S. E. Hospital Universitario Erasmo Meoz de la ciudad de Cúcuta* [tesis no publicada]. Universidad Francisco de Paula Santander, Colombia.

Lee, S. (2014). The role of hospital selection in ambulance logistics. *IIE Transactions on Healthcare Systems Engineering*, *4*(2), 105-117. http://doi.org/10.1080/19488300.2014.914608

Leo, G., Lodi, A., Tubertini, P. y Di Martino, M. (2015). Emergency department management in Lazio, Italy. *Omega*, *58*, 128-138. http://doi.org/10.1016/j.omega.2015.05.007

Luscombe, R. y Kozan, E. (2016). Dynamic resource allocation to improve emergency department efficiency in real time. *European Journal of Operational Research*, *255*(2), 593-603. http://doi.org/10.1016/j.ejor.2016.05.039

Masmoudi, M., Leclaire, P., Vincent, C. y Casalino, E. (2014). Modelling and simulation of the doctors' availability in emergency department with SIMIO software. Case of study: Bichat-Claude Bernard Hospital. En M. S. Abbes, J.-Y. Choley, F. Chaari, A. Jarraya y M. Haddar (Eds.), *Mechatronic systems: Theory and applications* (pp. 119-129). Springer. http://doi.org/10.1007/978-3-319-07170-1

Mathur, K. y Solow, D. (1996). *Investigación de operaciones: el arte de la toma de decisiones*. Prentice Hall Hispanoamericana.

Ministerio de Salud y Protección Social (MSPS). (2015a). *Plan estratégico sectorial 2014-2018: sector administrativo de salud y protección social*. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/PES/plan-estrategico-sectorial-2014-2018.pdf

Ministerio de Salud y Protección Social (MSPS). (2015b). *Informe nacional de calidad de la atención en salud 2015*. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/informe-nal-calidad-atencion-salud-2015.pdf

Ministerio de Salud y Protección Social (MSPS). (2017). *Encuesta nacional de evaluación de los servicios de las EPS: informe final Colombia 2017*. https://www.

minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/CA/Informe-encuesta-satisfaccion-eps-2017.pdf

Observatorio Nacional de Calidad en Salud. (2019). *Indicadores de calidad: Resolución 0256/16*. Ministerio de Salud y Protección Social. http://calidadensalud.minsalud.gov.co/Paginas/Indicadores.aspx

Rui, P. y Kang, K. (2015). *National hospital ambulatory medical care survey: 2015 emergency department summary tables*. https://www.cdc.gov/nchs/data/nhamcs/web_tables/2015_ed_web_tables.pdf

Rozo Lesmes, P. y Acosta Hernández, A. L. (2016). *Caracterización: registro especial de prestadores de servicios de salud (REPS)-IPS. Estudios sectoriales*. Ministerio de Salud y Protección Social. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/PES/caracterizacion-registro-especial-prestadores-reps.pdf

Sepehri, Z., Arabzad, S. M. y Sajadi, S. M. (2015). Analysing the performance of emergency department by simulation: The case of Sirjan Hospital. *International Journal of Services and Operations Management*, *20*(3), 289. http://doi.org/10.1504/IJSOM.2015.067769

Sinreich, D., Jabali, O. y Dellaert, N. P. (2012). Reducing emergency department waiting times by adjusting work shifts considering patient visits to multiple care providers. *IIE Transactions*, *44*(3), 163-180. http://doi.org/10.1080/0740817X.2011.609875

Tsai, C.-H., Weng, S.-J., Huang, C.-Y., Yang, C.-Y., Li, J.-Y. y Tsai, K.-T. (2010). The impact of different decision behavior models of emergency physicians on the performance of emergency departments. En *Proceedings of the 2010 Spring Simulation Multiconference* (pp. 1-10). http://doi.org/10.1145/1878537.1878614

Velásquez-Restrepo, P. A., Rodríguez-Quintero, A. K. y Jaén-Posada, J. S. (2011). Metodologías cuantitativas para la optimización del servicio de urgencias. *Gerencia y Políticas de Salud*, *10*(21), 196-218.

Weng, S.-J., Cheng, B.-C., Kwong, S. T., Wang, L.-M. y Chang, C.-Y. (2011). Simulation Optimization for emergency department resources allocation. En *Proceedings of the 2011 Winter Simulation Conference* (pp. 1231-1238). IEEE. https://doi.org/10.1109/WSC.2011.6147845

Zhang, X., Zhu, T., Luo, L., He, C., Cao, Y. y Shi, Y. (2013). Forecasting emergency department patient flow using Markov chain. En *2013 10th International Conference on Service Systems and Service Management (ICSSSM)* (pp. 278-282). IEEE. https://doi.org/10.1109/ICSSSM.2013.6602537

Zhao, Y., Peng, Q., Strome, T., Weldon, E., Zhang, M. y Chochinov, A. (2015). Bottleneck detection for improvement of emergency department efficiency. *Business Process Management Journal*, *21*(3), 564-585. http://doi.org/10.1108/BPMJ-06-2014-0060

Capítulo 5.
Impacto de la inoportunidad de ayudas diagnósticas e interconsultas en la duración de la estancia hospitalaria

Tatiana María Ceballos Acevedo,[1]
Paula Andrea Velásquez Restrepo[2]
y Juan Sebastián Jaén-Posada[3]

[1] Desarrollo organizacional
Institución prestadora de servicios
de salud, IPS Universitaria
Medellín, Colombia
tatiana.ceballosa@ipsuniversitaria.com.co

[2] Desarrollo organizacional
Institución prestadora de servicios
de salud, IPS Universitaria
Medellín, Colombia
paula.velasquezr@ipsuniversitaria.com.co

[3] Departamento de Ingeniería Industrial
Universidad de Antioquia
Medellín, Colombia
juan.jaen@udea.edu.co

Resumen

La estancia hospitalaria es una medida global de la eficiencia relacionada directamente con el uso de la cama y atención del paciente. La estancia puede prolongarse si los servicios de soporte a la atención presentan retrasos o no están en sincronía con la atención del paciente. En este capítulo se presenta una metodología para medir el impacto en días de estancia que tienen los retrasos en la entrega de ayudas diagnósticas e interconsultas a un grupo relacionado de diagnóstico específico (GRD) en la eficacia y eficiencia de la ronda médica y, por ende, en la estancia del paciente. Se desarrolló un modelo de simulación bajo dinámica de sistemas que representa la interacción de los procesos que afectan la estancia en una unidad de medicina interna. El análisis muestra que políticas de sincronización entre los tiempos de servicio de las ayudas diagnósticas, interconsultas y ronda médica podrían reducir hasta en un 6,2 % la estancia hospitalaria de pacientes bajo el GRD estudiado, lo que implica disminuir hasta en 0,71 días (17 horas) el promedio de la estancia hospitalaria de ese GRD. Esto permitiría una ganancia efectiva de aproximadamente 63 camas para la unidad de medicina interna que se estudió.

Palabras clave: estancia hospitalaria, sincronía de servicios de salud, gestión de la cama, flujo del paciente, dinámica de sistemas

Introducción

En los países latinoamericanos, la desproporción entre oferta y demanda de servicios de salud es un importante motivo de preocupación (García Fariñas et al. 2011). En el caso de Colombia, el sistema de salud atraviesa una crisis por la incontenible demanda de los servicios de salud, que supera la capacidad de los hospitales para atenderla (Afanador, 2012). Esta situación se ve también reflejada en la saturación de los servicios de urgencias, al punto de que se presentan casos de cierre temporal (Correa Arcila, 2010). Ante la desproporción entre oferta y demanda de servicios de salud que conlleva la saturación de los servicios, la eficiencia es una condición importante, debido a que los recursos son limitados. En los hospitales, la cama es uno de los recursos centrales para la prestación de servicios. Por eso, entre las medidas globales de la eficiencia hospitalaria están aquellas que se relacionan con el uso que se le dé a este recurso. El promedio de la estancia hospitalaria de pacientes se utiliza como indicador de eficiencia, porque resume el aprovechamiento de la cama y la agilidad de los servicios prestados en los hospitales (Jiménez Paneque, 2004). Al presentarse una mayor estadía, hay menor disponibilidad de camas y mayor saturación en los servicios de urgencias, por la espera de una cama (Kim et al., 2011). Además de la relación directa entre la prolongación de la estancia y la disponibilidad de camas, también se aumentan los costos de la atención, el riesgo de eventos adversos y la mortalidad (González-Angulo et al., 2009). Una estancia se considera prolongada cuando esta supera un estándar de tiempo de un tipo de paciente con determinada cantidad de diagnósticos en un contexto de atención específico (Ceballos-Acevedo et al., 2014).

Ante esto, las instituciones de salud tienen la necesidad de reducir la estancia hospitalaria para aumentar el flujo de pacientes a través de mayores niveles de eficiencia en la operación, sin afectar la calidad de la atención. Según el grupo de investigación del Advisory Board International (Tell et al., 2010), los hospitales que logran disminuir las estancias prolongadas de pacientes pueden agregar camas y aumentar la capacidad efectiva. El estudio afirma que en un hospital con 300 camas, en promedio, restar la duración de la estancia en un día equivale a contar con 49 nuevas camas en la institución (Tell et al., 2010). Esto tiene un impacto significativo en la accesibilidad de pacientes a los servicios de hospitalización, la descongestión la saturación de los servicios de urgencias y la generación de ganancias para los hospitales.

Las estancias prolongadas en los hospitales se originaron en las diferentes características y roles de cada uno de los actores del sistema: personal de salud, administración hospitalaria, pacientes y relación entre entidades de la red de aten-

ción (Ceballos-Acevedo et al., 2014). Frente a la administración hospitalaria, se observa que la larga estancia de pacientes está asociada a la eficiencia en la interacción del servicio de hospitalización con otros servicios del hospital, como los tiempos de respuesta de los procedimientos quirúrgicos, las ayudas diagnósticas, las interconsultas y la transferencia del paciente entre servicios (Kim et al., 2011; Zambrana García et al., 2001; Negro Álvarez et al., 2003; Silber et al., 2009). Las ayudas diagnósticas corresponden a todo tipo de pruebas médicas y exámenes físicos y de laboratorio que se le realizan al paciente para determinar un diagnóstico o la evolución de un tratamiento. Las interconsultas corresponden a la emisión de conceptos médicos o profesionales que orientan el proceso de diagnóstico o de evaluación de la condición del paciente al médico tratante (Ceballos-Acevedo et al., 2014). Se encontró que la inoportunidad (demora) en la entrega de ambas ha sido reportada en un 64 % de la literatura consultada, y que constituye la primera causa más indicada respecto de la estancia prolongada. Las dos segundas causas alcanzaron frecuencias semejantes del 38 %, y corresponden a la espera para trasladar a otro paciente a otro nivel de complejidad y las complicaciones fruto de la edad del paciente (Ceballos-Acevedo et al., 2014).

El propósito de este capítulo es presentar una metodología para establecer el impacto que la inoportunidad y la desincronización de servicios de ayudas diagnósticas e interconsultas tienen en la estancia hospitalaria. Como una forma de validar la metodología se utiliza la información de un servicio de medicina interna en la Clínica León XIII de la ciudad de Medellín, Colombia, una institución de nivel tres de complejidad que cuenta con 500 camas, 14 quirófanos, 100 camillas en urgencias y un promedio de egresos de 27 000 pacientes por año. En la primera sección se presenta la metodología implementada. En la segunda sección se detallan los resultados encontrados. En la tercera se plantean el análisis y las discusiones. Finalmente, en la cuarta sección se exponen las conclusiones de este trabajo.

Métodos

Elección del grupo relacionado de diagnóstico con más generación de estancia prolongada

El primer criterio para definir es cuándo un paciente presenta una estancia prolongada. Existen diferentes métodos para establecer esta condición. Uno de los más usados corresponde a los grupos relacionados de diagnóstico (GRD) que agrupan a los pacientes por diagnóstico, nivel de severidad, manejo y estancia, para después compararlos con los estándares internacionales y, así, determinar el nivel de eficiencia y eficacia de los servicios por tipo de diagnóstico (Panis et al., 2003; Tamargo Barbeito et al., 2012). Existen otros métodos como el sistema de clasificación que evalúa la extensión de la expansión de tumores malignos, conocido

como *TNM staging system*, el índice de Horn, el índice de Salemi y el APACHE, o método específico para las unidades de cuidados intensivos (Jiménez Paneque, 2004). Entre todos los anteriores, se eligieron los GRD como criterio para definir las estancias prolongadas, dado que esta es una de las metodologías más reconocidas y utilizadas a nivel mundial para comparar estancias y eficiencias entre los hospitales y los países (Jiménez Paneque, 2004).

Para analizar la estancia hospitalaria por GRD en la Clínica León XIII, se procesó la información de 26 800 egresos hospitalarios del 2013. El número de GRD encontrados en la institución fue de 585, de los cuales el 37,26 % se clasificó como quirúrgico, y el 62,74 %, como clínico. La comparación con el estándar internacional determinó que el 80 % de los egresos de GRD clínicos de la institución están por encima del estándar; mientras que los que están por encima en los GRD quirúrgicos corresponden al 70 %. Con el objetivo de establecer los GRD con mayor impacto en estancia para la institución se ponderaron los GRD según el volumen de egresos de la institución y la diferencia entre la estancia promedio observada y la estancia estándar internacional. En la tabla 1 se presentan los cinco primeros GRD que generan mayor impacto en la institución. Con ello se identificó que las infecciones de los riñones y el tracto urinario corresponden al GRD que cumple con los dos criterios: una estancia prolongada y un volumen de pacientes elevado. Obsérvese que si bien el segundo GRD (otros trastornos del sistema nervioso) da una mayor estancia prolongada, el volumen de pacientes es insuficiente para superar en impacto al primer GRD encontrado.

Tabla 1. Identificación del GRD con mayor impacto en la estancia en la institución

GRD	Nivel de severidad	Tipo GRD	Egresos	Estancia observada	Estancia estándar GRD	Diferencia de estancia	Ponderado
Infecciones de los riñones y el tracto urinario	3	M	392	11,18	7,71	3,47	1360,2
Otros trastornos del sistema nervioso	3	M	235	13,91	8,15	5,76	1353,6
Procedimientos no complejos del aparato respiratorio	3	Q	269	14,39	9,45	4,94	1328,9
Infecciones de los riñones y el tracto urinario	2	M	381	8,48	5,44	3,04	1158,2
Neumonía simple y tosferina (pertusis)	3	M	339	11,27	8,46	2,81	952,6

Nivel de severidad: 1: baja; 2: media; 3: alta

Tipo de GRD: M: clínico; Q: quirúrgico

Diferencia de estancia: (estancia observada − estancia estándar GRD)

Ponderado: Egresos * (estancia observada − estancia estándar GRD)

Fuente: elaboración propia.

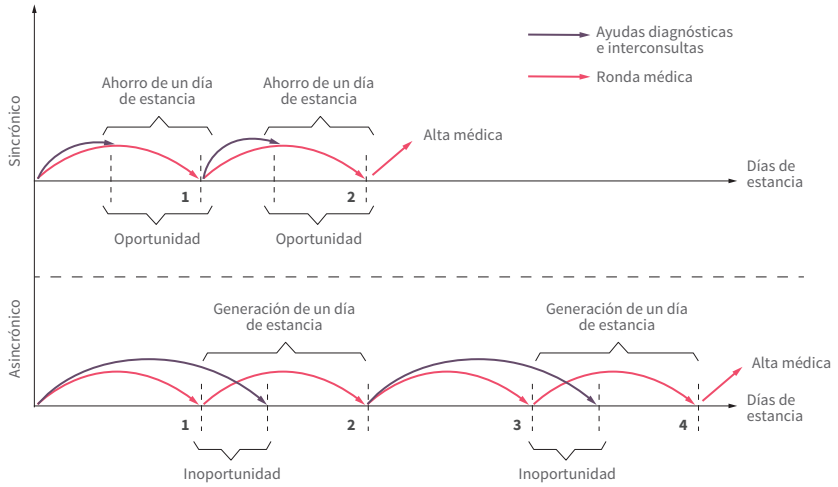
Descripción del proceso y elección de la metodología de simulación

Una vez identificado el GRD con mayor impacto en estancia en la institución, se simuló el proceso de atención de estos pacientes con el propósito de determinar en qué magnitud las demoras en los procesos de entrega de resultados de ayudas diagnósticas e interconsultas explican la prolongación de la estancia observada en ese GRD. En principio, se exploró la posibilidad de utilizar la simulación bajo eventos discretos, la cual ha sido ampliamente utilizada en el campo de la salud, como lo argumentan Jacobson et al. (2006), quienes proporcionaron una descripción general de las aplicaciones de simulación de eventos discretos en el sistema de atención de salud para los últimos cuarenta años de su estudio. Posteriormente, se encontró que la pregunta de investigación está en una dimensión táctico-estratégica, para la cual una aproximación bajo eventos discretos sería ineficiente en una primera etapa de indagación. El problema puede ser abordado de manera más eficiente desde el enfoque de la dinámica de sistemas (Sterman, 2006), dada

su capacidad para representar de manera agregada los procesos, los retardos y los ciclos que están presentes en este problema. Por otro lado, existe abundante literatura sobre el uso de la dinámica de sistemas en la simulación del flujo del paciente (Jaén, 2018; Rashwan et al., 2015; Vanderby y Carter, 2009).

En la figura 1 se describen los elementos del proceso que se van a simular, y se ilustra como caso hipotético la situación de un paciente que posee una condición que solo le demandaría dos días de estancia hospitalaria si los procesos de las ayudas diagnósticas y las interconsultas estuvieran listos a tiempo para confirmar el alta médica. A partir de aquí se pueden generar dos tipos de resultados, dependiendo de la sincronía (parte superior de la figura 1) o la asincronía (parte inferior de la figura 1) de las ayudas diagnósticas e interconsultas con el ciclo de rondas médicas.

Figura 1. Interacción entre la oportunidad de las ayudas diagnósticas e interconsultas y la generación de días de estancia



Fuente: elaboración propia.

Cada parte de la figura 1 cuenta con un eje horizontal en el que se enumeran los días de estancia del paciente. La línea roja corresponde al ciclo de ronda médica, que en la mayoría de los casos es una al día y coincide en número con los días de estancia del paciente. Por otro lado, la ronda médica es generadora de ayudas diagnósticas, interconsultas y de la alta médica, que depende del resultado de estas. Por esto, una condición ideal para generar el alta médica es el arribo previo de la ayuda diagnóstica y la interconsulta. En la parte superior de la figura 1 se presenta el caso ideal, en el cual las ayudas diagnósticas e interconsultas están

disponibles antes de que se presente la ronda médica y, por lo tanto, el alta médica se puede generar a los dos días de estancia.

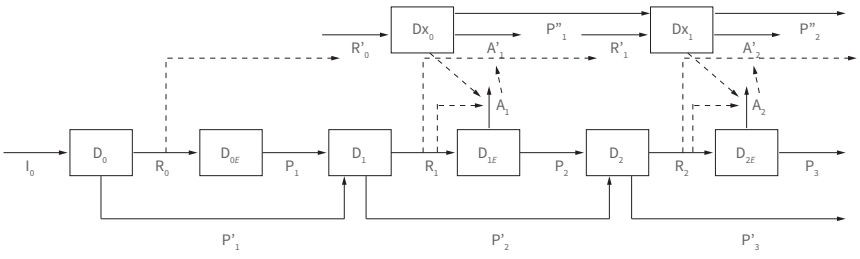
El caso opuesto es presentado en la parte inferior de la figura 1. Las ayudas diagnósticas e interconsultas arriban con posterioridad a la ronda médica y generan un nuevo día de estancia, por el hecho de que el médico tratante carece de información suficiente para evaluar la condición del paciente ese mismo día para generar el alta. Es posible que al otro día la ayuda diagnóstica o interconsulta ya se encuentre disponible, y es allí cuando el médico puede tener claridad sobre la condición del paciente, solicitar nuevas ayudas diagnósticas o interconsultas o generar el alta.

A partir de las posibles situaciones que pueden presentarse ante la sincronía o desincronía de las ayudas diagnósticas e interconsultas con la ronda médica, se realiza un modelo de simulación que permitiría evaluar, en un caso real, la magnitud de estas desincronías y su impacto en la estancia médica.

Simulación del proceso

El proceso se simula a partir del modelamiento de dos tipos de flujos: uno de pacientes y uno de ayudas diagnósticas e interconsultas. Los flujos de pacientes corresponden a cohortes diarias de pacientes donde cada nivel corresponde a la cantidad de pacientes que hay en un día de estancia determinado. Los flujos de ayudas diagnósticas e interconsultas corresponden a las cohortes diarias de ayudas diagnósticas e interconsultas disponibles en ese día. Como la estancia máxima en ese GRD, infecciones de los riñones y el tracto urinario, ha reportado un valor de 14 días, en la Clínica León XIII el modelo simula 14 cohortes de pacientes, a pesar de que la estancia promedio del GRD es de 11,5 días. El horizonte temporal de la simulación corresponde a un trimestre y el paso de la simulación está dado en horas. La estructura del modelo se ilustra en la figura 2, en la cual solo se presenta el modelamiento de las tres primeras cohortes. Las cohortes subsiguientes a la cohorte tres poseen la misma estructura y secuencia de las cohortes dos y tres.

Figura 2. Diagrama de flujos y niveles de las cohortes de pacientes y ayudas diagnósticas e interconsultas



Fuente: elaboración propia.

Donde:
I_0: ingreso por hora de los pacientes.
D_0: nivel de pacientes que acaban de ingresar.
R_i: ronda médica del día i.
D_i: pacientes sin revisión médica o ronda en el día i.
D_{iE}: pacientes revisados (con ronda) en el día i.
P_i: transición de $D_{(i\text{-}1)E}$ a D_i, que simula la permanencia de un día de estancia en la institución.
P'_i: pacientes que no son dados de alta en el día i de estancia, por falta de ronda.
A_i: pacientes dados de alta en el día i de estancia.
$Dx_{i\text{-}1}$: nivel de ayudas e interconsultas disponibles en la ronda i.
A'_i: ayudas diagnósticas e interconsultas utilizadas en la ronda i.
P''_i: ayudas diagnósticas e interconsultas que llegan tarde y son enviadas al nivel de disponibilidad del día siguiente $Dx_{i\text{-}1}$.

En la figura 2, la fila inferior representa el flujo de las cohortes de pacientes en los días de estancia. En la fila superior se representa el flujo de ayudas diagnósticas e interconsultas disponibles para cada día de estancia. Todos los días, los pacientes del GRD estudiado entran al hospital. El ingreso por hora de los pacientes I_0 está representado por medio de una función GRAPH que devuelve el valor del número de pacientes por hora (ecuación 1).

$$I_0 = GRAPH\,(t, t_0, \Delta t, \{a_1, a_2, a_3, ..., a_n\}) \tag{1}$$

Los pacientes ingresan al nivel D_0, que corresponde al nivel de pacientes que acaban de ingresar a la institución y aún no cuentan con una ronda médica. Al nivel D_0 ingresan cada hora nuevos pacientes y salen los pacientes a los cuales se les hace la ronda médica (ecuación 2). La ronda médica del día cero R_0 permite que los pacientes pasen del nivel D_0 (sin revisión médica o ronda) a pacientes revisados en ese día D_{0E}. La ronda está modelada como un flujo que compara la cantidad de pacientes que están sin revisar en D_0 en el instante t, y la capacidad de revisión de pacientes por hora K. Esta ecuación de la ronda médica R_0 es la misma para representar el proceso de ronda en todas las cohortes de pacientes y, por ende, se generaliza mediante la expresión R_i en la ecuación 3. La duración del proceso de ronda médica está condicionada por la ecuación 4, $F_1(t)$, la cual se restringe a que este proceso de revisión solo esté activo durante 7 horas, comenzando a las siete de la mañana.

$$D_0\,(t) = \int_{t_0}^{t} (I_0(t) - R_0(t) - P'_1(t))\,dt + D_0(t_0) \tag{2}$$

$$R_i(t) = F_1(t)\ MIN(D_i(t), K) / \tau \qquad i = 0, 1, ..., 13 \text{ y } \tau = 1\ hora \quad (3)$$

$$F_1(t) = IF(6 < Hour(t) < 14, 1, 0) \quad (4)$$

Después de ser revisados los pacientes en la ronda médica, ingresan al nivel de pacientes evaluados D_{0E} (ecuación 4). Los pacientes permanecen en D_{0E} hasta la finalización del día y pasan, a través del flujo P_1, al nivel del día siguiente D_1. La ecuación de P_1 es la misma en todas las cohortes y, por lo tanto, esta también puede ser generalizada mediante la expresión P_i (ecuación 5). Por otro lado, los pacientes a los cuales no se les va a hacer la ronda médica por falta de capacidad en médicos o porque llegan después de la hora de ronda y están en cualquiera de los niveles D_i, también deben ser pasados al nivel del día siguiente de estancia cuando cambia el día (ecuación 7). Esta ecuación también es genérica y se representa como P'_i (ecuación 7). P_1 permite la transición de D_{0E} a D_1, la cual simula la permanencia de un día de estancia en la institución. D_1 se representa en la ecuación 8 como D_i, dado que esta expresión es la misma para los dieciséis niveles siguientes.

$$D_{0E}(t) = \int_{t_0}^{t} (R_0(t) - P_1(t))\ dt + D_{0E}(t_0) \quad (5)$$

$$P_i(t) = IF(Hour(t) = 0, 1, 0) * D_{iE}(t) / \tau \qquad i = 0, 1, ..., 13 \text{ y } \tau = 1\ hora \quad (6)$$

$$P'_i(t) = IF(Hour(t) = 0, 1, 0) * D_i(t) / \tau \qquad i = 1, ..., 13 \text{ y } \tau = 1\ hora \quad (7)$$

$$D_i(t) = \int_{t_0}^{t} (P_i(t) + P'_i(t) - R_i(t))\ dt + D_i(t_0) \qquad i = 0, 1, ..., 13 \text{ y } \tau = 1\ hora \quad (8)$$

Después del primer día de estancia se presenta la probabilidad de que el paciente sea dado de alta una vez es revisado en la ronda médica; por esta razón, todos los niveles desde el D_{1E} hasta el D_{16E} poseen la misma estructura y ecuaciones y son generalizadas mediante la definición de D_{iE} (ecuación 9). El cálculo de los pacientes dados de alta en el día i de estancia A_i (ecuación 10) depende de la cantidad de ayudas diagnósticas e interconsultas que fueron pedidas el día anterior y están disponibles para ser revisadas Dx_{i-1}, la cantidad de ayudas diagnósticas e interconsultas por persona en el día de estancia *i-1*, α_{i-1} y la probabilidad

π_i de ser dado de alta el día i, dado que las ayudas diagnósticas e interconsultas están disponibles.

$$D_{iE}(t) = \int_{t_0}^{t} (R_i(t) - A_i(t) - P_{i+1}(t))\, dt + D_0(t_0) \qquad i = 1, ..., 13 \tag{9}$$

$$A_i(t) = \frac{Dx_{i-1}}{\alpha_{i-1}\,\tau}\,\pi_i \qquad i = 1, ..., 13 \text{ y } \tau = 1 \textit{ hora} \text{ y } \sum_{i=0}^{13} \pi_i = 1 \tag{10}$$

El flujo de las ayudas diagnósticas e interconsultas modela la cantidad de ayudas e interconsultas solicitadas R'_i (ecuación 11), las cuales son un retraso material de orden O y tiempo promedio de entrega T del número de pacientes revisados en la ronda R_i, multiplicado por el promedio de ayudas diagnósticas e interconsultas por persona α_{i-1}. Una vez que las ayudas e interconsultas están disponibles, estas son almacenadas en un nivel de ayudas e interconsultas disponibles para ser consultadas en el día de la ronda Dx_{i-1} (ecuación 12). La notación del índice *i-1* alude a que la ayuda proviene del día anterior. El nivel de ayudas e interconsultas disponibles tiene dos flujos de salida: el primero corresponde al flujo de las ayudas utilizadas A'_i (ecuación 13), las cuales dejan de ser importantes cuando termina la ronda en términos logísticos y "salen" del sistema. El otro, P''_i, corresponde a las ayudas e interconsultas que llegan tarde y son enviadas al nivel de disponibilidad del día siguiente (ecuación 14).

$$R'_i(t) = Delaymtr(R_i(t)\ \alpha,\ T,\ O,\ 0) \qquad i = 0, 1, ..., 13 \tag{11}$$

$$Dx_{i-1}(t) = \int_{t_0}^{t} (R'_i(t) - A'_i(t) - P''_i(t))\, dt + Dx_{i-1}(t)\ (t_0) \qquad i = 0, 1, ..., 13 \tag{12}$$

$$A'_i(t) = IF(A_i(t) \neq 0,\ Dx_{i-1}(t),\ 0) / \tau \qquad Cont\ \tau = 1\ hora \tag{13}$$

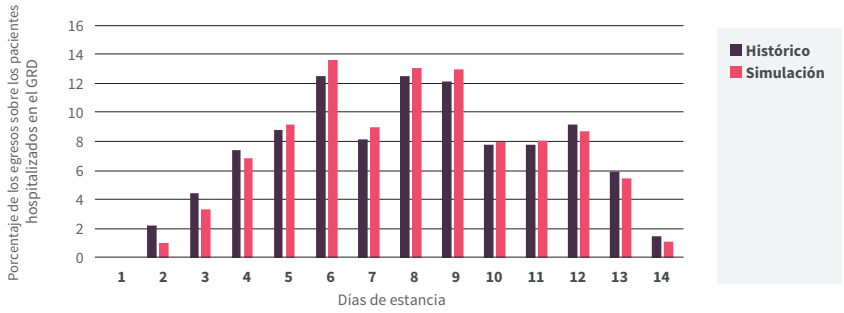
$$P''_i(t) = IF(Hour(t) = 0, 1, 0)\ Dx_{i-1}(t) / \tau \qquad Cont\ \tau = 1\ hora \tag{14}$$

Validación del modelo

En la fase de validación, se somete el modelo a una serie de pruebas y análisis para evaluar su validez y utilidad. En este caso, con la metodología establecida por Forrester y Senge (1980), un proceso incremental en el que se corroboraron la estructura, el comportamiento y las políticas. La estructura del modelo se validó

con base en la asesoría recibida del personal asistencial del hospital. Los datos que alimentan el modelo se tomaron de bases de datos que fueron depuradas y contrastadas directamente con la historia clínica, con el apoyo de un auditor médico. El comportamiento se validó realizando pruebas que evalúan el comportamiento del modelo y su capacidad de imitar el comportamiento real (Forrester y Senge, 1980). El promedio de ingresos de pacientes con el GRD de estudio es de 23 por mes, por lo que en el modelo se incluyeron 70 pacientes cuyo ingreso al sistema está dado por I_0 (ecuación 1) en un horizonte temporal de 3 meses. Para validar que el modelo reproduce el comportamiento previsto se tuvieron en cuenta dos variables: los egresos de pacientes con estancia de 1 a 14 días y la estancia promedio. Se comparó la similitud entre la proporción de egresos totales sobre el total de pacientes hospitalizados en cada día de estancia en el GRD del estudio con los simulados por el modelo (figura 3).

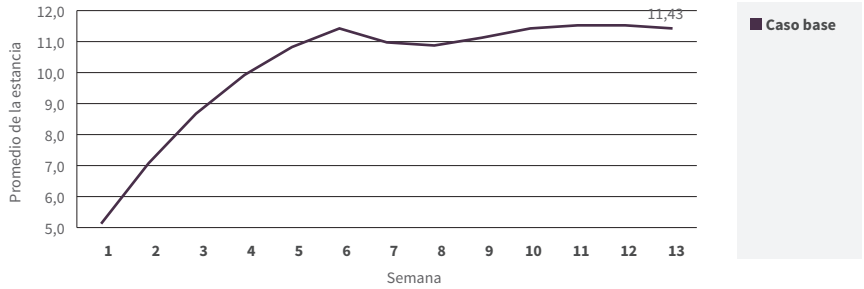
Figura 3. Porcentaje de egresos sobre el total de pacientes del GRD estudiado: histórico vs. simulado



Fuente: elaboración propia.

También se tomó como referente de validación la estancia promedio real. En la Clínica León XIII la estancia promedio de pacientes del GRD de estudio es de 11,2 días y el modelo arrojó un promedio de 11,43 días, lo cual presenta una diferencia del 2%. En la figura 4 se observa el comportamiento de la estancia promedio en el horizonte de tiempo simulado, indicando que esta tiende a acumularse hasta llegar a un nivel de estabilidad o decrecimiento, lo cual obedece a la dinámica de mayor demanda que oferta de cama, de recursos de especialistas y disponibilidad de ayudas diagnósticas.

Figura 4. Promedio de la estancia de pacientes en el escenario base



Fuente: elaboración propia.

Adicionalmente, en la tabla 2 se resumieron los estadísticos que evalúan el ajuste de los datos simulados a los reales. Para tal efecto, además de calcular el error cuadrático medio (RMS, por sus siglas en inglés) y el porcentaje del error cuadrático medio (RMSPE, por sus siglas en inglés), se calculó el estadístico de Theil (Sterman, 1984; Theil, 1966), el cual descompone el MSE para identificar si el error se debe al sesgo o diferencia de medias (UM), a diferencias en la variación o desviaciones estándar (US) o a la covarianza entre las series. Los resultados de la tabla 2 muestran el comportamiento esperado de un modelo que puede generar confianza, pues los errores que presentan están concentrados en US y UC (Sterman, 1984). Es decir, el modelo tiene la capacidad de reproducir el comportamiento promedio y tendencial de la realidad, pero con dificultades en la predicción exacta del valor. Este tipo de errores son comunes en la simulación bajo dinámica de sistemas y no son un factor determinante para rechazar los modelos cuando estos no están diseñados para predecir valores, sino para evaluar comportamientos (Sterman, 1984).

Tabla 2. Resumen estadístico para evaluar el ajuste de los resultados simulados a los datos históricos

Error cuadrático medio (RMS)	Porcentaje del error cuadrático medio (RMSPE)	Sesgo (UM)	Desviación estándar (US)	Covarianza (UC)
0,007	18	0	0,56	0,44

Fuente: elaboración propia.

Escenarios para el análisis

Con el fin de determinar el impacto de la respuesta de ayudas diagnósticas e interconsultas en la duración de la estancia hospitalaria, se plantearon varios escenarios que tuvieron en cuenta la variación en los tiempos de retraso de ayudas diagnósticas e interconsultas de especialistas. La Clínica León XIII ha acordado con los especialistas 24 horas como tiempo máximo de respuesta para las interconsultas. Para las ayudas diagnósticas, los estándares de tiempos dependen de lo pactado en la contratación con los proveedores y son acordados por tipo de estudio los tiempos máximos de respuesta según la tabla 3.

Tabla 3. Estándares de tiempos de ayudas diagnósticas

Ayuda diagnóstica	Estándar actual (horas)	Ayuda diagnóstica	Estándar actual (horas)
Rayos X	24	Urotac	24
Ecografía	24	Colonoscopia	36
Tomografía	24	Faringografía	12
Ecocardiograma	24	Espirometría	24
Dúplex	24	Ecocardiograma de estrés con prueba farmacológica	48
Endoscopia	36	Angiotomografía	12
Ultrasonografía	24	Colangiografía	36
Electroencefalograma	6	Arteriografía	36
Doppler	24	Cistografía	36
Holter	48	Ecocardiograma transesofágico	48

Fuente: elaboración propia.

Durante el 2013, para los pacientes atendidos con el GRD de infecciones de los riñones y el tracto urinario de alta severidad, se observó que el 20 % de las ayudas diagnósticas y el 40 % de las interconsultas con especialistas fueron inoportunas, es decir, no cumplieron con la promesa de servicio (estándares de tiempo) contratada por la Clínica León XIII con los proveedores de estos servicios. Según lo anterior, se plantearon escenarios que evalúan diferentes cambios sobre los estándares de tiempos de respuesta que ha definido la Clínica León XIII para las ayudas diagnósticas e interconsultas según el tipo de servicio. La variable respuesta que

identifica el desempeño del escenario propuesto es la estancia promedio de los pacientes en el GRD bajo estudio. Los escenarios que disminuyan la estancia con respecto al valor base de 11,43 días se consideran mejores que aquellos escenarios que generen valores mayores o iguales. Los escenarios analizados fueron:

Escenario 1: cumplimiento del estándar de tiempo de respuesta de las interconsultas médicas. En este primer escenario se observa el efecto en la duración de la estancia de pacientes si las interconsultas se realizaran en el tiempo establecido como promesa de servicio.

Escenario 2: cumplimiento del estándar de tiempo de respuesta de las ayudas diagnósticas. En este escenario se observa el efecto en la duración de la estancia de pacientes si las ayudas observadas como inoportunas se realizaran en el tiempo establecido.

Escenario 3: cumplimiento de los estándares de tiempo de respuesta de las ayudas diagnósticas e interconsultas médicas. Para este escenario se quiere observar el efecto en la duración de la estancia de pacientes si las ayudas diagnósticas e interconsultas que son inoportunas se realizaran en el tiempo establecido como promesa de servicio al igual que las interconsultas.

Escenario 4: reducción de los tiempos de respuestas de las ayudas diagnósticas más demandadas por los pacientes del GRD en estudio cuando las interconsultas cumplen la promesa de servicio (estándar de tiempo contratado). Para este escenario se explora el efecto en la duración de la estancia de pacientes, dado que algunas de las principales ayudas diagnósticas se realizan en un menor tiempo establecido al de la promesa de servicio actual (estándar de tiempo contratado). Las disminuciones como proporción de la actual promesa de servicio de las cuatro ayudas diagnósticas más demandadas son: rayos X (reducción de 24 horas a 6 horas), ecocardiograma transesofágico (reducción de 48 horas a 24 horas), Holter (reducción de 48 horas a 24 horas) y endoscopias (reducción de 36 horas a 24 horas).

Escenario 5: reducción de los tiempos de respuesta de las interconsultas de urología, nefrología e infectología de los pacientes de GRD bajo estudio y, además, las ayudas diagnósticas cumplen la promesa de servicio. Para este escenario se simula el efecto en la duración de la estancia de pacientes si algunas de las principales interconsultas se realizaran en un menor tiempo del establecido como promesa de servicio (estándar de tiempo contratado), y si las ayudas diagnósticas cumplieran la promesa de servicio. Se simula la reducción a 12 horas de las interconsultas de urología, nefrología e infectología, con lo cual se pretende que sean realizadas el mismo día en que son solicitadas durante la ronda. Las interconsultas a estas especialidades representaron el 34% de las interconsultas a los pacientes bajo el GRD estudiado.

Resultados

Los resultados de cada escenario resumidos en la tabla 4 muestran lo siguiente:

Escenario 1: cumplimiento del estándar de tiempo de respuesta de las interconsultas médicas. Para este escenario se observa que la estancia promedio de pacientes se reduce en 0,25 días respecto al histórico, lo cual equivale a una ganancia efectiva de 23 camas (Tell et al., 2010). Este escenario permite identificar que solamente el efecto de la inoportunidad de las interconsultas en la Clínica León XIII influye en la duración de la estancia.

Escenario 2: cumplimiento del estándar de tiempo de respuesta de las ayudas diagnósticas. La simulación sugiere que la estancia promedio de los pacientes no se reduce respecto al histórico. Este escenario permite identificar que, contrario a lo que se podría pensar, el impacto de las ayudas diagnósticas observadas como inoportunas en la Clínica León XIII no están prolongando la estancia hospitalaria.

Escenario 3: cumplimiento de los estándares de tiempo de respuesta de las ayudas diagnósticas e interconsultas médicas. La tabla 4 muestra que la estancia promedio de pacientes se reduce en 0,68 días respecto al histórico, lo cual equivale a una ganancia efectiva de 60 camas (Tell et al., 2010). Dado que esta política de cumplimiento conjunto en ayudas e interconsultas disminuye más el promedio de la estancia hospitalaria que el de una política aislada, se establece la necesidad de revaluar en los siguientes escenarios el impacto de la disminución de los tiempos de respuesta establecidos de las ayudas diagnósticas e interconsultas.

Escenario 4: reducción de los tiempos de respuestas de las ayudas diagnósticas más demandadas por los pacientes del GRD en estudio y, además, las interconsultas cumplen la promesa de servicio. El resultado en la tabla 4 indica que la estancia promedio de pacientes disminuye respecto al histórico en 0,71 días de estancia, equivalente a una ganancia efectiva de 63 camas (Tell et al., 2010).

Escenario 5: reducción de los tiempos de respuesta de las principales interconsultas (urología, nefrología e infectología) de los pacientes del GRD bajo estudio y, además, las ayudas diagnósticas cumplen la promesa de servicio. La tabla 4 muestra el resultado de este escenario. Se observa que la estancia promedio de pacientes merma respecto al histórico en 0,7 días de estancia, lo que equivale a una ganancia efectiva de aproximadamente 62 camas (Tell et al., 2010).

Tabla 4. Resumen del desempeño de los escenarios

Escenarios	Estancia promedio (días)	Cambio (en procentaje)	Reducción (días)	Ganancia efectiva (camas)
Escenario base: situación sin cambios en las condiciones iniciales	11,43	-	-	-
Escenario 1: cumplimiento del estándar de tiempo de respuesta de las interconsultas médicas	11,18	2,2	0,25	23
Escenario 2: cumplimiento del estándar de tiempo de respuesta de las ayudas diagnósticas	11,43	0	0	-
Escenario 3: cumplimiento de los estándares de tiempo de respuesta de las ayudas diagnósticas e interconsultas médicas	10,75	5,6	0,68	60
Escenario 4: reducción de los tiempos de respuestas de las ayudas diagnósticas más demandadas por los pacientes del GRD en estudio e interconsultas que cumplen la promesa de servicio	10,72	6,2	0,71	63
Escenario 5: reducción de los tiempos de respuesta de las interconsultas de urología, nefrología e infectología de los pacientes del GRD bajo estudio y, además, las ayudas diagnósticas que cumplen la promesa de servicio	10,73	6,1	0,7	62

Fuente: elaboración propia.

Análisis y discusión

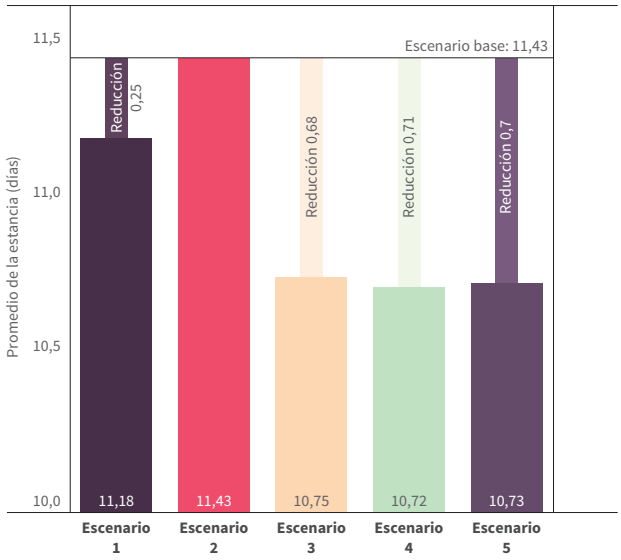
Con los resultados de las simulaciones, ilustrados en la figura 5, se establecieron comparaciones y se determinó el impacto de la inoportunidad y asincronía en las ayudas diagnósticas e interconsultas. Los dos primeros impactos son los de los escenarios 1 y 2, en los cuales la reducción del tiempo de inoportunidad en las ayudas diagnósticas no presentó ninguna mejoría, con respecto a la de disminuir el tiempo de inoportunidad de interconsultas, que sí las generó. Este resultado muestra la interdependencia de los dos procesos y sugiere el mayor impacto de la disponibilidad de interconsultas que el de las ayudas diagnósticas, pues no hay ganancia en disminuir el tiempo en las ayudas diagnósticas si no hay recurso humano para interpretarlas. Es la disminución en la oportunidad en interconsultas la que genera el impacto en este GRD.

El escenario 3, en el cual se reducen tanto tiempos de inoportunidad de las ayudas diagnósticas como de las interconsultas, confirma el hallazgo de la necesidad de realizar mejoras interdependientes. Al lograr una disminución conjunta en los tiempos, el indicador de estancia disminuye y sugiere que las políticas de mejora se deben presentar de manera conjunta. El escenario 3 plantea la estancia que debería tener el GRD, según la promesa de servicio del hospital, y muestra que el incumplimiento genera un impacto del 5,6 % adicional en los días de estancia (16,32 horas).

Los escenarios 4 y 5 muestran posibles alternativas de mejoras conjuntas, al reducir los tiempos de la promesa de servicio acordada con los proveedores y especialistas. Por un lado, (escenario 4), disminuyen los tiempos de las ayudas diagnósticas más solicitadas; por otro (escenario 5), disminuyen los tiempos de respuesta de las interconsultas. Ambas políticas incluyen la disponibilidad del recurso complementario, es decir, garantizan que la reducción en ayudas cuente con las interconsultas, y viceversa. Los resultados para ambas políticas son semejantes y sugieren mejoras en la reducción de la estancia en ese GRD de un 6 % (17 horas, aproximadamente).

Los resultados de los escenarios 3, 4 y 5 determinan una mejora en la estancia con un rango máximo de 43 minutos entre estos escenarios. Esto sugiere que al hospital le conviene, por costo-beneficio, acogerse a la política de cumplimiento de los estándares actuales de tiempo de respuesta de las ayudas diagnósticas e interconsultas médicas (escenario 3), puesto que los escenarios 4 y 5 implican un aumento en los costos de los servicios, ya que, al momento de negociar con los proveedores la reducción de los tiempos de respuesta, exigirán un aumento en las tarifas y, adicionalmente, se requieren más especialistas en interconsultas.

Figura 5. Comparación del desempeño de los escenarios



Fuente: elaboración propia.

Finalmente, si bien las mejoras permiten reducir la estancia del GRD bajo estudio, se encontró en los escenarios 3, 4 y 5 que estas no tienen un impacto considerable en la reducción de la estancia en ese GRD. Es decir, la simulación sugiere que no son los factores logísticos los principales y únicos determinantes de la estancia prolongada en ese GRD. El estudio no contempló factores como la situación sociofamiliar del paciente o la necesidad de atención en otra institución, entre otros factores que se presentan con frecuencia en el hospital por el tipo de población atendida. La razón para este alcance de la simulación es que se encontró pertinente iniciar el análisis de la estancia hospitalaria desde los procesos del hospital y directamente intervenibles por los administradores de este. Como los resultados del estudio muestran un impacto limitado en la reducción de la estancia y dado el contexto del hospital, conviene contemplar estrategias para abordar otras causas subsiguientes presentadas en la literatura, como lo son explorar estrategias con las aseguradoras e instituciones con fines sociales para la conformación de redes que faciliten la continuidad de la atención y alivios en las condiciones sociales del paciente.

Conclusiones

Los resultados encontrados confirman lo inicialmente presentado en la literatura, en la cual se establecía que la inoportunidad en ayudas e interconsultas es una de las causas de la estancia prolongada en hospitales (Kim et al., 2011; Zambrana García et al., 2001; Negro Álvarez et al., 2003; Silber et al., 2009). Sin embargo, en adición a lo establecido, se encuentra que las intervenciones conjuntas en la reducción de tiempos de oportunidad de ayudas diagnósticas e interconsultas sugieren caminos de mejora.

El modelo revela que hay una total interdependencia entre la necesidad de reducir ayudas diagnósticas e interconsultas de una manera sincronizada y conjunta para que se logren los beneficios de esta política. No obstante, el modelo sugiere que es más importante garantizar la oportunidad de las interconsultas que de las ayudas diagnósticas.

En el caso de estudio del GRD, infecciones del tracto urinario nivel tres, en la Clínica León XIII, la inoportunidad en las ayudas diagnósticas e interconsultas incrementa la estancia hospitalaria en un 5,6 %. El aumento de la estancia en este GRD puede atenuarse si se garantiza la oportunidad en las ayudas diagnósticas e interconsultas, sin incurrir en costos adicionales.

Por otro lado, el trabajo sugiere que el impacto de la sincronía en las ayudas diagnósticas e interconsultas en la unidad de medicina interna podría ser significativamente mayor en la reducción de la estancia total de la unidad de medicina interna si se aplicara la misma metodología a todos los GRD existentes.

Desde la perspectiva de procesos internos del hospital, este resultado plantea el trabajo futuro de identificar si la periodicidad de la ronda médica también puede contribuir en la disminución de la estancia hospitalaria. A su vez, sugiere explorar estrategias de relacionamiento entre entidades de la red de atención para intervenir otros factores causantes de prologar la estancia hospitalaria.

Agradecimientos

Se agradece de manera especial a todos los directores, coordinadores y auditores de la Clínica León XIII, quienes supervisaron, aconsejaron, validaron y apoyaron la realización del presente trabajo.

Referencias

Afanador, G. F. (2012, noviembre). *Atención segura y de calidad a través del manejo de la operación hospitalaria*. El Hospital. http://www.elhospital.com/temas/Atencion-segura-y-de-calidad-a-traves-del-manejo-de-la-operacion-hospitalaria+8090462

Ceballos-Acevedo, T. M., Velásquez-Restrepo, P. A. y Jaén-Posada, J. S. (2014). Duración de la estancia hospitalaria: metodologías para su intervención. *Gerencia & Políticas de Salud*, *13*(27). https://doi.org/10.11144/Javeriana.rgyps13-27.dehm

Correa Arcila, A. M. (2010, 18 de septiembre). Baja capacidad hospitalaria. *El Colombiano.* https://www.elcolombiano.com/historico/baja_capacidad_hospitalaria-LEEC_104825

Forrester, J. y Senge, P. (1980). Test for building confidence in system dynamics models. *TIMS Studies in the Management Sciences*, *14*, 209-228. https://www.albany.edu/faculty/gpr/PAD724/724WebArticles/ForresterSengeValidation.pdf

García Fariñas, A., Ramos Valle, I., García Rodríguez, J. F. y Gálvez González, A. M. (2011). El balance entre la oferta y la demanda en salud: el caso de los servicios de rehabilitación integral en La Habana, Cuba. 2009-2010. *Infodir, 7*(12). http://www.medigraphic.com/pdfs/infodir/ifd-2011/ifd111f.pdf

González-Angulo, I. J., Rivas-Gonzáles, C., Romero-Cedano, A. y Cortina-Gómez, B. (2009). Relación entre el prestador de servicio de salud y la estancia prolongada en el hospital. Relation between health services provider and the extended stay in the hospital. *Conamed*, *14*, 21-24.

Jacobson, S. H., Hall, S. N. y Swisher, J. R. (2006). Discrete-event simulation of health care systems. En R. W. Hall (Ed.), *Patient flow: Reducing delay in healthcare delivery. International Series in Operations Research & Management Science* (vol. 91, pp. 211-252). Springer. https://doi.org/10.1007/978-0-387-33636-7_8

Jaén, S. (2018). A goal-seeking system dynamics methodology for hospital bed capacity planning. En J. Figueroa-García, E. López-Santana y J. Rodríguez Molano (Eds.), *Applied Computer Sciences in Engineering. WEA 2018. Communications in Computer and Information Science* (pp. 280-291). Springer, Cham. https://doi.org/10.1007/978-3-030-00350-0_24

Jiménez Paneque, R. E. (2004). Indicadores de calidad y eficiencia de los servicios hospitalarios: una mirada actual. *Revista Cubana de Salud Pública*, *30*(1), 17-36. http://www.redalyc.org/articulo.oa?id=21430104

Kim, C. S., Hart, A. L., Paretti, R. F., Kuhn, L., Dowling, A. E., Benkeser, J. L. y Spahlinger, D. A. (2011). Excess hospitalization days in an academic medical center: Perceptions of hospitalists and discharge planners. *The American Journal of Managed Care*, *17*(2), 34-42. http://www.ncbi.nlm.nih.gov/pubmed/21473658

Negro Álvarez, J. M., Jiménez Molina, J. L., Félix Toledo, R., Pacual Camús, A., Miralles López, J. C., López Sánchez, J. D., Ferrándiz Gomis, R., Guerrero Fer-

nánez, M. y Hernández García, J. (2003). Adecuación de ingresos y estancias hospitalarias en pacientes con asma bronquial. Mejora de la gestión de las camas en la sección de alergología, tras la difusión de un estudio previo. *Alergología e Inmunología Clínica*, *18*(2), 61-70. http://revista.seaic.org/abril2003/61_original.pdf

Panis, L. J. G. G., Gooskens, M., Verheggen, F. W., Pop, P. y Prins, M. H. (2003). Predictors of inappropriate hospital stay: A clinical case study. *International Journal for Quality in Health Care: Journal of the International Society for Quality in Health Care*, *15*(1), 57-65. https://doi.org/10.1093/intqhc/15.1.57

Rashwan, W., Abo-Hamad, W. y Arisha, A. (2015). A system dynamics view of the acute bed blockage problem in the Irish healthcare system. *European Journal of Operational Research*, *247*(1), 276-293. https://doi.org/10.1016/j.ejor.2015.05.043

Silber, J. H., Rosenbaum, P. R., Rosen, A. K., Romano, P. S., Itani, K. M. F., Cen, L., Mi, L., Halenar, M. J., Even-Shoshan, O. y Volpp, K. G. (2009). Prolonged hospital stay and the resident duty hour rules of 2003. *Medical Care*, *47*(12), 1191-1200. https://doi.org/10.1097/MLR.0b013e3181adcbff

Sterman, J. D. (1984). Statical proffs for models. *Dynamica*, *10*(II), 51-66.

Sterman, J. D. (2006). *Business dynamics: System thinking and modeling for a complex world.* McGraw-Hill.

Tamargo Barbeito, T. O., Jiménez Paneque, R. E., Gutiérrez Rojas, Á. y Díaz Mora, I. (2012). Dos procedimientos de ajuste por riesgo para la estadía hospitalaria como indicador de desempeño. *Revista Cubana de Salud Pública*, *38*(1), 29-44. https://www.medigraphic.com/pdfs/revcubsalpub/csp-2012/csp121d.pdf

Tell, W., Nataraja, S., Keren, J., Fontana, E., Kennedy, E., Wyche, M. y Calvert, S. (2010). *Next generation capacity management: Collaborating for clinical appropriate and efficient inpatient throughput*. Clinical operations Board. Advisory Board International.

Theil, H. (1966). *Applied economics forecasting*. North Holland Publishing Company.

Vanderby, S. y Carter, M. W. (2009). An evaluation of the applicability of system dynamics to patient flow modelling. *Journal of the Operational Research Society*, *61*(11), 1572-1581. https://doi.org/10.1057/jors.2009.150

Zambrana García, J. L., Delgado Fernández, M., Cruz Caparrós, G., Martín Escalante, M. D., Díez García, F. y Ruiz Bailén, M. (2001). Factores predictivos de estancias no adecuadas en un servicio de medicina interna. *Medicina Clínica*, *117*(3), 90-92. https://doi.org/10.1016/S0025-7753(01)72026-X

Capítulo 6.
Árboles de modelos aditivos generalizados de localización, escala y forma para la predicción del tiempo de estancia hospitalaria

Juan Camilo España Lopera,[1]
Olga Cecilia Úsuga Manco[2] *y*
Juan Sebastián Jaén-Posada[3]

[1] Departamento de Ingeniería Industrial
Universidad de Antioquia
Medellín, Colombia
juan.espana@udea.edu.co

[2] Departamento de Ingeniería Industrial
Universidad de Antioquia
Medellín, Colombia
olga.usuga@udea.edu.co

[3] Departamento de Ingeniería Industrial
Universidad de Antioquia
Medellín, Colombia
juan.jaen@udea.edu.co

Resumen

El tiempo de estancia hospitalario, conocido en la literatura como *length of stay* (LOS), es una variable fundamental en la administración hospitalaria, debido a su impacto en la eficiencia de los recursos. Por eso, la predicción ha sido un tema de interés de los investigadores durante las últimas décadas. Aunque se han utilizado diferentes técnicas para la predicción de LOS, sigue existiendo una falencia en la exploración de técnicas que tengan en cuenta las características que tiene esta variable. En el presente trabajo se desarrolló la metodología de árboles de modelos GAMLSS (AMG), en la cual se utilizan los modelos aditivos generalizados de localización, forma y escala (GAMLSS), junto con la técnica de árboles de clasificación y regresión (CART), para predecir el LOS de cada paciente. La metodología se aplica en tres casos de estudio de tres hospitales con diferentes características. Dentro de los principales logros de la investigación se encuentran: el diseño e implementación en *software* de una metodología que facilita la generación y selección de modelos de predicción de variables continuas, la inclusión del

criterio de información de Akaike (AIC) como métrica para comparar modelos, la mejora encontrada en algunas métricas utilizadas con respecto a otras técnicas.
Palabras clave: estancia hospitalaria, modelos de predicción, modelos GAMLSS, árboles de decisión

Introducción

Continuamente, los hospitales se ven obligados a diseñar estrategias que les permitan: ser más eficientes con el uso de sus recursos (Weir et al., 2009); ser más efectivos con los servicios brindados, debido al aumento en la presión por parte de entidades regulatorias y usuarios cada vez más exigentes (Liu, 2013), y contar con estrategias para mejorar la experiencia del cliente, dado que es una ventaja competitiva y un factor de éxito crítico (Vaish et al., 2016).

La eficiencia en el uso de los recursos es uno de los objetivos más importantes en la literatura de administración hospitalaria (Al Taleb et al., 2017; Barnes et al., 2016; Nouaouri et al., 2015; Turgeman et al., 2017). A efectos de alcanzar este objetivo, los hospitales utilizan diversas estrategias, entre ellas programar de manera adecuada las cirugías y otros servicios prorrogables que generarán la hospitalización (Azari et al., 2015; Barnes et al., 2016; Gustafson, 1968). Este problema lo han abordado diferentes estudios por medio de la predicción del LOS (Al Taleb et al., 2017; Barnes et al., 2016; Nouaouri et al., 2015; Turgeman et al., 2017). Otra estrategia que se plantea continuamente dentro de los hospitales es la disminución del LOS, debido al impacto que tiene en los costos de atención de los pacientes (Awad et al., 2017; Liu et al., 2010; Widyastuti et al., 2012). Para lograr esta disminución es fundamental identificar las causas y las características de las personas que influyen tanto positiva como negativamente en el LOS y, de esta manera, facilitar las acciones encaminadas a disminuirlo (Liu et al., 2010; Turgeman et al., 2017; Widyastuti et al., 2012).

El LOS también se ha empleado con el propósito de mejorar la efectividad de los servicios médicos en la generación de alertas sobre posibles complicaciones o cambios en las condiciones del paciente, que permitan enfocarse en la atención prioritaria de algunos (Awad et al., 2017; Azari et al., 2012), ya que, cuando el tiempo real de un paciente ha superado el estimado con una diferencia significativa, esto se debe a que el paciente sufrió alguna complicación o existen ineficiencias en la atención prestada (Barnes et al., 2016).

Otro objetivo de las instituciones de salud para lograr el éxito es mejorar la satisfacción de sus usuarios (Vaish et al., 2016). Algunos autores coinciden en que el tiempo de espera desempeña un rol fundamental en dicha satisfacción, no solo por la cantidad de tiempo de espera, sino por el hecho de ser informado de dicho tiempo (Carter y Potts, 2014). Se afirma que la percepción del tiempo de espera y la información sobre este influye en la satisfacción de los usuarios (Fontova-Al-

mató et al., 2015). Otros estudios también han encontrado hallazgos respecto a los efectos negativos que tiene la incertidumbre en el contexto de los pacientes hospitalizados, ya que es un factor correlacionado positivamente con la ansiedad (Pahlevan Sharif, 2017). Esto evidencia cuán fundamental es contar con una predicción acertada del LOS para poder informar a los pacientes y lograr un avance en su satisfacción con la atención prestada.

En los objetivos mencionados y las estrategias para lograrlo se refleja la importancia de la predicción del LOS y por qué esta variable ha sido el interés de muchos estudios en literatura relacionada con la administración en salud (Azari et al., 2012; Barnes et al., 2016; Cai et al., 2016; Nouaouri et al., 2015; Pérez et al., 2006; Pilotto et al., 2016; Widyastuti et al., 2012). El LOS se calcula restando la fecha de egreso de la fecha registrada de ingreso al hospital, midiendo un episodio individual de hospitalización (Turgeman et al., 2017). La mayor parte de los estudios se enfocan en predecir el tiempo en el momento del ingreso, con la información disponible en ese punto (Azari et al., 2012; Pérez et al., 2006; Turgeman et al., 2017); sin embargo, también existen trabajos que han estudiado el problema de predecir el LOS, no solo al inicio de la estancia, sino actualizando esta predicción a medida que se registran eventos que permitan ajustarla (Barnes et al., 2016; Cai et al., 2016).

El problema de predecir el LOS no se ha investigado únicamente por su impacto en la administración hospitalaria, sino por el interés que despierta su modelamiento y predicción en el contexto académico, debido a sus características y complejidad (Turgeman et al., 2017). Una de las características que dificulta el modelamiento de esta variable es que su distribución es altamente asimétrica, por lo cual limita el uso de algunos modelos a causa de sus supuestos, por ejemplo, los modelos de regresión lineal, foco de críticas de algunos investigadores (Carter y Potts, 2014; Faddy et al., 2009). Otra característica es que existen muchas covaribles que pueden influir en el LOS, razón por la cual las técnicas utilizadas deben tener componentes claros de selección de variables. Estas razones han conseguido que la literatura relacionada con la predicción del LOS se centre en la exploración y comparación de técnicas (Tanuja et al., 2011; Turgeman et al., 2017; Verburg et al., 2014). Se identifican dos oportunidades en esta área de estudio: la primera es la inexistencia de acuerdos en la literatura relacionados con la idoneidad de alguna de las técnicas exploradas, que hacen que el diseño de nuevas técnicas sea un área de vigente interés. La segunda oportunidad se presenta con relación a la selección de los modelos que una técnica puede generar, ya que este no ha sido un tema que se haya abordado con profundidad.

Otro aspecto importante en la predicción del LOS es la dificultad para generalizar o extrapolar los modelos, debido a que puede cambiar según las características del país, ciudad o región. En primer lugar, por la información disponible, ya que cada hospital puede llevar registros diferentes, de acuerdo con sus nece-

sidades y regulaciones. En segundo lugar, por las características de la población, ya que la influencia de una variable puede ser diferente en poblaciones distintas. En tercer lugar, por la existencia de regulaciones capaces de influir en el LOS y que son difíciles de detectar y controlar en este tipo de estudios, por ejemplo, regulaciones que le exijan a un hospital cumplir con unos parámetros mínimos de atención o cumplir ciertas condiciones para autorizar el egreso del paciente. Dado esto, algunos autores mencionan la necesidad de que las investigaciones hagan hincapié en la metodología con la que se obtiene el modelo para que sea replicable en diferentes contextos (Awad et al., 2017; Pérez-Castro et al., 2017).

En diferentes investigaciones se han encontrado coincidencias en relación con ciertas características que describen el LOS. La primera es que solo toma valores positivos, y en algunas investigaciones se toma como una variable continua, al tener en cuenta la hora de ingreso y de egreso (Faddy et al., 2009;Liu et al., 2010); en otros casos se toma como un conteo, pero se omite la hora (Carter y Potts, 2014; Verburg et al., 2014). La segunda característica es que es una variable en la que hay presencia de valores extremos correspondientes a casos con estancias muy largas, aunque el promedio se mantenga en un rango relativamente pequeño. La tercera característica es el sesgo a la derecha, que se debe a la presencia de valores extremos y a la concentración de la variable en un rango pequeño, lo cual genera una distribución leptocúrtica. Algunos autores han analizado el ajuste del LOS a diferentes distribuciones, dentro de las que se encuentran la distribución exponencial (Awad et al., 2017), log-normal, gamma, Weibull (Liu et al., 2010; Marazzi et al., 1998); adicionalmente, se ha estudiado el ajuste a distribuciones discretas como la binomial negativa y la Poisson (Carter y Potts, 2014; Verburg et al., 2014). Por otro lado, se ha criticado el ajuste de modelos de regresión lineal, debido a que la estancia no se acopla a una distribución normal (Carter y Potts, 2014; Faddy et al., 2009; Turgeman et al., 2015).

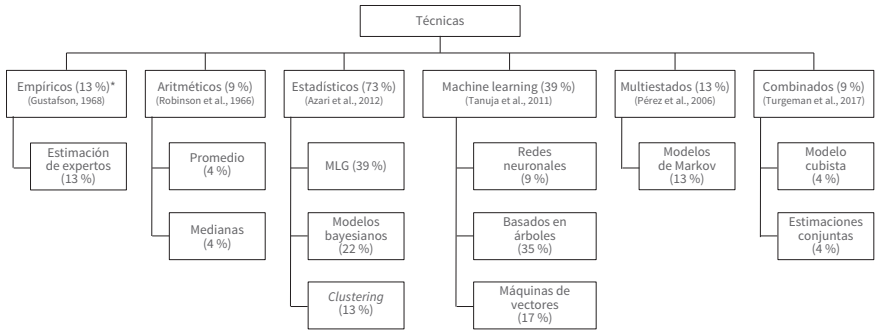
Existen dos enfoques con los que se ha estudiado el LOS: el primero busca entender la asociación entre el LOS y otras variables y el segundo se centra en modelarlo y predecirlo con base en unas variables predictoras. Los primeros son de tipo causal y fueron el enfoque inicial en su estudio. Algunos trabajos iniciales datan de los años setenta, cuando se documentaron variaciones en el LOS causadas por diferencias geográficas (Burns y Wholey, 1991). Posteriores trabajos, como el de Liu et al. (2010), también han dirigido sus esfuerzos hacia la identificación de variables que permitan explicar adecuadamente el LOS. Dentro de las variables que se han encontrado como significativas están la severidad y los factores relacionados con las instalaciones del hospital y de la comunidad que rodea al paciente, las cuales desempeñan un rol fundamental en las variaciones, aunque las relaciones causales no hayan sido comprendidas en su totalidad (Burns y Wholey, 1991). Con frecuencia, también se han incluido la comorbilidad y el diagnóstico

o enfermedad principal del paciente, como factores críticos para explicar el LOS (Azari et al., 2012; Liu et al., 2010).

El segundo enfoque, más común en la literatura reciente, es la predicción y modelamiento del LOS (Al Taleb et al., 2017; Azari et al., 2012; Faddy et al., 2009; Tanuja et al., 2011). Estas investigaciones son de tipo correlacional y pueden realizarse con dos objetivos: el primero, modelar la variable LOS como respuesta, explicada por un conjunto de variables predictoras; en este caso se estudia el tipo de relación que tiene cada una de las covariables con las variables respuesta y el grado de asociación entre ellas (Burns y Wholey, 1991; Marazzi et al., 1998). Cuando se realizan estudios con este objetivo, se presentan muchas similitudes con el enfoque anterior, ya que al final lo que se quiere es entender las variables que están impactando el crecimiento o decrecimiento del LOS, principalmente para identificar variables explicativas que pueden ser intervenidas a efectos de disminuirlo. Un ejemplo de este tipo de variables es la cantidad de pacientes por cada enfermera disponible, ya que se pueden modificar por decisiones en la administración hospitalaria. Sin embargo, muchos modelos tienen el objetivo principal de predecir acertadamente el LOS, sacrificando incluso la capacidad de entender las relaciones entre variables explicativas y variables respuesta para alcanzar una mayor exactitud en la predicción. Esto se presenta más en las técnicas computacionales, que se conocen como técnicas de tipo caja negra (Dayhoff y DeLeo, 2001; Mobley et al., 1995).

Con el fin de predecir el LOS se han utilizado técnicas de diferentes campos de estudio que han generado varias perspectivas en el abordaje del mismo problema. Para tener una estructura con la cual entender estas perspectivas y la evolución de esta área de estudio se han propuesto clasificaciones de técnicas, una de ellas es la propuesta por Awad et al. (2017). En esta clasificación se identificaron cuatro grupos que, los autores consideran, resumen de modo adecuado las diferentes aproximaciones que se encuentran en la literatura. Con base en esta clasificación y de acuerdo con la revisión de literatura hecha, se propone una clasificación propia para agrupar las técnicas de predicción del LOS. En la propuesta se incluyen dos nuevas categorías: los métodos empíricos y las técnicas combinadas. La figura 1 clasifica los grupos de técnicas encontradas en la revisión. Adicionalmente, se presenta el porcentaje de artículos en los cuales se utiliza cada técnica, con respecto al número total de artículos de la revisión de la literatura en el que se encontraron aplicaciones; esto con el objetivo de analizar su frecuencia de uso.

Figura 1. Clasificación de técnicas de predicción del LOS propuesta



Fuente: elaboración propia.

Adicionalmente a la identificación de estos grupos de técnicas, también pueden diferenciarse las técnicas en el tipo de variable respuestas que utilizan, algunas tienen en cuenta las categóricas y son conocidas como técnicas de clasificación. En este tipo de métodos se agrupan los días de estancia en rangos y se predice el rango en el que queda cada paciente (Mobley et al., 1995; Tanuja et al., 2011). Por otro lado, están las técnicas de regresión, que estiman el LOS para variables respuestas continuas, y la predicción, que, a diferencia de los métodos de clasificación, no pertenece a un grupo, sino que predice el valor exacto que va tomar la observación de acuerdo con las covariables (Marazzi et al., 1998; Turgeman et al., 2017). A continuación, se presentan algunas aplicaciones que se han encontrado de los diferentes tipos de técnicas mencionados.

En el caso de las técnicas de clasificación se encuentran varias aplicaciones. Liu et al. (2006), por ejemplo, predijeron el LOS en un hospital geriátrico aplicando un clasificador ingenuo de Bayes, apoyado en un árbol de clasificación C4.5 para la selección de las variables más importantes; adicionalmente, con el fin de mejorar el desempeño se utiliza una imputación ingenua de Bayes para datos faltantes (Liu et al., 2006). En el trabajo de Tanuja et al. (2011) se compararon modelos de clasificación con el propósito de predecir el LOS en un hospital especializado. Se emplearon cuatro modelos: redes neuronales, clasificador de Bayes ingenuo, modelo de K vecino más cercano y un árbol de decisión C4.5. Se encontró que la red neuronal es la de mejor desempeño (Tanuja et al., 2011). En el trabajo de Al Taleb et al. (2017), en un hospital de Arabia Saudita en pacientes con problemas cerebrovasculares, se usó un algoritmo J48 de árbol de clasificación que está basado en un árbol C4.5 y una red bayesiana, métodos que se han utilizado en otras aplicaciones médicas (Kaur y Chhabra, 2014; Kumari et al., 2014), y la red bayesiana tuvo un mejor desempeño (Al Taleb et al., 2017).

Para el caso de análisis de variables continuas, Marazzi et al. (1998) propusieron ajustar el LOS a tres distribuciones de probabilidad: Weibull, gamma y log-normal. En este estudio se ajustaron estas distribuciones a 3260 muestras, en las que se divide la población que, en principio, estaba formada por aproximadamente cinco millones de hospitalizaciones y se identificó que la distribución log-normal es la que se acomoda en la mayoría de las muestras. En este trabajo también se tuvo la limitación de no tener variables que explicaran el LOS, lo que reduce la capacidad de las instituciones de implementar estrategias para mejorar la eficiencia (Marazzi et al., 1998; Turgeman et al., 2017; Verburg et al., 2014). Otros trabajos abordan estas limitaciones con técnicas de regresión que tienen en cuenta las variables predictoras. Tal es el caso de la investigación de Carter y Potts (2014), quienes, utilizando una regresión lineal, analizaron la estancia hospitalaria de pacientes con operación total de rodilla en un hospital de servicios ortopédicos especializados. También se encuentra el trabajo propuesto por Verburg et al. (2014), en el que se compararon ocho modelos de regresión diferentes: regresión lineal sin transformaciones, regresión lineal sin transformaciones eliminando los LOS mayores a 30 días, regresión lineal con transformación logarítmica, modelo lineal generalizado con distribución gaussiana y una función de enlace logarítmica, regresión de Poisson, regresión binomial negativa, regresión gamma con función de enlace logarítmica y regresión de riesgo proporcional de Cox. En este estudio se encontró que el modelo de Cox y la regresión lineal con transformación logística tenían mejor desempeño para estancias inferiores a cuatro días, y en los otros casos el modelo lineal generalizado con función de enlace logarítmica tuvo mejor desempeño, aunque la diferencia fue poca. También se concluyó que los modelos planteados no permiten predecir de forma adecuada el LOS (Turgeman et al., 2017; Verburg et al., 2014).

Dentro de la familia de las regresiones también se han explorado métodos más sofisticados, comúnmente usados en el campo del *machine learning* y minería de datos. Pendharkar y Khurana (2014) analizaron y compararon tres métodos de regresión con los datos de 88 hospitales de Pensilvania, árboles CART, árboles CHAID y regresión de vector de soporte, y encontraron que aun cuando no había diferencias significativas, el método CART es el más fácil de interpretar (Pendharkar y Khurana, 2014). En el trabajo de Houthooft et al. (2015) también se compararon varios métodos de regresión con los datos de un hospital universitario en su departamento de cuidados intensivos. Los métodos utilizados fueron redes neuronales, regresión de vector de soporte, árbol CART, bosques aleatorios y regresión de vector de relevancia, y se halló que el método de mejor desempeño fue la regresión de vector de soporte.

Debido a las dificultades de encontrar un modelo que prediga apropiadamente el fenómeno del LOS, algunos autores han explorado métodos combinados. Por ejemplo, Turgeman et al. (2017) aplicaron uno de árbol cubista, que aun

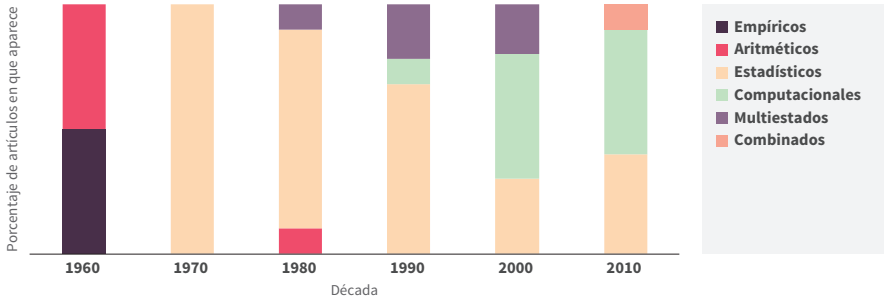
cuando no es nuevo y ya había sido propuesto con anterioridad (Quinlan, 1992), es menos común en la literatura de aplicación a la predicción del LOS. Permite estimar el valor de la estancia para cada paciente y no un rango de tiempo como los modelos de clasificación. Se empleó en un hospital de adultos mayores con problemas cardiovasculares. Este modelo obtuvo mejor desempeño que los otros cinco con los que se comparó (árbol CART, árbol CHAID, máquina de vector de soporte, redes neuronales o modelo lineal generalizado con función de enlace Poisson).

Los trabajos de predicción del LOS han evolucionado desde 1966, cuando se presentó uno de los trabajos seminales en el área (Robinson et al., 1966), por parte de investigadores del campo de la salud. En este estudio se realizaron análisis exploratorios para identificar estrategias que disminuyeran la incertidumbre con respecto al promedio del LOS, que era la métrica más usada. Los métodos que plantea se basan en una clasificación de pacientes por enfermedad o por variables demográficas como el sexo y la edad mediante reglas fijas. Posteriormente a esta clasificación, se compararon las diferencias entre los grupos formados y, verificando si existía una disminución en su variabilidad, también se realizó una predicción por parte de fisiatras y otra por parte de las enfermeras (Robinson et al., 1966). Gustafson (1968) fue uno de los primeros autores que se sirvieron de modelos estadísticos para predecir el LOS. En su estudio comparó cinco métodos, dentro de los cuales están la estimación por puntos basada en el criterio de expertos, los modelos de regresión lineal múltiple, un método basado en el histórico de medias, un método basado en *odds* de radio y un modelo subjetivo bayesiano (Gustafson, 1968). En el trabajo de Burns y Wholey (1991) también se analizó el LOS buscando las variables que mayor efecto tienen sobre el LOS sin plantear modelos estadísticos para predicción, pero generando resultados importantes para modelos posteriores que debieron seleccionar las variables con mayor efecto en el LOS (Burns y Wholey, 1991).

En 1995 comenzaron a aparecer aplicaciones computacionales que se estaban desarrollando en la época y empezaban a volverse más frecuentes. Estas metodologías se irían popularizando en las aplicaciones médicas, generando un campo de investigación relacionado con la computación y la estadística (Mobley et al., 1995). A finales de la década de 1990 también se aplicó uno de los métodos más utilizados en el campo de la predicción del LOS: las distribuciones tipo fase y modelos de Markov (Jiménez et al., 1999; Marshall et al., 2005). Después del 2000 empezaron a aparecer modelos aplicados desde un campo de investigación de minería de datos (Liu et al., 2006; Silva et al., 2008; Tanuja et al., 2011). Finalmente, en estudios más recientes, las mismas técnicas de minería de datos y algunas mejoras de estas dejaron de reportarse en la literatura de minería de datos y se comenzaron a asociar con el campo del *machine learning* (Awad et al., 2017; Azari et al., 2015; Turgeman et al., 2017). En la figura 2 se muestra cómo ha cam-

biado, a través de las décadas, la participación de los grupos de técnicas que se identificaron en el capítulo anterior. Se observa que los métodos computacionales surgieron en las últimas tres décadas y, aun así, como resultado de la popularidad que han ganado diferentes corrientes de este enfoque, se han vuelto los más usados en la actualidad. Por otro lado, el enfoque estadístico fue el que más se usó en las décadas de 1970 y 1980; sin embargo, fue perdiendo participación frente a las técnicas computacionales, aun cuando en la actualidad sigue estando presente en muchas investigaciones.

Figura 2. Evolución histórica de técnicas de predicción del LOS



Fuente: elaboración propia.

Dentro de la amplia gama de técnicas encontradas, es necesario resaltar la cubista, ya que tiene la bondad de aprovechar las ventajas de otras. Se basa en reglas que combinan diferentes metodologías publicadas anteriormente y que han ido evolucionando con el paso de los años (Kuhn y Johnson, 2013). La columna vertebral del modelo cubista se sustenta en la técnica de árboles de modelos, descrita por Quinlan (1992) en su algoritmo M5, donde se define un árbol de decisión y una regresión lineal en cada nodo. Posteriormente, este sufre algunos cambios para dar origen a la técnica cubista. Esta puede analizar gran cantidad de variables, que es una ventaja de los árboles de decisión, junto con la posibilidad de separar el conjunto total de observaciones en subgrupos con medias diferentes. Adicionalmente, puede aprovechar la precisión de algunos modelos de regresión basados en distribuciones paramétricas.

En las técnicas encontradas en la literatura existen muchos avances y algunas oportunidades. Se resalta la exploración de una gran variedad de técnicas, su combinación para utilizar las ventajas que cada una tiene y la mejora en cuanto a la disminución de residuales del modelo generado por la técnica cubista, que se comparó con las técnicas más competitivas de la literatura (Turgeman et al., 2017). Por otro lado, se identifican oportunidades en la técnica cubista, en cuanto a que puede incluir en cada nodo final modelos diferentes a la regresión lineal

múltiple; también se identifican oportunidades, respecto a la medición y comparación de los modelos con criterios como el AIC y el criterio de información bayesiano (BIC), que incluyen bondad de ajuste y penalización por complejidad, ya que muchos estudios únicamente usan métricas enfocadas en los residuales (Al Taleb et al., 2017; Barnes et al., 2016; Turgeman et al., 2017). También existe una oportunidad en el desarrollo de metodologías, que incluyan procedimientos no considerados y que faciliten el proceso de selección del modelo final. Por último, se identifica la oportunidad de que estas metodologías para el desarrollo de modelos hagan parte de componentes de *software* que simplifiquen la generación de modelos de predicción del LOS.

El objetivo general de la investigación fue diseñar una metodología con la cual generar y seleccionar un modelo de predicción del LOS, que recoja aprendizajes de estudios previos y que tenga un desempeño competitivo con otras técnicas, desde el punto de vista de bondad de ajuste y de residuales del modelo seleccionado.

En las secciones siguientes se presenta una breve revisión de la literatura sobre los métodos utilizados y se describen los métodos desarrollados por este trabajo para abordar el problema. Luego, se abordan los resultados de los casos de estudio para los que se aplicaron las técnicas, en que se compara con otras técnicas existentes. Después, se presentan los análisis y discusiones de los resultados obtenidos. Y se cierra con las conclusiones de la investigación.

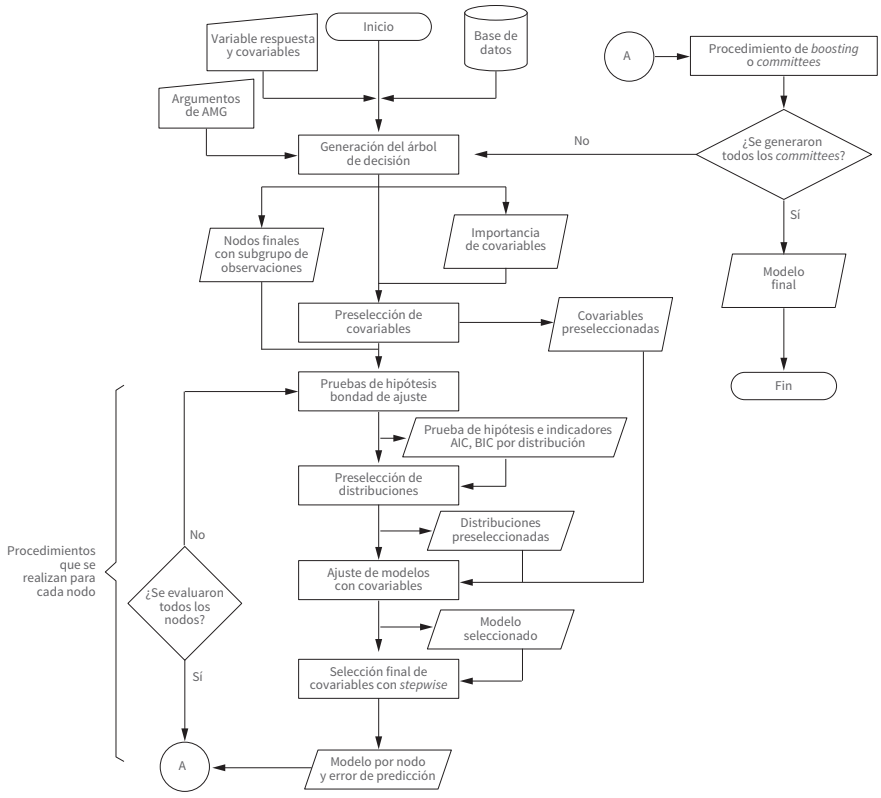
Métodos

En la revisión de la literatura se identificaron las características del LOS, las técnicas de predicción desde sus diferentes campos de estudio, su evolución, algunos criterios para seleccionar modelos y algunos aprendizajes y oportunidades. Con base en estos, se desarrolló la metodología árboles de modelos GAMLSS (AMG). Así se aprovechan ventajas encontradas con la técnica cubista y se define una metodología que incluye procedimientos que generalmente no incluyen las técnicas, como pruebas de hipótesis para conocer la naturaleza de la distribución del LOS, selección de distribución, selección de variables y selección de modelo final. Con este último se incluyen criterios que tienen en cuenta la naturaleza de la distribución del LOS y que penalizan modelos complejos, pero también se dejan criterios basados únicamente en el comportamiento de los errores encontrados en las predicciones. Por último, se implementa la metodología en un *software* estadístico con diferentes argumentos que permiten al usuario desarrollar un modelo utilizando los criterios que escoja y permitiéndole incluir o no ciertos pasos de la metodología propuesta. De este modo, con tal metodología se logran cubrir los principales retos encontrados en la literatura.

Las técnicas computacionales son las más comunes en la literatura reciente. Por lo general, no tienen supuestos de distribuciones de la variable respuesta como las técnicas estadísticas; esto hace que algunas se puedan aplicar desde el enfoque computacional y que desde el estadístico no se apliquen o que se deban hacer ajustes como transformaciones de potencia (Box y Cox, 1964). La técnica cubista proviene de un enfoque computacional, pero incorpora la regresión lineal, que es una técnica estadística. En dicha regresión no se analiza ni se brinda información sobre los supuestos. En la metodología AMG se agregan características como las pruebas de hipótesis de bondad de ajuste, los criterios AIC y BIC y un amplio conjunto de distribuciones paramétricas que permiten a los investigadores tener más información para analizar y seleccionar el modelo final sin perder la flexibilidad que tiene la técnica cubista. Además, se exploran algunas variaciones en los procedimientos utilizados por dicha técnica para analizar su impacto sobre el desempeño del modelo desde el punto de vista del análisis de los residuales y de la bondad de ajuste del modelo. Son precisamente estas características adicionales las que diferencian a la metodología AMG de otras como la técnica cubista o el algoritmo M5, ya que estos últimos solo ajustan una distribución normal y en muchas ocasiones las variables respuesta, como es el caso del LOS, pues no tienen las características de dicha distribución, incluso con transformaciones.

En la figura 3 se presenta un diagrama de flujo de la metodología AMG propuesta. En este se definen las actividades que se realizan en la metodología; las entradas manuales que requiere cada actividad, que en la implementación de *software* son argumentos definidos por el usuario; el resultado de cada actividad y otras actividades que utilizan cada uno de estos resultados, y los puntos de decisión del algoritmo que establecen los bucles que debe realizar el algoritmo repitiendo una serie de actividades para diferentes nodos o *committees* (Kuhn y Johnson, 2013).

Figura 3. Diagrama de flujo de la metodología AMG



Fuente: elaboración propia.

Generación de un árbol de decisión

En la metodología AMG se utilizó un árbol tipo CART. Para construirlo dentro de la función `gamlss_tree` se usa la rutina *Recursive Partitioning* del paquete `rpart` (Therneau y Atkinson, 2018), implementada en el *software* R (R Core Team, 2018). La construcción del árbol consiste en encontrar la variable que mejor divida a la población en dos grupos; los datos son separados con dicha variable y luego este mismo proceso es aplicado a cada nuevo grupo, y así de manera recursiva hasta que se alcance un mínimo de cinco observaciones en un subgrupo o hasta que no se encuentre posibilidad de mejora. El modelo final puede ser representado como un árbol binario (Therneau y Atkinson, 2019). Para identificar la variable que mejor divida la población, se evalúa, en el caso de árboles de regresión, la expresión $\Delta_{ss} = SS_T - (SS_R * P_R + SS_L * P_L)$ con $P_R + P_L = 1$ donde SS_T

representa la suma de cuadrado del nodo T, y SS_R representan la suma de cuadrado de los nodos hijos derecho e izquierdo, respectivamente, y P_R y P_L representan la proporción de observaciones en los nodos hijos derecho e izquierdo, respectivamente. Esto es equivalente a maximizar el error entre grupos de un análisis de varianza. Después de construir el árbol de decisión, la población inicial queda dividida en nodos finales, también conocidos como *hojas*, nodos que permiten obtener información interesante sobre el problema como algunas características que identifican observaciones con mayor dispersión y observaciones con valores más altos o más bajos en la variable respuesta.

Preselección de covariables

La selección de las variables explicativas o covariables es un tema de suma importancia en la literatura de modelos de predicción. En primer lugar, porque determinan la capacidad de un modelo de explicar la variabilidad de la variable respuesta y, en segundo lugar, porque definen la complejidad que puede alcanzar un modelo y junto con esta complejidad el tiempo de cómputo que requiere su ajuste.

En el presente trabajo se preseleccionaron las covariables que se consideraron para ajustar los modelos en cada nodo final. Se utiliza el procedimiento que tienen incorporado los árboles de decisión generados en el paso anterior: la función rpart, que genera una importancia para cada variable. Esta se calcula sumando las mejoras que aporta cada variable en cada una de las divisiones en las que se usa; dicha mejora se mide con base en la disminución del error de predicción (suma de cuadrados). El valor de importancia generado en este procedimiento es reescalado para que la suma sea 100; posteriormente, de la lista de importancia de variables se eliminan las de un valor menor a uno (Therneau y Atkinson, 2018). Las variables con importancia mayor a uno, según el árbol de decisión, son las preseleccionadas para ajustar los modelos en cada nodo final.

Pruebas de hipótesis bondad de ajuste

En AMG es posible aprovechar todas las distribuciones de probabilidad definidas en el paquete GAMLSS, dentro de las cuales se encuentran distribuciones para variables continuas, discretas o mixtas. Tales distribuciones pueden tener desde uno hasta cuatro parámetros con la ventaja de modelar incluso variables que no siguen una distribución de la familia exponencial, como variables leptocúrticas, variables platicúrticas, variables con sesgo positivo o negativo y variables que representan conteos con alta dispersión; de hecho, las distribuciones de este paquete se pensaron como una flexibilización de los modelos lineales generalizados (Stasinopoulos y Rigby, 2018). Estas características, como se mencionó, son adecuadas para la distribución del LOS. En AMG para cada una de las distribuciones que el usuario quiere analizar, se realiza la siguiente prueba de hipótesis:

$Ho: y \sim D(\mu, \sigma, \nu, \tau)$

$Ha: y \nsim D(\mu, \sigma, \nu, \tau)$

Donde D pertenece a una de las distribuciones del paquete `gamlss` que, como se mencionó, pueden tener hasta 4 parámetros. Llevar a cabo estas hipótesis implica emplear dos pruebas estadísticas: Kolmogorov-Smirnov, con la función `ks.test`, y Anderson-Darling, con la función `ad.test`. Para evaluar las pruebas de hipótesis se utiliza una significancia definida por el usuario. Este paso también puede usarse como restricción, definiendo en los argumentos de la función diseñada `gamlss_tree`, a fin de que solo se tengan en cuenta las distribuciones para las cuales la prueba de hipótesis es aceptada. De esta manera, los modelos que se ajusten en cada nodo final solo contemplarán distribuciones que se ajustan a los datos analizados.

Preselección de distribuciones

A efectos de seleccionar las distribuciones candidatas se utiliza la función `fitDist` (Rigby y Stasinopoulos, 2005), y con base en una variable y un conjunto de distribuciones definidas por el usuario, la función realiza, para cada distribución, la estimación de los parámetros sin covariables. Con estos ajustes se calculan indicadores para la selección de modelos, como el criterio de información de Akaike generalizado (GAIC) y el BIC. En el caso de AMG, se emplea el GAIC para ordenar las distribuciones de menor a mayor, es decir, la de mejor ajuste de primera. De estas distribuciones se seleccionan las n con mejor GAIC, donde n lo define el usuario. Este paso ayuda a preseleccionar las distribuciones con base en una idea del potencial que tiene cada una para acoplarse a los datos analizados. Además, tal preselección limita el número de modelos GAMLSS que se deben ajustar con covariables haciendo más eficiente la ejecución de la función.

Ajuste de modelos con covariables

En este paso se realiza, para cada una de las distribuciones preseleccionadas en los dos pasos anteriores, el ajuste de un modelo GAMLSS (Rigby y Stasinopoulos, 2005). Estos se introdujeron con el fin de superar las limitaciones asociadas a los modelos lineales generalizados y a los modelos aditivos generalizados. La ecuación (1) es el modelo del AMG. Aunque los modelos GAMLSS permiten modelar diferentes parámetros de una distribución como μ, σ, ν, y τ en el desarrollo de este algoritmo se acotó al modelamiento de μ.

$$g_1(\mu) = X_1 \beta_1 + \sum_{J=1}^{J_1} Z_{j1} \gamma_{j1} \qquad (1)$$

Donde g_1 es la función de enlace del parámetro μ, X_1, es una matriz de n observaciones por J_1' variables explicativas β_1, es el vector de parámetros de longitud J_1'. Z_{j1} es una matriz $n \times q_{j1}$. Finalmente, γ_{j1} es un vector de q_{j1} variables aleatorias. Estas dos últimas componentes son el efecto aleatorio que puede estar conformado, en la forma general, por múltiples efectos aleatorios.

Estos modelos se ajustan para cada una de las distribuciones preseleccionadas y con su ajuste se obtiene el GAIC, el error absoluto medio (MAE), el error cuadrático medio (MSE) y la raíz del error cuadrático medio (RMSE). De acuerdo con la necesidad del usuario se definen los indicadores con los que se selecciona el modelo final.

Selección final de covariables con procedimiento paso a paso

En el paso anterior se obtuvo un modelo con todas las covariables que se preseleccionaron, que garantiza que las utilizadas son importantes para todo el conjunto de datos; sin embargo, no necesariamente son importantes para las observaciones que pertenecen a un nodo. Por esto, se adiciona este componente, que permite seleccionar, con base en las observaciones del nodo final, las variables de mayor importancia. Se realiza un procedimiento paso a paso o *stepwise* aplicando la función stepGAIC con pasos hacia adelante y hacia atrás que permiten obtener un modelo con parsimonia sin sacrificar de manera significativa su desempeño (Rigby y Stasinopoulos, 2005). Con este procedimiento se descartan algunas variables irrelevantes para las observaciones del nodo que se está evaluando.

Procedimiento de **boosting** *o* **committees**

Este procedimiento es opcional, ya que el usuario define el número de iteraciones que se realizan, aun cuando en caso de que el número de iteraciones sea uno, se omite este paso. Cada iteración representa un nuevo árbol generado, pero antes de iniciar el nuevo árbol se modifica el valor de la variable respuesta de acuerdo con la expresión $y_m = y - (\hat{y}_{m-1} - y)$, donde y_m representa el valor ajustado de la variable respuesta que se utilizará en la generación del árbol de decisión número m, y es el valor real de la variable, $\hat{y}_{m-1}$ es el valor predicho en el árbol m - 1. Esta expresión tiene el objetivo de penalizar las observaciones que son sobreestimadas o subestimadas, de tal manera que el próximo modelo genere una predicción menor o mayor, respectivamente. Por último, cuando todos los árboles son generados, la predicción final es el promedio simple del valor predicho de cada uno de ellos (Kuhn y Johnson, 2013). Este procedimiento es conocido como *committees*, nombre que se le asigna en la técnica cubista de donde se toma.

Implementación de la metodología AMG en R

Para la metodología descrita en la sección anterior se desarrolló un paquete en el lenguaje de programación R, que comprende un grupo de funciones que ejecuta los pasos definidos sobre un conjunto de datos. Este paquete puede consultarse en el repositorio de GitHub.[1] Su objetivo es generar múltiples modelos AMG y seleccionar el de mejor desempeño de acuerdo con el indicador definido por el usuario (AIC, BIC, MAE y RMSE). La función `gamlss_tree` ajusta modelos AMG compuestos de un árbol de decisión tipo CART y un modelo de regresión GAMLSS en cada hoja o nodo final creados por el árbol de decisión. La función `pred_gamlss` se emplea en predicciones de observaciones nuevas con base en el modelo ajustado en la función `gamlss_tree`. La función `test_dist` analiza el ajuste de las distribuciones del paquete `gamlss` que defina el usuario, utilizando dos pruebas de bondad de ajuste Anderson-Darling y Kolmogorov-Smirnov. La función `split_sample` divide un conjunto de datos en dos subconjuntos: un conjunto para entrenamiento de modelos y otro para evaluación. El usuario define el porcentaje de datos que va a tener el conjunto de entrenamiento, y el porcentaje de datos restantes corresponderá al grupo de evaluación.

Resultados

En esta sección se analizan tres casos de estudio de predicción del LOS que se manejaron para validar la metodología AMG y para compararla con otras técnicas de la literatura. En la primera parte se describen los casos de estudio, su procedencia y algunas características de los conjuntos de datos. También se comparan los tres casos de estudio y se describen las razones por las que se seleccionaron, basado en sus diferencias y cómo estas pueden impactar en el desempeño de la metodología desarrollada. En la segunda parte se describe brevemente y se compara la distribución de la variable respuesta de cada uno de los casos de estudio, analizando diferentes medidas de resumen y un análisis gráfico. En la tercera parte se presentan las funciones y los argumentos utilizados para generar el modelo de la metodología AMG y los modelos de las otras técnicas con las que se contrastaron los resultados. Por último, se equipara el desempeño del modelo seleccionado por la metodología AMG con los modelos generados por las demás técnicas; esto mediante tres métricas diferentes (AIC, MAE e RMSE) para tener una visión más amplia de las fortalezas y debilidades de cada una.

1 Véase https://github.com/juancamiloespana/gmtree

Descripción de los casos de estudio

Los tres casos de estudio seleccionados tenían características diferentes que permitían analizar el comportamiento de la metodología desarrollada bajo condiciones diversas. En primer lugar, los datos provienen de países diferentes. En segundo lugar, la cantidad de observaciones en cada caso varía de manera importante, lo que puede impactar en el ajuste de los modelos; más aún, teniendo en cuenta que la población se divide con el componente de árboles de decisión y que se reduce la cantidad de observaciones utilizadas para ajustar cada modelo. En tercer lugar, la naturaleza y cantidad de variables explicativas difiere para cada caso.

El primer caso de estudio, que llamaremos Medellín, proviene de los registros médicos de un hospital de tercer nivel de Medellín, Colombia. El tercer nivel de atención resuelve los problemas que exigen mayor grado de especialización, de mayor complejidad y que requieren uso de alta tecnología (Vignolio et al., 2011). La información corresponde a todos los pacientes hospitalizados que tuvieron egresos en el año 2015. La base de datos cuenta con 27 872 registros, 16 variables predictoras y una variable respuesta, que es el LOS en días, medida como variable continua, teniendo en cuenta las horas de ingreso y egreso. De las 16 variables predictoras, 11 son categóricas y 5 son discretas. Dentro de las variables más relevantes encontradas se resalta la que indica si el paciente tuvo tratamiento con antibióticos, y, en el caso de los pacientes con este tipo de tratamientos, era mayor el tiempo de estancia. También resultaron importantes la variable que indicaba la especialidad y el diagnóstico que recibían los pacientes.

El segundo caso de estudio, que llamaremos Microsoft, proviene de un proyecto de Microsoft para la predicción del LOS en el que se empleó una base de datos capaz de evaluar varios modelos (Microsoft, 2018). Tales datos fueron dados por un tercero y se les realizó un remuestreo de las variables *rcount* y *lengthofstay*, que son el número de readmisiones en los últimos 180 días y el LOS, respectivamente. También se realizó un remuestreo de las variables continuas para hacer que sus distribuciones fueran normales y se corrigieron problemas que se tenía con las fechas de altas y número de documentos. En este caso de estudio, con los modelos analizados se encontró que la variable que más influía en la estancia hospitalaria era el número de readmisiones en los últimos 180 días, donde a mayor número de readmisiones, mayor era el tiempo de estancia. También se encontró la variable que indicaba el tipo de servicio donde se encontraba hospitalizado el paciente, la presencia de trastorno psicológico severo y el día cuando se realizaba la visita, que puede afectar debido a la ocupación del hospital en ciertos días.

El tercer caso de estudio, que llamaremos Arizona, proviene de los registros de hospitalización del seguro médico nacional de Estados Unidos, ubicados en una base de datos conocida como Medpar. Los datos pertenecen a una muestra aleatoria de pacientes hospitalizados en 1991, en el estado de Arizona, con problemas cardiovasculares y que recibieron uno de dos procedimientos: *coronary*

artery bypass graft (CABG) y *percutaneous transluminal coronary angioplasty* (PTCA) para ciertos problemas cardiovasculares. La base de datos cuenta con 3589 registros, 5 variables predictoras y una variable respuesta, correspondiente al tiempo de estancia (Hilbe, 2016). En este caso, las variables que influyen principalmente en el tiempo de estancia fueron el procedimiento que se le realiza al paciente y el tipo de admisión del paciente.

Un aspecto importante es que las variables que están disponibles en cada caso de estudio son muy diferentes; solo coinciden en cuestiones demográficas, como la edad y el sexo. Esto evidencia la realidad de la predicción del LOS, donde cada hospital lleva un registro de diferentes tipos. En la introducción se mencionaban tres variables importantes en este tipo de estudios: el diagnóstico, la severidad de la enfermedad y la comorbilidad. En los tres casos estudiados se tiene el diagnóstico, pero no la severidad y la comorbilidad solo está en el caso de Medellín. Esto refleja uno de los principales problemas en la predicción del LOS: la disponibilidad de las variables que, según la literatura, tienen relación fuerte con el LOS. Esta es una oportunidad para los hospitales y, probablemente, es de las que más pueden impactar en el desempeño de los modelos de predicción del LOS.

Comparación de técnicas en los casos de estudio

Como se mencionó en la sección anterior, para cada uno de los casos de estudio se utilizaron cinco técnicas diferentes (AMG, cubista, CART, LM y GAMLSS); sin embargo, dos de estas (AMG y GAMLSS) se evalúan con dos variantes, que consisten en cambiar el criterio con el que se selecciona el modelo final. En una de las variantes se escoge el modelo con mejor MAE y, en la otra, el modelo con mejor AIC. De cada una de las técnicas resulta un modelo, y a este se le miden tres indicadores (AIC, MAE y RMSE) y con los resultados se comparan las técnicas. El primer indicador (MAE) se basa en la bondad de ajuste con penalización dada por el número de parámetros y los dos últimos (MAE y RMSE) como medida únicamente de los residuales. Estas tres perspectivas permiten una comparación más amplia sobre los criterios para los que funciona mejor cada una de las técnicas, ya que, como se mencionó, según la necesidad del investigador, el criterio de selección puede cambiar. El indicador AIC no se calculó para los árboles CART y modelos cubistas, ya que los modelos generados no permiten su cálculo. En los modelos generados por la metodología AMG, que tienen varios modelos paramétricos ajustados, uno por cada nodo, se suman los AIC del modelo de cada nodo. Esto implica una penalización mayor en estos, porque, a pesar de usar las mismas variables, cada una de ellas ajusta un parámetro diferente en cada nodo. Las métricas usadas corresponden a la medición en la muestra de evaluación. En la tabla 1 se presentan los resultados del caso de Medellín; en la tabla 2, los del caso de Microsoft, y, en la tabla 3, los del caso de Arizona.

Tabla 1. Comparación de desempeño de técnicas en el caso de estudio Medellín

Técnicas	RMSE	AIC	MAE	MAE/ media (en porcentaje)	MAE/ mediana (en porcentaje)
AMG-MAE	5,64	135 197	3,51	59	88
AMG-AIC	7,42	115 026	4,57	76	114
Cubista	5,67	NA	3,32	55	83
CART	5,68	NA	3,63	61	91
LM	5,47	140 600	3,55	59	89
GAMLSS-AIC	7,26	140 600	4,55	76	114
GAMLSS-MAE	5,48	115 645	3,56	59	89

Fuente: elaboración propia.

Tabla 2. Comparación de desempeño de técnicas en el caso de estudio de Microsoft

Técnicas	RMSE	AIC	MAE	MAE/ media (en porcentaje)	MAE/ mediana (en porcentaje)
AMG-MAE	1,1	121 722	0,85	21	21
AMG-AIC	2,09	94 418	1,58	40	40
Cubista	1,2	NA	0,82	21	21
CART	1,24	NA	0,95	24	24
LM	1,18	126 311	0,9	23	23
GAMLSS-AIC	1,83	111 416	1,34	34	34
GAMLSS-MAE	1,18	126 311	0,9	23	23

Fuente: elaboración propia.

Tabla 3. Comparación de desempeño de técnicas en el caso de estudio Arizona

Técnicas	RMSE	AIC	MAE	MAE/ media (en porcentaje)	MAE/ mediana (en porcentaje)
AMG-MAE	5,34	17 182	3,41	39	43
AMG-AIC	9,27	15 113	6,83	77	85
Cubista	5,71	NA	3,23	37	40
CART	5,35	NA	3,43	39	43
LM	5,57	17 843	3,45	39	43
GAMLSS-AIC	9,43	15 356	6,81	77	85
GAMLSS-MAE	5,57	17 843	3,45	39	43

Fuente: elaboración propia.

Análisis y discusión

En los casos de estudio analizados se lograron resultados muy diferentes. Para el caso de Microsoft, los modelos llegaron a predicciones muy buenas que se desvían, en promedio, menos de un día, es decir, para un paciente que se predice va a estar cinco días hospitalizado, la estancia real podría estar en cuatro días, por ejemplo, lo cual sería un resultado muy bueno con el cual se podrían definir diferentes estrategias. Por otro lado, en casos como el de Medellín, la desviación de la predicción en promedio puede ser de tres días, que es una predicción menos ajustada. Esto refleja la importancia de tener un registro de las variables adecuadas que permitan tener errores menores y de esta manera facilitar el accionar de diferentes estrategias.

En la tabla 1 se muestran los resultados de los tres indicadores evaluados para cada uno de los siete modelos seleccionados en el caso de estudio de Medellín. Como se observa, el MAE es menor que el RMSE en todos los casos; esto se debe a la presencia de valores extremos a la que es más sensible el RMSE. En el indicador MAE/media se observa que los modelos seleccionados tienen un porcentaje de error alto en todas las técnicas. Tal vez ello es consecuencia de la baja correlación entre las variables predictoras y el LOS en este caso de estudio.

El modelo seleccionado con AMG-AIC es el que mejor indicador AIC tiene; sin embargo, es el de menor desempeño en MAE y RMSE. Por otra parte, el modelo seleccionado con AMG-MAE es el segundo que mejor indicador MAE obtuvo y el tercero en RMSE, superado por el modelo seleccionado de regresión

lineal y GAMLSS-MAE. Esto último es lo esperado, ya que en estos dos métodos los parámetros se estiman con el objetivo de minimizar la suma de cuadrados del error.

En la tabla 2 se presentan los resultados de los indicadores de la ejecución de las siete técnicas sobre el caso de estudio de Microsoft. Aquí, el error MAE está entre el 21 % y el 40 % de la media, lo que muestra un porcentaje de error bajo, debido a que para este conjunto de datos las variables predictoras sí son significativas explicando el LOS. El modelo seleccionado con la técnica AMG-AIC obtuvo el mejor AIC de los siete modelos seleccionados; no obstante, fue el que tuvo un error más alto en MAE y RMSE. Por otro lado, el modelo seleccionado de la técnica AMG-MAE fue el segundo mejor en el indicador de MAE después de la técnica cubista, por una diferencia muy pequeña, y fue el de mejor desempeño en el indicador RMSE.

Al igual que en el caso anterior, el indicador MAE también es menor que el RMSE para todos los modelos seleccionados, aunque para este caso las diferencias entre estos dos indicadores son menores, debido a que en este conjunto de datos la asimetría es menor y, en consecuencia, el impacto que este tiene sobre el RMSE también.

En la tabla 3 se presentan los indicadores de los siete modelos seleccionados para el caso de estudio de Arizona. Se obtuvieron resultados muy similares al caso de estudio anterior, en cuanto a las mejores técnicas en cada métrica: en AIC el mejor desempeño lo tuvo el modelo de AMG-AIC y, al igual que en los casos anteriores, fue el de mayor error en las dos métricas restantes. En el indicador MAE, el mejor modelo fue seleccionado por la técnica cubista, al igual que en los dos casos anteriores, teniendo en segundo lugar el modelo de la técnica AMG-MAE, que generó el modelo que obtuvo el indicador RMSE. En este caso de estudio, al igual que en el primero, las diferencias entre las técnicas más competitivas de cada indicador son más estrechas que en el caso de estudio de Microsoft. Esto puede ser consecuencia de una menor capacidad predictora de las variables explicativas, que se refleja en el hecho de que en este conjunto de datos el error MAE está entre el 37 % y el 78 % de la media, un poco mejor que en el caso de estudio de Medellín, pero con mayor porcentaje de error que el caso de estudio de Microsoft.

En los resultados se observa que la técnica AMG-AIC generó el modelo con el mejor indicador AIC en los tres casos de estudio. Esto representa un resultado muy importante en el contexto de predicción del LOS, no solo por la mejora en el indicador, sino también por la introducción de un indicador que no es común en la literatura reciente del área y que, como se mencionó en la revisión de la literatura, es uno de los criterios de selección de modelos más importantes, por tener en cuenta la bondad de ajuste y por la penalización que se realiza por la complejidad del modelo, medida en el número de parámetros que se estiman.

En los indicadores basados en los residuales RMSE y MAE, el modelo de la técnica AMG-MAE tuvo resultados competitivos, ya que alcanzó el mejor RMSE en dos de los casos y en el tercero logró el segundo mejor indicador. Este mismo modelo obtuvo el segundo mejor indicador MAE en los tres casos. En el caso de estudio de Microsoft consiguió los mejores indicadores de los modelos generados por las técnicas AMG, lo cual indica que en escenarios con mayor cantidad de datos y mejor calidad de variables es donde toma mayor ventaja sobre otras técnicas. Un aspecto importante que se debe resaltar es la diferencia que se presenta en los modelos generados por las técnicas AMG-AIC y AMG-MAE, ya que el modelo con mejor AIC fue el que peor indicador MAE y RMSE obtuvo, y viceversa, lo que muestra la importancia de definir el criterio con el que se escogen los modelos.

El caso de estudio de Microsoft fue el que mejores indicadores logró. Esto resalta la importancia de tener variables explicativas que influyan en la variable respuesta. Adicionalmente, en este caso se tiene mayor cantidad de variables continuas, que son más fáciles de procesar por parte de las técnicas de modelado. En el caso de estudio de Arizona se tenían pocos registros y pocas variables; sin embargo, los indicadores de los modelos fueron mejores que en el primer caso, lo que evidenció que existía información más relevante para la predicción. Otro tema que se observa en el análisis de los casos es que la diferencia entre la métrica MAE y RMSE es mayor mientras mayor sea la presencia de valores extremos, como se da en el caso de Medellín y Arizona, con una mayor brecha entre las dos métricas. En contraste, se tiene el caso de estudio de Microsoft, que con menor dispersión presenta una menor brecha entre ambos indicadores.

El desarrollo e implementación de la metodología AMG permitió comprender que para el caso del LOS y otros problemas con las mismas características, es importante tener presente que un conjunto de datos, por su naturaleza, puede provenir de poblaciones con distribuciones diferentes y que, por lo tanto, una técnica que permita identificar y segmentar estas poblaciones ayuda a modelar de manera más acertada el comportamiento de la variable respuesta. Adicionalmente, se identificó una diferencia importante entre los modelos seleccionados basados en métricas de minimización de residuales, como MAE y RMSE, respecto a los modelos seleccionados con base en AIC; diferencia en cuanto a sus métricas de desempeño y en cuanto a la naturaleza de la distribución ajustada. En muchos problemas, conocer la naturaleza de la distribución de los datos es fundamental, por lo cual utilizar una métrica como el AIC será indispensable. Finalmente, se resalta la importancia de tener covariables adecuadas que permitan tener una predicción ajustada, ya que sin estas covariables ninguna técnica arrojará un resultado bueno.

Conclusiones

En este trabajo se identificaron algunas oportunidades en el contexto de predicción del LOS que se abordaron con el desarrollo del trabajo. La primera es tener un estudio enfocado en la metodología que generaba un modelo de predicción, más que en el modelo mismo y sus resultados. Esta oportunidad se trabajó con el diseño e implementación en *software* de la metodología AMG, que facilita la selección de modelos para la predicción del LOS en un espectro amplio de posibles modelos. La segunda oportunidad está relacionada con el uso de métricas más enfocadas en la bondad de ajuste que en la minimización de residuales, como era el caso de la mayoría de artículos encontrados en la literatura. Para esto se utilizó, entre otras, la métrica AIC y esto llevo a la necesidad de incluir diferentes tipos de distribuciones, además de pruebas de hipótesis de ajustes de distribuciones que permitieran mejorar el ajuste de los modelos basados en AIC. Finalmente, se identificó la oportunidad de mejorar el desempeño de los modelos de predicción del LOS, lo cual se logró generando el mejor modelo en todos los casos evaluados en la métrica AIC y el mejor modelo en la mayoría de casos evaluados con la métrica MAE.

En la literatura de predicción del LOS, principalmente desde los métodos computacionales que son los más comunes en la actualidad, no se aborda con profundidad el ajuste de los datos a las distribuciones paramétricas sobre las que se basan muchos modelos empleados. Como consecuencia se tienden a ajustar regresiones que no cumplen los supuestos. Con la presente investigación se aborda esta problemática y se genera una metodología que utiliza métodos computacionales, pero que permite el análisis de ciertas características importantes en los modelos de regresión como el ajuste de los datos a la distribución sobre la que se basa el modelo, generando dentro de la metodología AMG la versatilidad para que el usuario defina la rigurosidad con que se ajustará el modelo y sea consciente del incumplimiento de algunos supuestos para las conclusiones que se generen de estos.

Las construcciones de modelos predictivos con un desempeño adecuado dependen fundamentalmente de tener variables predictoras que expliquen gran parte de la variabilidad de la variable respuesta; no tenerlas implica errores más grandes y, tal vez, la imposibilidad de realizar una predicción acertada. Esto se evidencia dentro del primer caso de estudio, para el cual los modelos generaron predicciones con un margen de error muy amplio y, aunque existían diferencias en los modelos ajustados, eran insuficientes para lograr un desempeño apropiado. Este desempeño inferior al esperado se asocia con las variables explicativas, ya que las mismas técnicas en el caso de estudio 2 mostraron un desempeño muy superior. Este aspecto se debe resolver desde las instituciones hospitalarias, garantizando que capturan la información adecuada.

En los modelos construidos con la metodología AMG se encontró que los que fueron seleccionados a través del AIC, que prioriza la bondad de ajuste, mos-

traban un mejor comportamiento en este indicador con respecto a las otras técnicas, pero no fueron los que tuvieron mejor desempeño en los indicadores MAE y RMSE. Estos modelos mostraban una tendencia a ajustarse mejor al modelo real de los datos, de acuerdo con la naturaleza del AIC que mide la bondad de ajuste. Además, la metodología desarrollada tenía la variante de seleccionar un modelo enfocado en los indicadores MAE y RMSE, que permitía tener predicciones muy competitivas en comparación con las técnicas que se comparó.

El desempeño del modelo generado con la técnica AMG-MAE superó los resultados que se obtienen en los modelos generados por las técnicas en las que se basa (árboles de decisión y modelos GAMLSS), analizándolos individualmente y superando en casi todos los escenarios a la regresión lineal. Esto muestra que es una metodología competitiva que puede generar resultados de mucho interés en las investigaciones que se utilice. Adicionalmente, la metodología desarrollada tiene un potencial que no se alcanzó a explotar en el presente trabajo, potencial que se encuentra en el uso de los componentes aditivos de los modelos GAMLSS y en el ajuste de los parámetros de escala y forma de cada distribución. Explotar estos componentes adicionales puede mejorar de manera considerable el desempeño de los modelos generados y permitir una flexibilidad mayor en cuanto al tipo de fenómenos que se modelan.

La presente investigación se enfocó en generar las bases para un campo de investigación futuro en el que se puedan proponer metodologías que tengan en cuenta tanto el enfoque computacional como el enfoque estadístico en la construcción de modelos, para que no se conciban como campos de investigación independientes, sino como complementarios. De esta manera, se podrá tener un conocimiento más amplio del fenómeno estudiado.

Referencias

Al Taleb, A. R., Hoque, M., Hasanat, A. y Khan, M. B. (2017). Application of data mining techniques to predict length of stay of stroke patients. En *2017 International Conference on Informatics Health and Technology* (pp. 1-5). IEEE. https://doi.org/10.1109/ICIHT.2017.7899004

Awad, A., Bader-El-Den, M. y McNicholas, J. (2017). Patient length of stay and mortality prediction: A survey. *Health Services Management Research*, *30*(2), 105-120. https://doi.org/10.1177/0951484817696212

Azari, A., Janeja, V. P. y Levin, S. (2015). Imbalanced learning to predict long stay Emergency Department patients. En *International Conference on Bioinformatics and Biomedicine (BIBM) Imbalanced* (pp. 807-814). IEEE. https://doi.org/10.1109/BIBM.2015.7359790

Azari, A., Janeja, V. P. y Mohseni, A. (2012). Predicting hospital length of stay (PHLOS): A multi-tiered data mining approach. En *IEEE 12th International Con-*

ference on Data Mining Workshops (pp. 17-24). IEEE. https://doi.org/10.1109/ICDMW.2012.69

Barnes, S., Hamrock, E., Toerper, M., Siddiqui, S. y Levin, S. (2016). Real-time prediction of inpatient length of stay for discharge prioritization. *Journal of the American Medical Informatics Association* (23), e2-e10. https://doi.org/10.1093/jamia/ocv106

Box, G. E. P. y Cox, D. R. (1964). An analysis of transformations: Applying the Box-Cox transformation. *Journal of the Royal Statistical Society*, *Series B* (26), 211-252.

Burns, L. R. y Wholey, D. R. (1991). The effects of patient, hospital, and physician characteristics on length of stay and mortality. *Medical Care*, *29*(3), 251-271.

Cai, X., Pérez-Concha, O., Coiera, E., Martín-Sánchez, F., Day, R., Roffe, D. y Gallego, B. (2016). Real-time prediction of mortality, readmission, and length of stay using electronic health record data. *Journal of the American Medical Informatics Association*, *23*(3), 553-561. https://doi.org/10.1093/jamia/ocv110

Carter, E. M. y Potts, H. W. (2014). Predicting length of stay from an electronic patient record system: A primary total knee replacement example. *BMC Medical Informatics and Decision Making*, *14*(1), 14-26. https://doi.org/10.1186/1472-6947-14-26

Dayhoff, J. E. y DeLeo, J. M. (2001). Artificial neural networks: Opening the black box. *Cancer*, *91*(S8), 1615-1635. https://doi.org/10.1002/1097-0142(20010415)91:8+<1615::AID-CNCR1175>3.0.CO;2-L

Faddy, M., Graves, N. y Pettitt, A. (2009). Modeling length of stay in hospital and other right skewed data: Comparison of phase-type, gamma and log-normal distributions. *Value in Health*, *12*(2), 309-314. https://doi.org/10.1111/j.1524-4733.2008.00421.x

Fontova-Almató, A., Juvinyà-Canal, D. y Suñer-Soler, R. (2015). Influence of waiting time on patient and companion satisfaction. *Revista de Calidad Asistencial*, *30*(1), 10-16. https://doi.org/10.1016/j.cali.2014.12.009

Gustafson, D. H. (1968). Length of stay: Prediction and explanation. *Health Services Research*, *3*(1), 12-34.

Houthooft, R., Ruyssinck, J., van der Herten, J., Stijven, S., Couckuyt, I., Gadeyne, B., Ongenae, F., Colpaert, K., Decruyenaere, J., Dhaene, T. y De Turck, F. (2015). Predictive modelling of survival and length of stay in critically ill patients using sequential organ failure scores. *Artificial intelligence in medicine*, *63*(3), 191-207. https://doi.org/10.1016/j.artmed.2014.12.009

Hilbe, J. M. (2016). *COUNT: Functions, data and code for count data*. Cambridge University Press. https://cran.r-project.org/package=COUNT

Jiménez, R., López, L., Domínguez, D. y Fariñas, H. (1999). Difference between observed and predicted length of stay as an indicator of inpatient care inefficiency.

International Journal for Quality in Health Care, *11*(5), 375-384. https://doi.org/10.1093/intqhc/11.5.375

Kaur, G. y Chhabra, A. (2014). Improved J48 classification algorithm for the prediction of diabetes. *International Journal of Computer Applications*, *98*(22), 13-17. https://doi.org/10.5120/17314-7433

Kuhn, M. y Johnson, K. (2013). *Applied predictive modeling*. Springer. https://doi.org/10.1007/978-1-4614-6849-3

Kumari, M., Vohra, R. y Arora, A. (2014). Prediction of diabetes using Bayesian network. *International Journal of Computer Science and Information Technologies*, *5*(4), 5174-5178.

Liu, H. C. (2013). A theoretical framework for holistic hospital management in the Japanese healthcare context. *Health Policy*, *113*(1-2), 160-169. https://doi.org/10.1016/j.healthpol.2013.08.009

Liu, P., Lei, L., Yin, J., Zhang, W., Naijun, W. y El-Darzi, E. (2006). Healthcare data mining: Prediction inpatient length of stay. En *3rd International IEEE Conference Intelligent Systems* (pp. 832-837). IEEE. https://doi.org/10.1109/IS.2006.348528

Liu, V., Kipnis, P., Gould, M. K. y Escobar, G. J. (2010). Length of stay predictions improvements through the use of automated laboratory and comorbidity variables. *Medical Care, 48*(8), 739-744. http://www.jstor.org/stable/25701529

Marazzi, A., Paccaud, F., Ruffieux, C. y Beguin, C. (1998). Fitting the distributions of length of stay by parametric models. *Medical Care*, *36*(6), 915-927. http://www.jstor.org/stable/3767008

Marshall, A., Vasilakis, C. y El-Darzi, C. (2005). Length of stay-based patient flow models: Recent developments and future directions. *Health Care Management Science*, *8*(3), 213-220. https://doi.org/10.1007/s10729-005-2012-z

Microsoft. (2018). *Predicting hospital length of stay*. https://microsoft.github.io/r-server-hospital-length-of-stay/index.html

Mobley, B. A., Leasure, R. y Davidson, L. (1995). Artificial neural network predictions of lengths of stay on a post-coronary care unit. *Heart and Lung. The Journal of Cardiopulmonary and Acute Care, 24*(3), 251-256. https://doi.org/10.1016/S0147-9563(05)80045-7

Nouaouri, I., Samet, A. y Allaoui, H. (2015). Evidential data mining for length of stay (LOS) prediction problem. En *IEEE International Conference on Automation Science and Engineering* (pp. 1415-1420). IEEE. https://doi.org/10.1109/CoASE.2015.7294296

Pahlevan Sharif, S. (2017). Locus of control, quality of life, anxiety, and depression among Malaysian breast cancer patients: The mediating role of uncertainty. *European Journal of Oncology Nursing*, *27*(2017), 28-35. https://doi.org/10.1016/j.ejon.2017.01.005

Pendharkar, P. C. y Khurana, H. (2014). Machine learning techniques for predicting hospital length of stay in pennsylvania federal and specialty hospitals. *International Journal of Computer Science and Applications*, *11*(3), 45-56. http://www.tmrfindia.org/ijcsa/v11i33.pdf

Pérez, A., Chan, W. y Dennis, R. J. (2006). Predicting the length of stay of patients admitted for intensive care using a first step analysis. *Health Services Outcomes Research Methodology*, *6*(2006), 127-138. https://doi.org/10.1007/s10742-006-0009-9

Pérez-Castro, N., Márquez-Grajales, A., Acosta-Mesa, H. G. y Mezura-Montes, E. (2017). Full model selection issue in temporal data through evolutionary algorithms: A brief review. En *IEEE Congress on Evolutionary Computation* (pp. 2451-2457). IEEE. https://doi.org/10.1109/CEC.2017.7969602

Pilotto, A., Sancarlo, D., Pellegrini, F., Rengo, F., Marchionni, N., Volpato, S. y Ferrucci, L. (2016). The multidimensional prognostic index predicts in-hospital length of stay in older patients: A multicentre prospective study. *Age and Ageing*, *45*(1), 90-96. https://doi.org/10.1093/ageing/afv167

Quinlan, J. R. (1992). Learning with continuous classes. En *5th Australian joint conference on artificial intelligence* (vol. 92, pp. 343-348). https://sci2s.ugr.es/keel/pdf/algorithm/congreso/1992-Quinlan-AI.pdf

R Core Team. (2018). *R Project*. *R Foundation for Statistical Computing*. https://www.r-project.org/

Rigby, R. A. y Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *54*(3), 507-554. https://doi.org/10.1111/j.1467-9876.2005.00510.x

Robinson, G. H., Davis, L. E. y Leifer, R. P. (1966). Prediction of hospital length of stay. *Health Services Research, 1*(3), 287-300. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1067345/

Silva, Á., Cortez, P., Santos, M. F., Gomes, L. y Neves, J. (2008). Rating organ failure via adverse events using data mining in the intensive care unit. *Artificial Intelligence in Medicine*, *43*(3), 179-193. https://doi.org/10.1016/j.artmed.2008.03.010

Stasinopoulos, M. y Rigby, R. (2018). *gamlss.dist: Distributions for generalized additive models for location scale and shape*. https://cran.r-project.org/web/packages/gamlss.dist/gamlss.dist.pdf

Tanuja, S., Acharya, D. U. y Shailesh, K. R. (2011). Comparison of different data mining techniques to predict hospital length of stay. *Journal of Pharmaceutical and Biomedical Sciences*, *7*(7), 1-4. http://eprints.manipal.edu/757/1/Shailesh_K_R_et._al.%281%29.pdf

Therneau, T. y Atkinson, B. (2018). *RPART: Recursive partitioning and regression trees*. https://cran.r-project.org/package=rpart

Therneau, T. y Atkinson, E. J. (2019). *An introduction to recursive partitioning: Using the RPART routines*. https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf

Turgeman, L., May, J. H. y Sciulli, R. (2017). Insights from a machine learning model for predicting the hospital length of stay (LOS) at the time of admission. En *78 Expert Systems with Applications* (pp. 376-385). Elsevier. https://doi.org/10.1016/j.eswa.2017.02.023

Turgeman, L., May, J., Ketterer, A., Sciulli, R. y Vargas, D. (2015). Identification of readmission risk factors by analyzing the hospital-related state transitions of congestive heart failure (CHF) patients. *IIE Transactions on Healthcare Systems Engineering*, *5*(4), 255-267. https://doi.org/10.1080/19488300.2015.1095823

Vaish, A., Vaish, A., Vaishya, R. y Bhawal, S. (2016). Customer relationship management (CRM) towards service orientation in hospitals: A review. *Apollo Medicine*, *13*(4), 1-5. https://doi.org/10.1016/j.apme.2016.11.002

Verburg, I. W. M., De Keizer, N. F., De Jonge, E. y Peek, N. (2014). Comparison of regression methods for modeling intensive care length of stay. *PLoS ONE*, *9*(10), 1-11. https://doi.org/10.1371/journal.pone.0109684

Vignolio, J., Vacarezza, M., Álvarez, C. y Sosa, A. (2011). Niveles de atención, de prevención y atención primaria de la salud. *Prensa Médica Latinoamericana*, *33*(1), 11-14. http://www.scielo.edu.uy/pdf/ami/v33n1/v33n1a03.pdf

Weir, E., D'Entremont, N., Stalker, S., Kurji, K. y Robinson, V. (2009). Applying the balanced scorecard to local public health performance measurement: Deliberations and decisions. *BMC Public Health*, *9*(1), 127-133. https://doi.org/10.1186/1471-2458-9-127

Widyastuti, Y., Stenseth, R., Wahba, A., Pleym, H. y Videm, V. (2012). Length of intensive care unit stay following cardiac surgery: Is it impossible to find a universal prediction model? *Interactive Cardiovascular and Thoracic Surgery*, *15*(5), 825-833. https://doi.org/10.1093/icvts/ivs302

Capítulo 7.
Marco de trabajo para la vigilancia epidemiológica en Santander, bajo el paradigma *big data*

Yuly Andrea Ramírez Sierra[1]
y Henry Lamos Díaz[2]

[1] Escuela de Estudios Industriales y Empresariales
Universidad Industrial de Santander
Bucaramanga, Colombia
yaramsie@correo.uis.edu.co

[2] Escuela de Estudios Industriales y Empresariales
Universidad Industrial de Santander
Bucaramanga, Colombia
hlamos@uis.edu.co

Resumen

La vigilancia es un rasgo esencial de la práctica epidemiológica y de la salud pública, y es el resultado de la recolección, validación, análisis e interpretación de datos de salud y enfermedades. El valor de la vigilancia implica la entrega efectiva y eficiente de información útil. Los sistemas de vigilancia deben ser flexibles frente al incremento de las necesidades de información y adoptar tecnologías y herramientas apropiadas que permitan gestionar la información oportunamente. En esta investigación, con base en las etapas del proceso de descubrimiento de conocimiento en bases de datos (KDD), se construyó un marco de trabajo como guía en las etapas de análisis e interpretación de datos y difusión de la información de la vigilancia epidemiológica de enfermedades infecciosas y no transmisibles (o crónicas) en Santander, considerando el paradigma *big data* y adoptando técnicas de minería de datos y de aprendizaje automático, con el propósito de apoyar a los tomadores de decisiones en salud pública. Esta metodología se adapta al estudio de caso del dengue, que evidencia las ventajas de integrar analítica descriptiva y predictiva, a fin de analizar el nivel de incidencia del evento y apoyar a los tomadores de decisiones en salud pública.

Palabras clave: salud pública, vigilancia epidemiológica, minería de datos, aprendizaje automático, *big data*

Introducción

El campo de la epidemiología es considerado el centro de la salud pública (Ogino et al., 2015). Durante décadas, ha proporcionado una base científica para la salud en prevención de enfermedades y ha contribuido a importantes descubrimientos científicos (Khoury et al., 2013). Se destaca que, a través de la historia, la epidemiología se ha preocupado por tres aspectos: 1) la extensión espacial de los brotes, 2) el progreso de la enfermedad y los posibles controles y 3) la identificación del origen de la enfermedad y la comparación con brotes ocurridos en el pasado (Pyne et al., 2015).

Por consiguiente, se requiere que la epidemiología amplíe su alcance más allá de la perspectiva histórica de la etiología, para abarcar la detección temprana de la enfermedad, el tratamiento, el pronóstico, y se convierta en una traducción más eficaz de los descubrimientos científicos, generando impacto en la salud individual y poblacional. Tales retos le permitirán enfocarse en la precisión del análisis de los factores de riesgo para abordar cuestiones de gran importancia social actual, como la economía de los servicios de salud, el envejecimiento de la población, la creciente carga de enfermedades crónicas comunes, la persistencia de las desigualdades sanitarias y la salud mundial (Khoury, 2015; Khoury et al., 2013).

Como rasgo esencial de la práctica epidemiológica se destaca la vigilancia (Porta, 2008), que es el resultado de la recolección, la validación, el análisis y la interpretación de datos de salud y enfermedades (Amato-Gauci y Ammon, 2008; Tsui et al., 2011), de gran importancia para que los interesados basados en la evidencia implementen estrategias encaminadas a la prevención y control de estas últimas (Amato-Gauci y Ammon, 2008).

Según Robertson (2017), el impacto global de las enfermedades se ha ido desplazando desde aquellas transmisibles a las no transmisibles y lesiones, aunque las enfermedades infecciosas continúan produciendo persistentes y episódicas amenazas a la salud pública. Así, la vigilancia y el control de enfermedades transmisibles es fundamental para el estatus de salud de una comunidad (O'Connell et al., 2010). La eficiencia de la transmisión de enfermedades infecciosas depende del nivel de infección del huésped infectado y de la susceptibilidad de los individuos no infectados, pero que están expuestos a la infección. Así, tanto la infecciosidad como la susceptibilidad involucran los componentes biológico, comportamental y ambiental (Grassly y Fraser, 2008). En el caso de las enfermedades arbovirales, su transmisión depende de las interacciones entre virus, mosquitos y huéspedes (*hosts*), y estas interacciones son afectadas por factores abióticos —como temperatura, precipitaciones y humedad— y por factores bióticos —como abundancia de huéspedes vertebrados y mosquitos vectoriales— (Gu et al., 2008).

Por otro lado, las enfermedades no transmisibles, también conocidas como enfermedades crónicas (Instituto Nacional de Salud, 2019; Organización Mundial de la Salud [OMS], 2018), se caracterizan por que las epidemias tardan décadas

en establecerse por completo, dado que se originan a edades tempranas, requieren un enfoque sistemático a largo plazo para el tratamiento y, por su larga duración, existen múltiples oportunidades para su prevención (World Health Organization, 2017). La OMS propone nuevas metas para el 2025, relacionadas con las enfermedades no transmisibles, a fin de lograr los objetivos de desarrollo sostenible, enfocándose en reducir el impacto de los factores de riesgo como el consumo de alimentos grasosos, el bajo consumo de frutas y verduras, el sobrepeso, la obesidad, el uso nocivo del alcohol, la inactividad física, el estrés psicológico y los determinantes socioeconómicos. Chen et al. (2016) proponen la cadena de factores para las causas principales de las enfermedades crónicas e incluyen tres tipos: factores no cambiantes (edad y herencia), factores cambiantes y factores difíciles de ser cambiados.

Entonces, el análisis de la situación de la salud puede dar conocimiento de oportunidades y prioridades para la prevención, investigación, políticas y desarrollo (Forouzanfar et al., 2016). La prevención puede mejorar significativamente la salud humana, y tiene como objetivo central reducir o modificar los factores de riesgo que influyen en la salud de la población. Así, tanto la cuantificación de estos riesgos como el monitoreo actualizado de la evidencia basada en factores de riesgo son claves para la salud pública y fundamentales para modificar el riesgo individual a través de la atención primaria y la autogestión (GBD 2016 Risk Factors Collaborators, 2017). En la tabla 1 se muestran los factores de riesgo generales para cada grupo de las enfermedades infecciosas y para enfermedades no transmisibles o crónicas.

Tabla 1. Factores de riesgo para enfermedades infecciosas o crónicas

Enfermedad	Factores de riesgo
Enfermedades transmitidas por vectores	• Temperatura de la superficie de la tierra • Aumento de la temperatura de la superficie del mar • Elevación del nivel del mar • Comportamiento del vector • Componente biológico del patógeno • Ambiente del individuo infectado • Comportamiento de la población • Latitud, longitud y altitud del área de estudio • Cambios antropogénicos y demográficos • Oscilación sureña de El Niño (ENSO) • Factores abióticos como temperatura, precipitaciones y humedad • Factores bióticos como abundancia de huéspedes vertebrados y mosquitos vectoriales
Enfermedades inmunoprevenibles	• Información sobre programas de vacunación

Enfermedad	Factores de riesgo
Enfermedades infecciosas	• Precipitaciones • Temperatura • Tiempo de vacaciones • Factores biológicos: carga viral o bacteriana del patógeno infeccioso • Factores comportamentales: patrones de contacto de un individuo infectado y los patrones de contacto del huésped o vector intermediario • Factores ambientales: la ubicación y ambiente del individuo infectado • Oscilación sureña de El Niño (ENSO) • Factores ecológicos • Factores socioeconómicos: densidad de la población, uso de medicamentos antibióticos, prácticas agrícolas, nivel de pobreza
Influenza	• Contacto directo o indirecto con una persona infectada • Enfermedad crónica como asma, diabetes, enfermedad del corazón o una condición inmune comprometida • Vivir con niños, ser mujer, pertenecer a un grupo de edad joven, tener mascotas en casa (gatos o perros) y ser fumador
Enfermedades transmitidas por alimentos	• Proceso de cocción o calentamiento de los alimentos • Prácticas de higiene personal • Alimentos preparados con materias primas contaminadas • Limpieza de utensilios y equipo de la cocina • Condiciones ambientales que permiten el crecimiento de patógenos selectivos e inhiben los microrganismos competidores • Alimentos obtenidos de fuentes no confiables • Prácticas de almacenamiento • Uso de utensilios o recipientes que contienen materiales tóxicos • Adición de sustancias químicas tóxicas a los alimentos • Agua no potable o agua de una fuente suplementaria no controlada
Infecciones asociadas a la atención en salud (IAAS)	• Infecciones asociadas a dispositivos (IAD) en unidades de cuidado intensivo • Resistencia bacteriana a los antimicrobianos (RA) en el ámbito hospitalario • Consumo de antibióticos (CAB) en el ámbito hospitalario
Infecciones de transmisión sexual (Connolly et al., 2004)	• Ausencia de tratamientos • Escasez de preservativos • Violencia sexual
Enfermedades por microbacterias	• Reemergencia de la enfermedad en países donde se consideraba casi eliminada • Aparición de casos de tuberculosis farmacorresistente y tuberculosis extensivamente drogorresistente (MDR-XDR TB) • Personas infectadas con el VIH • Debilitamiento de los programas de control

Enfermedad	Factores de riesgo
Enfermedades no transmisibles o crónicas	• Edad, herencia, fumar, falta de actividad física, estado emocional, nivel de pobreza, estándar de vida, ocupación, lugar de residencia, consumir alcohol • Hábitos alimenticios: bajo consumo de frutas y verduras, consumo excesivo de sal • Estudio excesivo • Ambientales y sociales: envejecimiento de la población, urbanización y globalización • Obesidad, sobrepeso, presión arterial elevada, alta glucosa en la sangre, lípidos sanguíneos anormales

Fuente: elaboración propia.

En varios estudios se ha explorado el uso de nuevas fuentes de información con el fin de plantear metodologías efectivas y rentables para la vigilancia de la salud (Raghupathi y Raghupathi, 2014), otorgándole importancia a las redes sociales como apoyo a la vigilancia de enfermedades, el análisis en tiempo real de las enfermedades, el análisis espacio-temporal de las epidemias (Broniatowski et al., 2013; Ginsberg et al., 2009; Ram et al., 2015; Yang et al., 2015). Así mismo, los teléfonos móviles pueden apoyar la recolección de información en tiempo real para la toma de decisiones ante una epidemia (Grossglauser y Saner, 2014). Además, los datos clínicos no estructurados y los datos ambientales y de cambio climático proporcionan oportunidades para mejorar el conocimiento de la salud, dado que otorgan más detalles de los eventos de interés con respecto a tiempo, lugar y persona (Hall et al., 2012).

Debido al potencial dado al *big data*, se ha generado un nuevo campo de investigación que ha sido atractivo para diversos sectores como industrias, gobiernos y comunidad de investigadores (Hu et al., 2014). Según un informe del Instituto Global McKinsey (Manyika et al., 2011): "Big Data son conjuntos de datos cuyo tamaño está más allá de la capacidad de las típicas herramientas de software de base de datos para capturar, almacenar, administrar y analizar", y para Hu et al. (2014) esta definición es subjetiva, porque no define al *big data* en términos de una métrica, pero engloba un aspecto dinámico, ya sea en el tiempo o en el sector, respecto a la magnitud del conjunto de datos para ser considerado grande.

En este sentido, en Colombia, en la Ley 1122 del 2007, mediante la cual se realizan ajustes a su Sistema General de Seguridad Social en Salud, se establece que la salud pública está guiada por políticas que buscan garantizar la salud de la población mediante acciones dirigidas de manera individual y colectiva, conformando indicadores de condiciones de vida, bienestar y desarrollo del país. Adicionalmente, en Colombia se llevan a cabo *análisis de la situación de la salud*, con el propósito de definir necesidades, prioridades, políticas en salud y evaluar su

pertinencia y cumplimiento, para formular estrategias de promoción, prevención, control de daños a la salud y construir escenarios prospectivos de salud (Ministerio de Salud y Proteccion Social, 2014). Por otro lado, el Sistema de Vigilancia de Salud Pública (Sivigila) integra en la Bodega de Datos del Sistema Integrado de Información de la Protección Social (Sispro) la información relacionada con eventos de interés en salud pública, lo cual facilita su uso y consulta por parte de los diferentes usuarios.

Por consiguiente, el objetivo de esta investigación fue construir un marco de trabajo con base en el proceso de descubrimiento de conocimiento en bases de datos (KDD, por sus siglas en inglés), definido como "el proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos" (Fayyad et al., 1996, pp. 40-41; traducción de los autores). Así, con este proceso es posible preparar los datos para el análisis y, consecuentemente, para extraer patrones y modelos que, mediante un proceso de evaluación e interpretación, se convierten en conocimiento. Este marco será guía para la vigilancia epidemiológica, en las etapas de análisis e interpretación de datos y difusión de la información, con apoyo del paradigma *big data*, adoptando técnicas de minería de datos para descubrir patrones y tendencias, y de aprendizaje automático, para construir modelos predictivos, con el fin de apoyar la toma de decisiones de las autoridades de salud pública.

Considerando que un marco de trabajo está orientado a los usos del conocimiento y las herramientas existentes que proporciona una base metodológica al estudio que se esté llevando a cabo (Robertson, 2017), el objeto de este capítulo es plantear una metodología general y extrapolable que se adapte al análisis de eventos específicos y cuyas preguntas de investigación comprendan:

¿La integración de técnicas de minería de datos como agrupamiento y las técnicas de aprendizaje automático para tareas de clasificación permiten identificar patrones y tendencias de la tasa de incidencia (niveles de riesgo) del evento de interés en salud pública, durante un periodo en el departamento de Santander? ¿La metodología ajustada para el análisis, interpretación y difusión de la información apoya parte del proceso de vigilancia epidemiológica en Santander?

En este capítulo, inicialmente, se detalla el método que menciona los componentes del marco de trabajo; luego, se encuentran los resultados del estudio de caso específico, el análisis y discusión de los resultados, y, por último, las conclusiones.

Métodos

En esta sección se menciona la arquitectura *big data*, que apoya el marco de trabajo (anexo 1) y seguidamente se explican los otros componentes del marco propuesto.

Arquitectura de apoyo para la analítica de datos

En la arquitectura *big data* se integran herramientas y tecnologías gratuitas para el almacenamiento de información y para el preprocesamiento, procesamiento y visualización de la información. Se destaca el apoyo de PostgreSQL, Rstudio, Spark, Shiny y Google My Maps.

Caracterización del área de estudio

La caracterización del área de estudio está relacionada con información que presente una descripción en cuanto a división política, extensión y otras características. Se selecciona el evento crítico de salud pública en Santander y se detalla su epidemiología.

Seleccionar los eventos de salud pública

Seleccionar uno o varios eventos de salud pública críticos de Santander en los cuales se enfocará el proceso de descubrimiento de conocimiento. Los eventos de interés en salud pública son aquellos eventos que el Ministerio de Salud y Protección Social considera importantes o trascendentes para la salud, con base en criterios de frecuencia, gravedad, comportamiento epidemiológico, posibilidades de prevención, costo-efectividad de las intervenciones e interés público, y que además requieren ser enfrentados con medidas de salud pública.

Epidemiología del evento

Detallar la epidemiología del evento o los eventos objeto de estudio y su impacto en la población. Cabe resaltar que la epidemiología es el estudio de la ocurrencia y distribución de estados o eventos relacionados con la salud en poblaciones específicas, incluyendo el estudio de los determinantes que influyen en dichos estados, y la aplicación de este conocimiento al control de los problemas de salud (Porta, 2008).

Proceso de extracción de conocimiento

Las siguientes fases permiten orientar los aspectos clave para descubrir conocimiento en bases de datos, que implica la integración y recopilación de la información, así como el proceso de selección, limpieza y transformación de dicha información.

Integración y recopilación de la información

Esta fase está conformada por la identificación de las bases de datos pertinentes para el estudio, su gestión y la respectiva recopilación, teniendo en cuenta las siguientes actividades:

Definir variables. Las que se incluirán en el análisis, relacionadas con los casos del evento de interés en salud pública seleccionado y con los factores de la tabla 1. Cabe resaltar que, además de las variables mencionadas en esa tabla, es posible incluir otras como los determinantes intermedios de la salud, que se distribuyen según la estratificación social y determinan las diferencias en cuanto a la exposición y la vulnerabilidad a las condiciones perjudiciales para la salud (Organización Panamericana de la Salud, 2012).

Fuentes de datos. A partir del evento y las variables que pueden influir, identificar las bases de datos en las cuales se encuentra la información requerida y gestionar su adquisición ante las respectivas entidades. En la tabla 2 se detallan las fuentes de datos que se deben tener en cuenta.

Tabla 2. Fuentes de información

Fuente	Descripción de la información	Entidad	Observaciones
Cubos Sivigila	Número de casos por eventos de interés en salud pública	Ministerio de Salud y Protección Social	Solicitud de clave y credenciales mediante correo electrónico
Laboratorio Departamental de Salud Pública	Datos del vector *Aedes aegypti* en el departamento de Santander	Secretaría de Salud de Santander	Solicitud oficial dirigida a la Secretaría de Salud
Centro de Investigaciones en Enfermedades Tropicales	Datos entomológicos del vector *Aedes aegypti* en Bucaramanga, Piedecuesta y Floridablanca	Universidad Industrial de Santander	Solicitud por medio de correo electrónico
Grupo de Epidemiología y Demografía de la Secretaría de Salud de Santander	Información sobre eventos de salud pública	Secretaría de Salud de Santander	Solicitud oficial dirigida a la Secretaría de Salud
Instituto de Hidrología, Meteorología y Estudios Ambientales (Ideam)	Precipitaciones, temperatura baja, media y alta, humedad	Ideam	Solicitud mediante correo electrónico
Estadísticas de demografía y población: datos poblacionales	Estimación y proyección de la población nacional, departamental y municipal total por área 1985-2020. Proyecciones de la población con base en el censo del 2005	Departamento Administrativo Nacional de Estadística (DANE)	Acceso público

Fuente	Descripción de la información	Entidad	Observaciones
Divipola-DANE	Codificación de la división político-administrativa de Colombia	Geoportal-DANE	Acceso público
Marco Geoestadístico Nacional	Mapa de Santander con sus correspondientes municipios	Geoportal-DANE	Acceso público
Páginas web de municipios y del departamento	Información sobre municipios y del departamento	Páginas web de municipios y del departamento	Acceso público

Fuente: elaboración propia.

Se resalta que en la base de datos Sivigila reposa información con un rezago de aproximadamente dos años, lo cual dificulta analizar de manera oportuna información actualizada; sin embargo, se destaca que cuenta con un protocolo de validez que da confiabilidad para considerarla una fuente principal de datos. Por otro lado, se puede presentar dificultad en el momento de interpretar las variables de las bases de datos que poseen información entomológica, porque es posible que no cuenten con un diccionario de variables que oriente el entendimiento de la información. En cuanto a la información proporcionada por el Instituto de Hidrología, Meteorología y Estudios Ambientales (Ideam), se puede presentar el reto de invertir tiempo considerable en la conversión del archivo tipo texto a un archivo de datos estructurado para facilitar su análisis.

Adicionalmente, la información característica de todos los municipios de Santander no se encuentra unificada, lo que conduce a explorar la página web de cada municipio para obtener los datos relacionados con las variables de interés. Se destaca la disponibilidad y estructuración de las bases de datos del Departamento Administrativo Nacional de Estadística (DANE) y su facilidad para integrarlas con la demás información.

Cabe resaltar que no se consideran las variables relacionadas con factores socioeconómicos, como nivel de pobreza, debido a que la información no se encontraba desagregada por municipio para el periodo objeto de estudio y no fue posible que las fuentes consultadas la suministraran.

Selección, limpieza y transformación

La calidad del conocimiento descubierto no depende solamente de la técnica de análisis utilizada, sino también de la calidad de los datos analizados. Así, en esta fase se realizan actividades que corresponden a la exploración y preprocesamiento de los datos.

Consulta del cubo Sivigila. Realizar la consulta en el cubo Sivigila del Sispro (Ministerio de Salud y Protección Social, 2017) sobre el número de casos del evento de interés de salud pública por cada uno de los 87 municipios de Santander. Después, esta información se transforma en una tasa de ocurrencia del evento por cada uno de los municipios. La tasa es una medida de frecuencia de ocurrencia de un fenómeno y, específicamente en epidemiología, demografía y estadísticas vitales, corresponde a la frecuencia con la que ocurre un evento en una población definida, usualmente en un periodo específico (Porta, 2008). La tasa se obtiene utilizando la ecuación (1), que utiliza un multiplicador de una potencia de 10 que convierte la tasa de una fracción o decimal complicado a un número entero (Porta, 2008):

$$Tasa = \frac{Número\ de\ eventos\ en\ un\ periodo\ específico}{Promedio\ de\ la\ población\ durante\ el\ periodo} \times 10^{n} \qquad (1)$$

La interpretación a este cálculo es dada como un indicador de incidencia del evento, que corresponde al número de casos del evento que se desarrolla en una población durante un periodo determinado.

Construcción de atributos. A partir de los atributos originales, se transforman algunos datos para apoyar la interpretación de la información. En este caso, las edades simples de las personas que padecen el evento se transforman en etapas del ciclo vital.

Las etapas del ciclo vital permiten entender las vulnerabilidades y las oportunidades de invertir en cada etapa del desarrollo humano, destacando que las experiencias se acumulan a lo largo de la vida y que las intervenciones en una generación repercutirán en las siguientes. Se clasifican en primera infancia (0 a 5 años), infancia (6 a 11 años), adolescencia (12 a 18 años), juventud (14 a 26 años), adultez (27 a 59 años) y vejez (más de 60 años) (Ministerio de Salud y de Protección Social, 2014).

Análisis de agrupamiento. El análisis de clúster o de agrupamiento es una herramienta útil que permite descubrir grupos no conocidos previamente en los datos (Han et al., 2012). Se supone que en los datos existe un "verdadero" o un agrupamiento natural, de manera que al recopilar objetos similares en grupos, se intenta reconstruir el grupo desconocido con la esperanza de que cada agrupación encontrada represente un tipo o categoría real de objetos, donde los puntos de datos dentro de un grupo tengan alta similaridad entre sí en comparación con los puntos de datos que pertenecen a otros grupos. Sin embargo, son desconocidas la asignación de objetos a las clases y la descripción de estas clases (Doring et al., 2006; Grekousis y Thomas, 2012). El agrupamiento también es conocido como un análisis de datos exploratorio y de reducción de dimensiones. El agrupamiento exitoso depende del algoritmo utilizado; se destaca que se obtienen resultados

diferentes de cada uno de los algoritmos y la elección del mejor método de agrupamiento depende de factores como (Grekousis y Thomas, 2012):

- El entendimiento de los objetivos para llevar a cabo el proceso de agrupamiento.
- La comparación del proceso de agrupamiento utilizando varios algoritmos, con el fin de evaluar cuál sería el preferible para el proceso de analítica que se va a desarrollar.
- La aplicación de criterios para medir el rendimiento de los algoritmos.

Adicionalmente, el agrupamiento puede utilizarse para la detección de atípicos, que corresponden a aquellos valores "alejados" de algún grupo y que podrían ser más interesantes que los casos comunes (Han et al., 2012). Los métodos de agrupamiento se pueden diferenciar en el proceso para asignar datos a los grupos, es decir, qué tipo de particiones formar. Existen dos tipos de particiones de datos: duros o nítidos y suaves o difusos (Ponce Cruz, 2010).

Agrupamientos nítidos o duros. Cada dato debe asignarse exactamente a un grupo. Estos métodos clásicos producen particiones exhaustivas; pero la asignación de datos al grupo puede ser inadecuada en presencia de puntos de datos que son igualmente distantes a dos o más grupos. Una partición dura obliga a la asignación completa de tales puntos de datos a uno de los grupos (Doring et al., 2006).

En este tipo de agrupamiento, el más utilizado es el K-medias (en inglés, *k-means*), conocido por la sensibilidad ante datos atípicos, pero es eficiente en términos del tiempo computacional empleado para el análisis de la información (Azar et al., 2013).

K-medias tiene variantes (Nidheesh et al., 2017), con objetivos diferentes y, por tanto, resultados diferentes (Morissette y Chartier, 2013). Dentro de estas se encuentran: algoritmo de Lloyd (1982), algoritmo de Forgy (1965) (Hamerly y Drake, 2015; Kapageridis, 2015), algoritmo de MacQueen (Hamerly y Drake, 2015; Macqueen, 1967) y algoritmo de Hartigan y Wong (1979).

Agrupamientos suaves o difusos. Este tipo de algoritmos surge de la necesidad de resolver una deficiencia del agrupamiento exclusivo, el cual considera que cada elemento se puede agrupar inequívocamente con los elementos de su grupo y que, por lo tanto, no se asemeja al resto de los elementos. Tras la introducción de la lógica difusa por Zadeh, en 1965, surge una solución para este problema, caracterizando la similitud de cada elemento a cada uno de los grupos (Bezdek et al., 1984). Es decir, el análisis de conglomerados difusos permite asociaciones graduales de puntos de datos a grupos en el intervalo [0, 1], lo que da la flexibilidad para expresar que los puntos de datos pertenecen a más de un grupo al mismo tiempo (Doring et al., 2006), y el concepto de estos grados de pertenencia está respaldado por la teoría de conjuntos difusos (Azar et al., 2013). Dentro de los algoritmos de agrupamiento difuso se encuentran el K-medias difuso y el Gustafson-Kessel (Grekousis y Thomas, 2012).

Índices de validación del agrupamiento. Cómo métodos de validación se tienen en cuenta índices que definen la compactación (los miembros de cada uno de los grupos deberían estar tan cerca de cada uno de los otros miembros como sea posible) y la separación (los grupos deberían estar ampliamente separados). La primera validación utiliza solo la matriz de pertenencia, pero la segunda validación emplea la matriz de pertenencia junto con el conjunto de datos crudos. Los índices de validación más utilizados son:

- Coeficiente de partición (PC, por sus siglas en inglés): es una medida del solapamiento entre los grupos (Azar et al., 2013; Georgiev et al., 2016).
- Coeficiente de entropía (CE, por sus siglas en inglés): mide el nivel difuso de la partición de los grupos (Azar et al., 2013; Georgiev et al., 2016).
- Índice de Xie y Beni (XB): cuantifica la tasa de la variación total dentro de los grupos y la separación de los grupos (Azar et al., 2013; Xie y Beni, 1991). Para grupos mejor separados y más compactos, se obtienen valores mínimos del índice de Xie y Beni. A medida que aumenta el número de agrupaciones, el índice decrece. Desgraupes (2017) muestra este índice para el caso del agrupamiento nítido.
- Índice de silueta en agrupamiento nítido (Campello y Hruschka, 2006; Desgraupes, 2017): el intervalo del índice se encuentra entre −1 y 1, donde valores cercanos a 1 indican que el punto está ubicado en el grupo correcto, y en el caso de que el valor se acerque a −1 indica que el punto debería pertenecer a otro grupo.
- Silueta difusa (en inglés, *fuzzy silhouette* [FS]): permite revelar las regiones con altas densidades de datos, al poner el relieve en la importancia de los datos concentrados en la vecindad de los prototipos (centroides) del grupo, y, al mismo tiempo, busca reducir la importancia de los individuos que se encuentran en áreas superpuestas (Campello y Hruschka, 2006).

Procesamiento, evaluación e interpretación de la información

El objetivo de esta fase es descubrir nuevo conocimiento para apoyar la toma de decisiones. Se busca un modelo con base en los datos recopilados que describa los patrones y las tendencias de los datos y que, a la vez, permita tener un mejor entendimiento de estos o explicar situaciones pasadas. De acuerdo con el tipo de datos y con apoyo del *software* R y del paquete *sparklyr* para interactuar con el marco de trabajo de Spark, se obtiene un modelo utilizando técnicas de aprendizaje automático para tareas de clasificación, con el fin de evaluar el riesgo del evento, teniendo en cuenta que este provee el riesgo de que ocurra un evento en condiciones específicas (Corley et al., 2014). Para esto se relacionan los niveles de incidencia del evento y el comportamiento de variables que permitan descubrir patrones para caracterizar estas etiquetas (niveles de incidencia); así mismo, se

identifican aquellas variables importantes que más influyen en la generación de la clasificación.

Los algoritmos recomendados para el proceso de clasificación son los mencionados en el anexo 2 y se seleccionan para su aplicación, por lo menos, dos que permitan involucrarse en el proceso de análisis. Dentro de estos algoritmos se destaca el desempeño de los árboles de decisión, los bosques aleatorios y los métodos de regresión con fines de clasificación. Después del ajuste de cada uno de los modelos utilizando métodos como validación cruzada para identificar sus mejores parámetros, se selecciona el modelo que tenga mejor desempeño en las métricas de evaluación como exactitud, precisión, recuperación, sensibilidad y especificidad.

Difusión y uso

Realizar la visualización de la información para mostrar los resultados más importantes del estudio. Además, documentar los resultados con el fin de apoyar a las autoridades de salud pública municipal y departamental.

Resultados

De acuerdo con el marco de trabajo planteado en la anterior sección, a continuación se ajusta la metodología al caso de estudio de un evento de interés en salud pública del departamento de Santander.

Arquitectura big data

A partir de la interacción entre el gestor de bases de datos PostgreSQL, RStudio Server y Spark, se realiza la integración, el preprocesamiento y el procesamiento de la información, y mediante la plataforma shyniapp.io se visualizan los resultados del estudio.

Caracterización del área de estudio

El área de estudio es Santander, uno de los 32 departamentos que conforman la República de Colombia. Según los *Lineamientos y directrices de ordenamiento territorial para el departamento de Santander al 2030*, adoptados por el Decreto 264 del 20 de agosto del 2014, emitido por la Gobernación de Santander, el departamento tiene una extensión de 30 537 km^2, esto es, participa con el 2,7 % de la extensión territorial nacional y con el 40 % de la región nororiental. Cuenta con diversos pisos térmicos con alturas entre 100 y 4200 m s. n. m., y las temperaturas promedio oscilan entre los 4 y los 28 °C. Santander está conformado por 87 municipios (Gobernación de Santander, 2014; Esquiaqui Felipe, 2016).

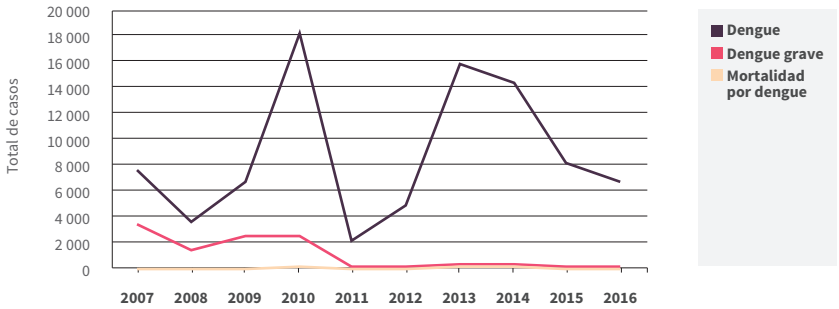
Selección del evento de salud pública

En el departamento de Santander, del total de casos de eventos de salud pública reportados en el periodo 2007-2016, el mayor evento ocurrido es el dengue, con el 33,26 % (88 421); seguido por la varicela, con el 15,66 % (41 627), y el dengue grave, que se encuentra en la quinta posición, con el 3,93 % (10 458). A nivel nacional, en este mismo periodo (2007-2016), en cuanto a casos de dengue, Santander se ubica como el segundo departamento en el cual ocurren la mayoría de los casos (11,55 %), después del departamento del Valle del Cauca, en el que ocurre el 15,02 % (114 957 casos).

En los municipios de Santander el evento del dengue varía entre 2 y 31 255 casos, la mayoría de los cuales se presenta en la ciudad de Bucaramanga (35,35 %), y el mínimo, en los municipios Macaravita y Albania (0,0023 %). Con respecto a los casos de dengue grave, el municipio donde ocurren la mayoría de los casos (39,83 %) es Bucaramanga, y el menor número de casos, en los municipios Macaravita, San Andrés, El Peñón, Molagavita, Bolívar, San Miguel, Palmar y Carcasí, cada uno con el 0,01 % (1) de los casos. Adicionalmente, el número total de casos de mortalidad por dengue en el periodo 2007-2016 ha sido de 103, presentándose el 33,01 % (34) de estos en el municipio de Bucaramanga.

En la figura 1 se evidencia que el comportamiento del dengue es variable: el menor número de casos (2053) ocurrió en el 2011, y dos picos de mayor presencia correspondieron a los años 2010 (18 417 casos) y 2013 (15 892 casos). Se observa una tendencia creciente desde el 2008 hasta el 2010 y del 2011 al 2013, y una tendencia decreciente en la presencia de casos desde el 2013 al 2016. En relación con el dengue grave, se observa que disminuye significativamente del 2010 (23,9 %) al 2011 (1,06 %); sin embargo, del 2011 al 2014 hay una tendencia creciente, pero para los años 2015 y 2016 se presenta la menor cantidad de casos, 44 y 69 casos, respectivamente.

Figura 1. Comportamiento del dengue, dengue grave y mortalidad por dengue, periodo 2007-2016



Fuente: cubo Sivigila de la base de datos Sispro.

De acuerdo con la información detallada, se ajusta el marco de trabajo propuesto al análisis del dengue, con el fin de explorar patrones y tendencias que apoyen a los tomadores de decisiones en salud pública tanto a nivel departamental como nacional.

Epidemiología del dengue

El dengue es una infección vírica transmitida por mosquitos hembra principalmente de la especie *Aedes aegypti* y, en menor grado, de *Aedes albopictus*. Estos mosquitos también transmiten la fiebre chikunguña, la fiebre amarilla y la infección por el virus de zika. El dengue se presenta en los climas tropicales y subtropicales, sobre todo en zonas urbanas y semiurbanas (OMS, 2017).

Se conocen cuatro serotipos del arbovirus causante del dengue (DEN-1, DEN-2, DEN-3 y DEN-4), y es la virosis humana más importante transmitida por artrópodos (Organización Panamericana de la Salud, 2015). Cuando una persona se recupera de la infección, adquiere inmunidad de por vida contra el serotipo en particular. Sin embargo, la inmunidad cruzada a los otros serotipos es parcial y temporal. Las infecciones posteriores causadas por otros serotipos aumentan el riesgo de padecer el dengue grave (Instituto Nacional de Salud, 2017; World Health Organization, 2017).

Proceso de extracción de conocimiento

Integración y recopilación de la información

De acuerdo con la tabla 3, las variables y las fuentes de información que se incluyeron en el estudio son las siguientes:

Tabla 3. Variables del caso de estudio

Variable	Unidades
Casos de dengue (Minsalud, 2017)	Número de casos de incidencia
Características asociadas a los casos: municipio, área de ocurrencia, edad, sexo, régimen de afiliación a salud y periodo epidemiológico (Minsalud, 2017)	Número de casos de incidencia
Información geográfica (Gobernación de Santander, 2017)	Latitud y longitud de cada municipio
Precipitaciones (Ideam, 2017)	Milímetros
Datos poblacionales (DANE, 2017)	Habitantes por año a nivel: total, cabecera, área rural dispersa y centro poblado.

Variable	Unidades
Mapa para georreferenciación (Geoportal-DANE, 2017)	*Shapefile* con coordenadas geográficas
Altitud (Gobernación de Santander, 2017)	Metros sobre el nivel del mar (m s. n. m.)

Minsalud: Ministerio de Salud y Protección Social.

Fuente: elaboración propia.

Selección, limpieza y transformación

Transformación de variables. Para efectos de los análisis, los números de casos de ocurrencia del dengue en el departamento de Santander durante el periodo 2007-2016 se transforman a tasas de ocurrencia del evento por 100 000 habitantes. Este cálculo se realiza para cada uno de los municipios teniendo en cuenta la estimación y la proyección de población nacional, departamental y municipal total por área 1985-2020, elaborada por el DANE.

Cálculo de índices de validación interna del agrupamiento. Los algoritmos K-medias difuso y algoritmo de Gustafson-Kesel, para la aplicación de agrupamiento difuso, y los algoritmos de Hartigan, MacQueen, Lloyd y Floyd para aplicar el agrupamiento K-medias, se validan mediante índices que se procesan con apoyo del *software* R y Spark; por otro lado, se utiliza Google My Maps para visualizar las agrupaciones de municipios. El análisis se efectúa para cada año del periodo 2007-2016.

De acuerdo con resultados de los índices, para cada uno de los agrupamientos de K-medias y difuso se elige el mejor algoritmo asociado a valores que tiendan a cero en los índices PE y XB, y valores que tiendan a 1 en los índices PC y silueta. Según el anexo 3, el mejor algoritmo de agrupamiento para cada uno de los años es el K-medias difuso. Entonces, considerando sus resultados, en la tabla 4 se muestra la proporción de municipios que conforman cada una de las agrupaciones obtenidas por cada año. En la tabla 5 se presentan las etiquetas asociadas con los niveles de incidencia promedio del dengue como: muy alta, alta, media, baja y muy baja. Estas etiquetas son establecidas a partir del cálculo de los quintiles de las incidencias promedio anual de dengue por 100 000 habitantes de todas las agrupaciones.

Tabla 4. Agrupaciones de municipios por año del periodo 2007-2016

Año	Casos de dengue por 100 000 habitantes	Total de casos	Grupo	Porcentaje municipios por grupo	Promedio de incidencia anual de dengue por 100 000 habitantes	Nivel de incidencia según el periodo 2007-2016
2007	254,94	113	1	19,54	400,32	Baja
			2	80,46	43,58	Muy baja
2008	236,85	3664	1	82,76	29,74	Muy baja
			2	17,24	297,39	Media
2009	285,53	6616	1	12,64	449,75	Alta
			2	87,36	40,21	Muy baja
2010	434,58	18 417	1	64,37	128,38	Media
			2	19,54	643,97	Alta
			3	16,09	1226,31	Muy alta
2011	114,02	2053	1	63,22	16,64	Muy baja
			2	36,78	135,97	Media
2012	60,54	4797	1	73,56	39,2	Muy baja
			2	26,44	322,24	Media
2013	405,8	15 892	1	14,94	1391,89	Muy alta
			2	21,84	413,05	Alta
			3	52,87	70,39	Baja
			4	10,34	958,69	Muy alta
2014	91,99	14 313	1	27,59	753,61	Muy alta
			2	72,41	107,21	Baja
2015	75,89	8192	1	25,29	672,19	Muy alta
			2	74,71	89,5	Baja
2016	113,29	6865	1	24,14	637,06	Alta
			2	75,86	101,54	Baja

Fuente: elaboración propia.

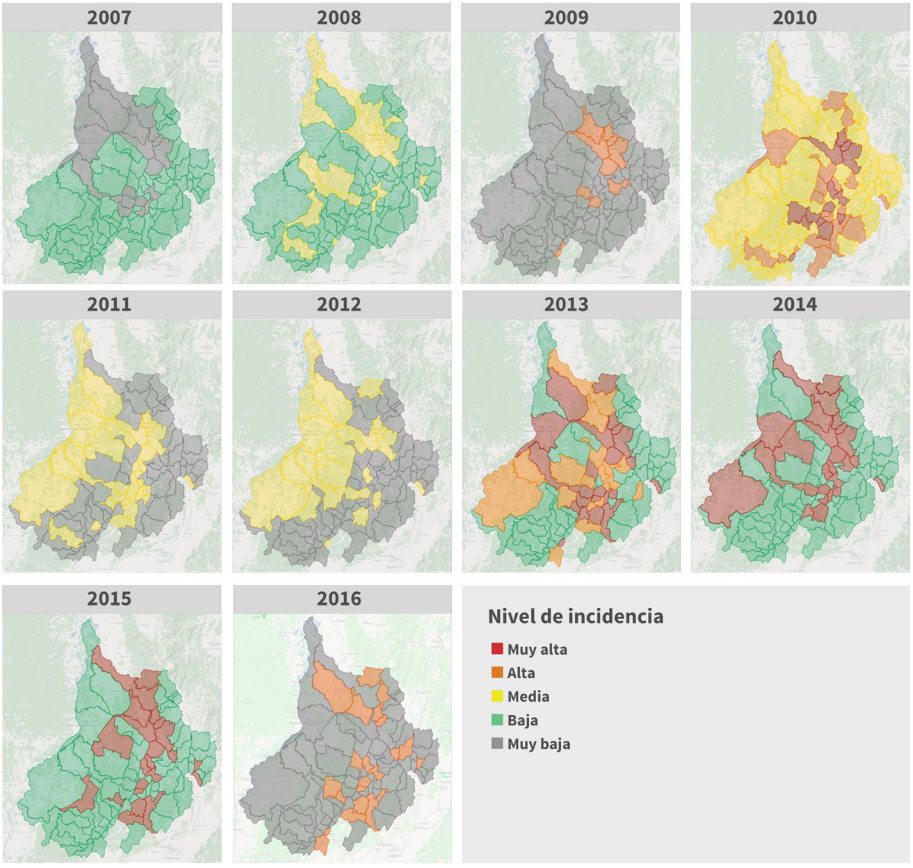
Tabla 5. Niveles de incidencia

Quintil	Incidencia promedio de dengue por 100 000 habitantes	Nivel de incidencia
1	0-54,304	Muy baja
2	54,304-124,146	Baja
3	124,146-400,23	Media
4	400,23-661,16	Alta
5	661,16-1391,89	Muy alta

Fuente: elaboración propia.

En la tabla 4 se muestra que el 2007 presenta agrupaciones de municipios con niveles de incidencia de dengue muy bajos y bajos; en contraste con los años 2008, 2011 y 2012, que además de tener niveles de incidencia muy bajos o bajos también muestran una incidencia en categoría media. Para los años 2009, y del 2014 al 2016, hay niveles de incidencia tanto bajos como altos o muy altos. Los años 2010 y 2013 registran el mayor promedio de incidencia de dengue por 100 000 habitantes y se caracterizan por niveles de incidencia media, alta y muy alta (figura 2). De acuerdo con los niveles de incidencia, en el anexo 4, se muestra el perfil que caracteriza a cada uno de ellos.

Figura 2. Representación de los niveles de incidencia de dengue en Santander, periodo 2007-2016



Fuente: mapas construidos con Google My Maps (2007-2009, 2010-2012, 2013-2015 y 2016).

Procesamiento, evaluación e interpretación de la información

Para el análisis del aprendizaje automático, se consideraron 47 variables predictoras relacionadas con el número de casos del evento por periodo epidemiológico, atributos sociodemográficos, características geográficas, precipitaciones y la variable dependiente, que corresponde a las etiquetas de niveles de incidencia construidas a partir del proceso previo de agrupamiento. Las variables de los modelos de clasificación con sus respectivos atributos son:

- Municipio: nombre del municipio.

- Número de casos: número de casos del evento.
- Tasa de casos: incidencia de casos del evento por 100 000 habitantes.
- Año de ocurrencia: periodo 2007-2016.
- Área de ocurrencia: cabecera; área rural dispersa; centro poblado.
- Edad: en quinquenios.
- Sexo: masculino; femenino.
- Régimen de afiliación: contributivo; subsidiado; excepción; especial; no afiliado; no definido.
- Periodo epidemiológico: conformado por cuatro semanas epidemiológicas. El calendario epidemiológico incluye 52 semanas epidemiológicas y en total son 12 o 13 periodos epidemiológicos.
- Población: área cabecera; resto (centro poblado y área rural dispersa).
- Altitud media: en m s. n. m.
- Coordenadas: latitud y longitud.
- Precipitación: total (mms) y máxima (mms).
- Nivel de incidencia: muy alta, alta, media, baja, muy baja.

Los modelos de aprendizaje automático supervisados para clasificación se obtienen con apoyo del *software* R y las funciones de *sparklyr*, con el fin de elegir, a partir de árboles de decisión, bosques aleatorios y regresión logística, y se evalúa cuál es el más adecuado o que tiene la capacidad de desempeñarse mejor en las diferentes métricas como exactitud, precisión, recuperación, sensibilidad y especificidad. Inicialmente, para cada uno de los métodos de clasificación se compara la exactitud lograda entre el modelo en el que se seleccionan todas las variables consideradas y otro en el que no se consideran la latitud ni la longitud de los municipios. Para la optimización de los parámetros de los modelos se utiliza validación cruzada repetida 10-fold, y para evaluar la capacidad predictiva del modelo se considera un 80 % de los datos para el conjunto de entrenamiento y el restante 20 % para el conjunto de prueba.

De acuerdo con los resultados obtenidos, a efectos de seleccionar los parámetros adecuados para cada uno de los modelos, en la tabla 6 se observa que en cuanto a las métricas globales de los modelos a un nivel de confianza del 95 %, la mayor exactitud se obtiene en el clasificador de regresión logística multinomial, lo que lleva a deducir que este sería el más adecuado, aunque sea un valor bajo de exactitud.

Tabla 6. Resultados de los clasificadores

Algoritmo	Parámetros	Modelo seleccionado	Exactitud	Intervalo de confianza Exactitud (95 %)	Valor de *p*
Árboles de decisión	Gini; entropía Factor de costo de complejidad (cp); minsplit	Minsplit = 5; cp = 0,009859155; Información	57,43	(47,44648; 66,45201)	$4{,}470037 \times 10^{-5}$
Bosque aleatorio	Mtry ntree nodezise	Mtry = 37; ntree = 82; nodesize = 6	58,93	(49,24; 68,14)	$8{,}654 \times 10^{-6}$
Regresión logística múltiple	Decay	Decay = 0,1	59,82	(50,14; 68,97)	$3{,}612 \times 10^{-6}$

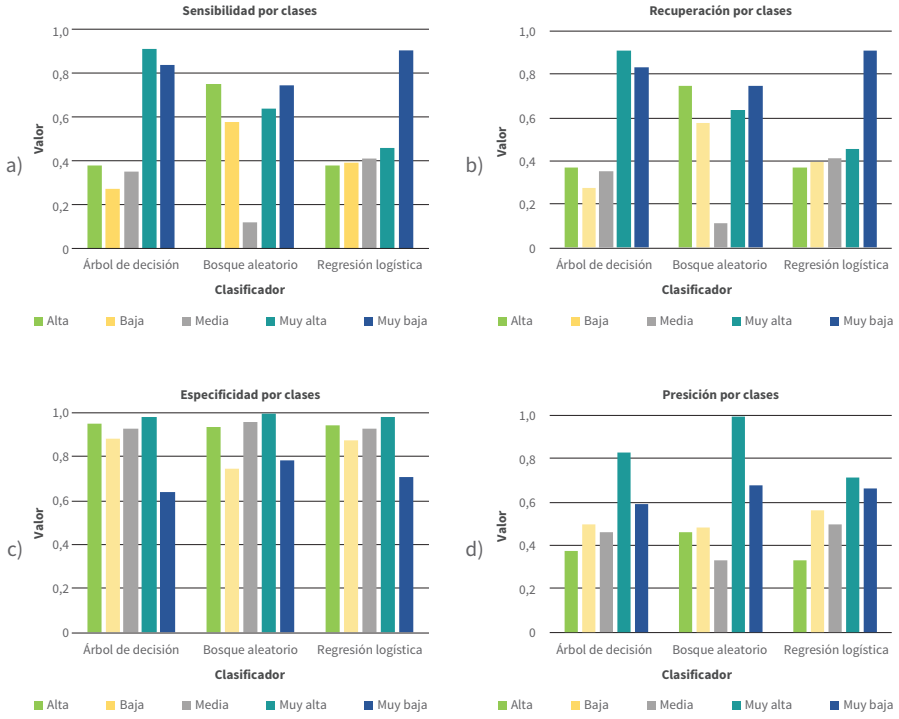
Fuente: elaboración propia.

Sin embargo, se procede a revisar las métricas obtenidas para cada una de las clases de niveles de incidencia analizando métricas como sensibilidad y especificidad, precisión y recuperación, teniendo en cuenta que las clases no están balanceadas, es decir, las proporciones de presencia de clases en los datos difieren: muy alta (9,42 %), alta (7,81 %), media (14,48 %), baja (29,54 %) y muy baja (38,73 %). Cabe resaltar que seleccionar un clasificador que catalogue las clases de incidencia muy alta o alta, como baja o muy baja, es un hecho que implica altos costos para los planes de salud pública, porque se puede llegar a invertir pocos recursos en planes preventivos en zonas que posiblemente necesiten otras estrategias para mitigar la presencia del evento.

Por tanto, al analizar la sensibilidad o recuperación de las clases en cada uno de los clasificadores (figura 3), se muestra que la proporción de observaciones en la clase media son bien clasificadas por la regresión logística multinomial (41,2 %). Así, de las observaciones de incidencia muy baja, el 90,7 % son clasificados como pertenecientes a esta. Sin embargo, para la clase de incidencia muy alta, el mejor clasificador es el árbol de decisión, con una sensibilidad o recuperación de aproximadamente el 91 %, aunque este algoritmo no tiene mejor desempeño en la clasificación de las demás etiquetas, en comparación con la regresión logística multinomial. En relación con el bosque aleatorio, hay un pobre desempeño en clasificar las observaciones con nivel de incidencia media, en comparación con los otros algoritmos; pero, en cuanto a las clases de incidencia muy baja, muy

alta, baja y alta, el modelo ajustado es capaz de clasificar el 74,4 %, el 63,6 %, el 75 % y el 57,6 %, respectivamente, como pertenecientes a esta clase.

Figura 3. Métricas de desempeño



Nota: a) sensibilidad por clases; b) recuperación por clases; c) especificidad por clases; d) precisión por clases.

Fuente: construido por los autores en RStudio.

En cuanto a la especificidad (figura 3) se muestra que, aun cuando todos los algoritmos ajustados poseen una exactitud entre el 57 % y el 60 %, los valores de especificidad son mayores al 60 %, lo cual significa que los algoritmos reconocen muy bien los datos que no pertenecen a la clase que se está analizando. Por ejemplo, si se revisa la incidencia muy baja, el bosque aleatorio clasifica bien el 78,3 % de las observaciones no pertenecientes a este nivel de incidencia, como incidencia baja, media, alta y muy alta. Así, con el bosque aleatorio ajustado se obtienen valores más altos para esta métrica en la clasificación de los niveles de incidencia muy baja, muy alta y media; pero tiene el más bajo desempeño en la clase de nivel de incidencia baja (74,7 %), y el árbol de decisión da el valor más alto al evaluar las clasificaciones de incidencia alta.

Por otro lado, en cuanto a la precisión (figura 3) de los modelos ajustados, se muestra que el bosque aleatorio tiene alta precisión en la clasificación de niveles de incidencia muy alta, alta y muy baja, en comparación con los otros algoritmos. Sin embargo, en cuanto a la clasificación de las clases baja y media hay valores más bajos que los otros algoritmos.

Por consiguiente, de acuerdo con los resultados de métricas de desempeño diferentes a la exactitud, el modelo ajustado que se podría adoptar es el de bosque aleatorio, dado que se destaca en su desempeño respecto a la clasificación de clases de incidencia alta, muy alta, y baja y muy baja. Ello permite considerar que las etiquetas asignadas por este clasificador generan confianza, en cuanto a los planes estratégicos que se van a orientar en salud pública, porque este clasificador identifica las clases de incidencia altas y muy altas, que de no clasificarse bien traería como consecuencia el incurrir en más costos a largo plazo, porque no se invertiría en planes preventivos hacia situaciones de alto riesgo, sino que se implementarían planes preventivos como si fuese una zona de bajo riesgo. Además, en el caso de la clasificación de etiquetas de incidencia baja y muy baja, esta apoyaría con alta confianza a las autoridades de salud pública.

Aunque es poco destacable el modelo de bosque aleatorio en la clasificación de niveles de incidencia media, al analizar este nivel de incidencia se presenta que tiene alto desempeño en la clasificación de las observaciones que no pertenezcan a esta clase.

Por otro lado, al analizar la importancia de las variables, en los árboles de decisión, las de mayor influencia son sexo (femenino y masculino), régimen de afiliación subsidiado y los quinquenios de edad 10 a 14, 15 a 19 y 35 a 39 años; además, el área de ocurrencia cabecera presenta una influencia destacable en el modelo ajustado de árbol de decisión. El modelo de bosque aleatorio (figura 4), utilizando el índice de Gini, que está relacionado con la disminución total en las impurezas de los nodos al dividirse en la variable en específico, promediando todos los árboles, muestra que las variables predictoras de mayor importancia son el sexo (femenino y masculino), similar al árbol de decisión, el régimen de afiliación contributivo, la precipitación total (en milímetros), la población del área de cabecera y centro poblado. Así mismo, el quinquenio de edad más importante es de 10 a 14 años, el área de ocurrencia cabecera y la altitud (m s. n. m.) de los municipios.

Figura 4. Importancia de variables en cada modelo



Nota: a) importancia de las variables en árbol de decisión; b) importancia de las variables en bosque aleatorio; c) importancia de las variables en la regresión logística multinominal.

Fuente: elaboración propia en RStudio.

El modelo ajustado de regresión logística multinomial muestra que las variables predictoras más importantes son los periodos epidemiológicos 1, 4, 2, 7, 3, 13 y 5, así como la población del área cabecera; los quinquenios de edad de 40 a 44, de 70 a 74 y de 75 a 79, y el régimen de afiliación de excepción (figura 4).

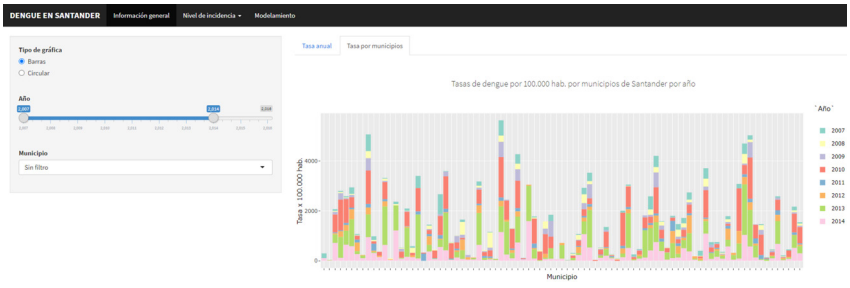
Visualización de la información

Los resultados de la información objeto de estudio se utilizan para realizar su visualización, que está conformada por tres secciones (figura 5):

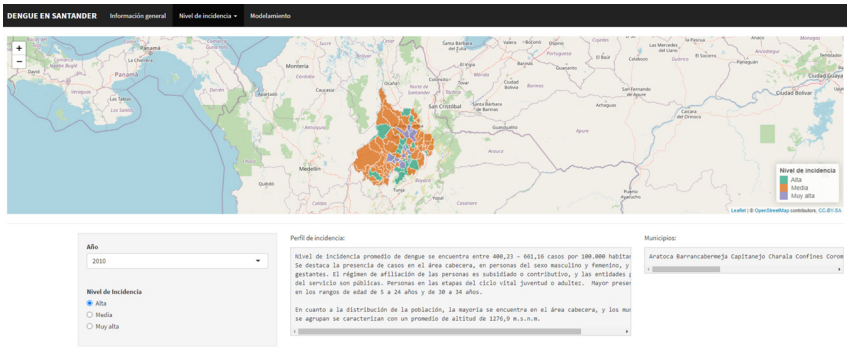
- Sección "Información general": se encuentran gráficas y tablas que resumen las tasas de dengue en el periodo 2007-2016, en el departamento de Santander.
- Sección "Nivel de incidencia": se encuentran los resultados del agrupamiento en un mapa en el cual se muestran los niveles de incidencia en cada municipio por cada uno de los años del periodo 2007-2016.
- Sección "Modelamiento": se encuentra información sobre los resultados de los modelos de aprendizaje automático.

Figura 5. Visualización de la información

a)



b)



c)



Nota: a) sección "Información general"; b) sección "Nivel de incidencia"; c) sección "Modelamiento".

Fuente: elaboración propia a partir de *Dengue en Santander* (s. f.).[1]

Análisis y discusión

En Santander (Colombia), los casos de dengue han experimentado un comportamiento variable, pues en algunos años hay poca presencia de este evento, pero en otros la presencia de casos aumenta de manera considerable. En este estudio, se analiza la distribución del dengue utilizando un algoritmo de agrupamiento difuso en el departamento de Santander durante el periodo 2007-2016. Las mayores tasas de incidencia promedio de dengue son 1391,89 casos y 1226,31 casos de dengue por 100 000 habitantes, y corresponden a los grupos 3 y 1 de los años 2010 y 2013, respectivamente. Además, en estos grupos se encuentran cerca del 15 % de los municipios del departamento. Se evidencia que en los niveles de incidencia alta y muy alta hay grupos conformados por una pequeña proporción de municipios, en comparación con los grupos de municipios con niveles de incidencia muy baja. Por otro lado, de los resultados obtenidos se identifica que la presencia de dengue era más alta en hombres que en mujeres —similar a lo encontrado en otro estudio, en el que se expresaba que esto podría deberse a que los hombres están más expuestos a las picaduras de mosquitos en comparación con las mujeres y a que, tal vez, los hombres no utilizan ropa para protegerse y tampoco repelentes, y esto se podría dar por diferencias en patrones de comportamiento y por el estilo de vida— (Mutheneni et al., 2018).

Además del análisis de la importancia de variables en los modelos de aprendizaje automático, se muestra que variables como régimen de afiliación, edad, sexo, precipitaciones de los municipios son determinantes para caracterizar el nivel de incidencia del dengue y ante esto Rodríguez-Morales et al. (2017)

1 Véase https://proyectodengue.shinyapps.io/Dengue_en_Santander/

mencionan que el zika y otras enfermedades arbovirales podrían estar asociadas con múltiples factores, incluyendo cambios en las condiciones ambientales y sociales que posiblemente influyen en su caracterización epidemiológica.

Varias son las fortalezas de este estudio: analizar información de un periodo de diez años, lo cual indica que los modelos de analítica de datos (como el agrupamiento) permiten obtener resultados importantes para asignar unas etiquetas de incidencia: muy alta, alta, media, baja y muy baja, a los municipios por cada uno de los años y ver a través del periodo evaluado los cambios en estos niveles de incidencia por medio de representaciones en mapas. Esta información sobre la caracterización de los perfiles de los niveles de incidencia, el ajuste de los modelos de aprendizaje automático para predecir el nivel de incidencia del dengue en cualquier municipio de Santander y la relación de las variables más importantes es información que evidencia el soporte destacable de otros métodos de analítica a la vigilancia en salud pública del dengue.

Debido a la importancia tanto de la vigilancia en salud pública como de los sistemas que respaldan ese proceso, este marco de trabajo propuesto no solamente es útil para las entidades departamentales de salud pública, sino también para las autoridades municipales y grupos de investigación, porque les permitirá disponer de una guía clara y objetiva que conduzca a la adecuada adopción de los paradigmas que aquí se tratan y de diferentes métodos y fases vitales para obtener conocimiento oportuno y de calidad, con el fin de implementar estrategias que mejoren el bienestar de la comunidad objeto de estudio y, por tanto, el bienestar de las comunidades aledañas.

Además, este marco permitirá evaluar, de acuerdo con los resultados obtenidos, los recursos invertidos en planes de prevención y promoción de la salud que busquen mitigar el impacto de esta enfermedad arboviral. Por otro lado, a diferencia del marco de trabajo propuesto por Ali et al. (2016), este aplica tanto para enfermedades infecciosas como para enfermedades no transmisibles, cuyo fin principal es analizar el riesgo del evento de interés en una población específica. Y aunque no tiene en cuenta la captura de datos de fuentes primarias ni en tiempo real, se podrían adicionar complementos que permitan mejorar lo propuesto. Cabe destacar que los resultados deben ser interpretados de acuerdo con el análisis de la información objeto de estudio y que el objetivo en común es analizar la evolución de la incidencia del evento de interés en salud pública a través de un periodo, así como identificar las variables o factores de riesgo que más influyen en la incidencia del evento.

Conclusiones

El agrupamiento permite caracterizar los niveles de incidencia del dengue e identificar las variables que más los diferencian. Así, esta clasificación es una fuente

de información para la toma de decisiones relacionadas con las posibles acciones que se deben orientar para mitigar la presencia de este evento, así como la evaluación de los planes de prevención actuales, con el fin de contribuir a la salud de la población desde las diferentes iniciativas por parte de las entidades de salud pública.

Adicionalmente, las etiquetas asignadas por el clasificador bosque aleatorio generarían confianza en cuanto a los planes estratégicos que van a orientar la salud pública, porque permite identificar las clases de incidencia altas y muy altas, que de no clasificarse bien traerían como consecuencia el incurrir en más costos a largo plazo, porque no se invertiría en planes preventivos hacia situaciones de alto riesgo, sino que se implementarían planes preventivos en zonas catalogadas erróneamente como de bajo riesgo. Además, en el caso de la clasificación de etiquetas de incidencia baja y muy baja, está apoyaría con alta confianza a las autoridades de salud pública.

El estudio de caso evidenció que, a partir del marco de trabajo propuesto y siendo este una guía de apoyo específico para el análisis de datos que involucra métodos de minería de datos y de aprendizaje automático bajo el paradigma *big data*, se destaca la importancia de integrar el agrupamiento a nivel descriptivo con modelos de aprendizaje automático, para identificar patrones y tendencias con los cuales entender el evento específico que se va a analizar y las variables que tienen más importancia o influencia en el nivel de incidencia o riesgo del evento en una población específica.

Agradecimientos

Agradecimientos al Ministerio de Salud y Protección Social y al Ideam, por compartir bases de datos importantes con el fin de integrarlas dentro del repositorio objeto de estudio de este proyecto.

Referencias

Ali, M. A., Ahsan, Z., Amin, M., Latif, S., Ayyaz, A. y Ayyaz, M. N. (2016). ID-Viewer: A visual analytics architecture for infectious diseases surveillance and response management in Pakistan. *Public Health*, *134*, 72-85. https://doi.org/10.1016/j.puhe.2016.01.006

Amato-Gauci, A. y Ammon, A. (2008). The surveillance of communicable diseases in the European Union: A long-term strategy (2008-2013). *Eurosurveillance*, *13*(26), 1-3. https://www.eurosurveillance.org/content/10.2807/ese.13.26.18912-en

Azar, A. T., El-Said, S. A. y Hassanien, A. E. (2013). Fuzzy and hard clustering analysis for thyroid disease. *Computer Methods and Programs in Biomedicine*, *111*(1), 1-16. https://doi.org/10.1016/j.cmpb.2013.01.002

Bezdek, J. C., Ehrlich, R. y Full, W. (1984). FCM: The fuzzy *c*-means clustering algorithm. *Computers & Geosciences*, *10*(2-3), 191-203. https://doi.org/10.1016/0098-3004(84)90020-7

Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., … Hay, S. I. (2013). The global distribution and burden of dengue. *Nature*, *496*(7446), 504-507. https://doi.org/10.1038/nature12060

Broniatowski, D. A., Paul, M. J. y Dredze, M. (2013). National and local influenza surveillance through twitter: An analysis of the 2012-2013 influenza epidemic. *PLoS ONE*, *8*(12), e83672. https://doi.org/10.1371/journal.pone.0083672

Buczak, A. L., Koshute, P. T., Babin, S. M., Feighner, B. H. y Lewis, S. H. (2012). A data-driven epidemiological prediction method for dengue outbreaks using local and remote sensing data. *BMC Medical Informatics and Decision Making*, *12*(1), 124. https://doi.org/10.1186/1472-6947-12-124

Campello, R. J. G. B. y Hruschka, E. R. (2006). A fuzzy extension of the silhouette width criterion for cluster analysis. *Fuzzy Sets and Systems*, *157*(21), 2858-2875. https://doi.org/10.1016/j.fss.2006.07.006

Chen, M., Ma, Y., Song, J., Lai, C.-F. y Hu, B. (2016). Smart clothing: Connecting human with clouds and big data for sustainable health monitoring. *Mobile Networks and Applications*, *21*(5), 825-845. https://doi.org/10.1007/s11036-016-0745-1

Connolly, M. A., Gayer, M., Ryan, M. J., Salama, P., Spiegel, P. y Heymann, D. L. (2004). Communicable diseases in complex emergencies: Impact and challenges. *Lancet*, *364*, 1974-1983.

Corley, C. D., Pullum, L. L., Hartley, D. M., Benedum, C., Noonan, C., Rabinowitz, P. M. y Lancaster, M. J. (2014). Disease prediction models and operational readiness. *PLoS ONE*, *9*(3), 1-9. https://doi.org/10.1371/journal.pone.0091989

Departamento Administrativo Nacional de Estadística (DANE). (2017). *Estimación y proyección de población nacional, departamental y municipal total por* área *1985-2020.* https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacion

Desgraupes, B. (2017). *ClusterCrit: Clustering indices.* CRAN Package. https://cran.r-project.org/web/packages/clusterCrit

Doring, C., Lesot, M.-J. y Kruse, R. (2006). Data analysis with fuzzy clustering methods. *Computational Statistics & Data Analysis*, *51*, 192-214. https://doi.org/10.1016/j.csda.2006.04.030

Esquiaqui Felipe, R. E. (2016). *Análisis de situación de salud con el modelo de los determinantes sociales de salud del departamento de Santander, 2015*. Alcaldía de Bucaramanga y Secretaría de Salud y Ambinete de Bucaramanga.

Fayyad, U., Piatetsky-Shapiro, G. y Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, *17*(3), 37-54. https://doi.org/10.1609/aimag.v17i3.1230

Forouzanfar, M. H., Afshin, A., Alexander, L. T., Biryukov, S., Brauer, M., Cercy, K., … Zhu, J. (2016). Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: A systematic analysis for the Global Burden of Disease Study 2015. *Lancet*, *388*(10053), 1659-1724. https://doi.org/10.1016/S0140-6736(16)31679-8

GBD 2016 Risk Factors Collaborators. (2017). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2016: A systematic analysis for the Global Burden of Disease Study 2016. *Lancet*, *390*(10100), 1345-1422. https://doi.org/10.1016/S0140-6736(17)32366-8

Geoportal DANE. (2017). *Marco geoestadístico nacional: Santander*. https://geoportal.dane.gov.co/?descarga-marco-geoestadistico-nacional

Georgiev, G., Gueorguieva, N., Chiappa, M. y Krauza, A. (2016). Feature selection using Gustafson-Kessel fuzzy algorithm in high dimension data clustering [presentación de ponencia]. *2015 IEEE 14th International Conference on Machine Learning and Applications, ICMLA 2015*, 1-6. https://doi.org/10.1109/ICMLA.2015.57

Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S. y Brilliant, L. (2009). Detecting influenza epidemics using search engine query data. *Nature*, *457*(7232), 1012-1014. https://doi.org/10.1038/nature07634

Gobernación de Santander. (2014). *Lineamientos y directrices de ordenamiento territorial del Departamento de Santander*. Universidad Santo Tomás. http://historico.santander.gov.co/index.php/gobernacion/documentacion/finish/359-lineamientos-y-directrices-del-ordenamiento-territorial/5969-lineamientos-y-directrices-de-ordenamiento-territorial-del-departamento-de-santander

Gobernación de Santander. (2017). *Directorio municipios*. http://www.santander.gov.co/index.php/atencion-al-ciudadano/directorios/directorio-municipios

Grassly, N. C. y Fraser, C. (2008). Mathematical models of infectious disease transmission. *Nature Reviews Microbiology*, *6*, 477-487. https://doi.org/10.1038/nrmicro1845

Grekousis, G. y Thomas, H. (2012). Comparison of two fuzzy algorithms in geodemographic segmentation analysis: The fuzzy C-means and Gustafson-Kessel methods. *Applied Geography*, *34*, 125-136. https://doi.org/10.1016/j.apgeog.2011.11.004

Grossglauser, M. y Saner, H. (2014). Data-driven healthcare: From patterns to actions. *European Journal of Preventive Cardiology*, *21*(25), 14-17. https://doi.org/10.1177/2047487314552755

Gu, W., Unnasch, T. R., Katholi, C. R., Lampman, R. y Novak, R. J. (2008). Fundamental issues in mosquito surveillance for arboviral transmission. *Transactions*

of the Royal Society of Tropical Medicine and Hygiene, *102*(8), 817-822. https://doi.org/10.1016/j.trstmh.2008.03.019

Hall, H. I., Correa, A., Yoon, P. W. y Braden, C. R. (2012). Lexicon, definitions, and conceptual framework for public health surveillance. *Morbidiy and Mortality Weekly Report 61*(3), 10-14. https://www.cdc.gov/mmwr/preview/mmwrhtml/su6103a3.htm

Hamerly, G. y Drake, J. (2015). Accelerating Lloyd's algorithm for k-means clustering BT-Partitional clustering algorithms. En M. E. Celebi (Ed.), *Partitional clustering algorithms* (pp. 41-78). Springer. https://doi.org/10.1007/978-3-319-09259-1_2

Han, J., Kamber, M. y Pei, J. (2012). Cluster analysis: Basic concepts and methods. En *Data mining* (pp. 443-495). Morgan Kaufman. https://doi.org/10.1016/B978-0-12-381479-1.00010-1

Hartigan, J. A. y Wong, M. A. (1979). Algorithm AS 136: A *k*-means clustering algorithm. *Applied Statistics*, *28*(1), 100. https://doi.org/10.2307/2346830

Hay, S. I., George, D. B., Moyes, C. L. y Brownstein, J. S. (2013). Big data opportunities for global infectious disease surveillance. *PLoS Medicine*, *10*(4). https://doi.org/10.1371/journal.pmed.1001413

Hu, H., Wen, Y., Chua, T.-S. y Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE Access*, *2*, 652-687. https://doi.org/10.1109/ACCESS.2014.2332453

Instituto de Hidrología, Meteorología y Estudios Ambientales (Ideam). (2017). *Archivo de información hidrometeorológica*.

Instituto Nacional de Salud. (2017). Dengue Código: 210 – 220 – 580. En *Protocolo de Vigilancia en Salud Pública*. https://www.ins.gov.co/buscador-eventos/Lineamientos/Dengue PROTOCOLO.pdf

Instituto Nacional de Salud. (2019). *Enfermedades crónicas no transmisibles*. https://www.ins.gov.co/Direcciones/Vigilancia/Paginas/No-Transmisibles.aspx

Jones, K. E., Patel, N. G., Levy, M. A., Storeygard, A., Balk, D., Gittleman, J. L. y Daszak, P. (2008). Global trends in emerging infectious diseases. *Nature*, *451*(7181), 990-993. https://doi.org/10.1038/nature06536

Kapageridis, I. K. (2015). Variable lag variography using k-means clustering. *Computers and Geosciences*, *85*, 49-63. https://doi.org/10.1016/j.cageo.2015.04.004

Khoury, M. J. (2015). Planning for the future of epidemiology in the era of big data and precision medicine. *American Journal of Epidemiology*, *182*(12), 1-5. https://doi.org/10.1093/aje/kwv228

Khoury, M. J., Lam, T. K., Ioannidis, J. P. A., Hartge, P., Spitz, M. R., Buring, J. E., … Schully, S. D. (2013). Transforming epidemiology for 21st century medicine and public health. *Cancer Epidemiology Biomarkers and Prevention*, cebp-0146. https://doi.org/10.1158/1055-9965.EPI-13-0146

Lloyd, S. P. (1982). Least squares quantization in PCM. *IEEE Transactions on Information Theory*, *28*(2), 129-137. https://doi.org/10.1109/TIT.1982.1056489

López, D., Gunasekaran, M., Senthil Murugan, B., Kaur, H. y Abba. (2014). Spatial big data analytics of influenze epidemic in Vellore, India. En *IEEE International Conference on Big Data, 2014* (pp. 19-24). IEEE. https://ieeexplore.ieee.org/document/7004422

Macqueen, J. (1967). Some methods for classification and analysis of multivariate observations. En *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, *1*(233), 281-297.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. y Hung, A. (2011). *Big data: The next frontier for innovation, competition, and productivity*. McKinsey Global Institute. https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation#

Ministerio de Salud y Proteccion Social. (2014). *Guía conceptual y metodológica para la construcción del ASIS de las entidades territoriales. Colombia 2014*. https://www.minsalud.gov.co/salud/publica/epidemiologia/Paginas/analisis-de-situacion-de-salud-.aspx

Ministerio de Salud y Protección Social. (2017). *Bodega de datos de SISPRO: Sivigila*.

Morissette, L. y Chartier, S. (2013). The *k*-means clustering technique: General considerations and implementation in *Mathematica*. *The Quantitative Methods for Psychology*, *9*(1), 15-24. https://doi.org/10.20982/tqmp.09.1.p015

Mutheneni, S. R., Mopuri, R., Naish, S., Gunti, D. y Upadhyayula, S. M. (2018). Spatial distribution and cluster analysis of dengue using self organizing maps in Andhra Pradesh, India, 2011-2013. *Parasite Epidemiology and Control*, *3*(1), 52-61. https://doi.org/10.1016/j.parepi.2016.11.001

Nidheesh, N., Abdul Nazeer, K. A. y Ameer, P. M. (2017). An enhanced deterministic K-Means clustering algorithm for cancer subtype prediction from gene expression data. *Computers in Biology and Medicine*, *91*(June), 213-221. https://doi.org/10.1016/j.compbiomed.2017.10.014

O'Connell, E. K., Zhang, G., Leguen, F., Llau, A. y Rico, E. (2010). Innovative uses for syndromic surveillance. *Emerging Infectious Diseases*, *16*(4), 669-671. https://doi.org/10.3201/eid1604.090688

Ogino, S., Campbell, P. T., Nishihara, R., Phipps, A. I., Beck, A. H., Sherman, M. E., … Begg, C. B. (2015). Proceedings of the second international molecular pathological epidemiology (MPE) meeting. *Cancer Causes & Control*, *26*(7), 959-972. https://doi.org/10.1007/s10552-015-0596-2

Organización Mundial de la Salud (OMS). (2017). *Dengue*. https://www.who.int/topics/dengue/es/

Organización Mundial de la Salud (OMS). (2018). *Enfermedades no transmisibles*. https://www.who.int/topics/noncommunicable_diseases/es/

Organización Panamericana de la Salud (OPS). (2012). Determinantes e inequidades en salud. En *Salud en las Américas*. https://www.paho.org/salud-en-las-americas-2012/index.php?option=com_docman&view=download&category_slug=sa-2012-volumen-regional-18&alias=163-capitulo-2-determinantes-e-inequidades-salud-163&Itemid=231&lang=en

Organización Panamericana de la Salud. (2015). *Dengure: guías para la atención de enfermos en la región de las Américas* (vol. 2). https://iris.paho.org/handle/10665.2/28232

Ponce Cruz, P. (2010). *Inteligencia artificial con aplicaciones a la ingeniería*. Alfaomega.

Porta, M. (Ed.). (2008). *Dictionary of epidemiology*. Oxford University Press. https://www.oxfordreference.com/view/10.1093/acref/9780195314496.001.0001/acref-9780195314496

Pyne, S., Vullikanti, A. K. S. y Marathe, M. V. (2015). Big data applications in health sciences and epidemiology. En *Handbook of statistics: Big data analytics* (vol. 33, pp. 171-202). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-444-63492-4.00008-3

Quintero-Herrera, L. L., Ramírez-Jaramillo, V., Bernal-Gutiérrez, S., Cárdenas-Giraldo, E. V., Guerrero-Matituy, E. A., Molina-Delgado, A. H., … Rodríguez-Morales, A. J. (2015). Potential impact of climatic variability on the epidemiology of dengue in Risaralda, Colombia, 2010-2011. *Journal of Infection and Public Health*, *8*(3), 291-297. https://doi.org/10.1016/j.jiph.2014.11.005

Raghupathi, W. y Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems*, *2*(3), 1-10. https://doi.org/10.1186/2047-2501-2-3

Ram, S., Zhang, W., Williams, M. y Pengetnze, Y. (2015). Predicting asthma-related emergency department visits using big data. *IEEE Journal of Biomedical and Health Informatics, 19*(4), 1216-1223. https://doi.org/10.1109/JBHI.2015.2404829

Robertson, C. (2017). Towards a geocomputational landscape epidemiology: Surveillance, modelling, and interventions. *GeoJournal*, *82*(2), 397-414. https://doi.org/10.1007/s10708-015-9688-5

Rodríguez-Morales, A. J., Ruiz, P., Tabares, J., Ossa, C. A., Yepes-Echeverry, M. C., Ramírez-Jaramillo, V., … Grobusch, M. P. (2017). Mapping the ecoepidemiology of Zika virus infection in urban and rural areas of Pereira, Risaralda, Colombia, 2015-2016: Implications for public health and travel medicine. *Travel Medicine and Infectious Disease*, *18*, 57-66. https://doi.org/10.1016/j.tmaid.2017.05.004

Sinka, M. E., Bangs, M. J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M., … Hay, S. I. (2012). A global map of dominant malaria vectors. *Parasites & Vectors*, *5*(1), 69. https://doi.org/10.1186/1756-3305-5-69

Tsui, K. L., Wong, S. Y., Jiang, W. y Lin, C. J. (2011). Recent research and developments in temporal and spatiotemporal surveillance for public health. *IEEE Transactions on Reliability*, *60*(1), 49-58. https://doi.org/10.1109/TR.2010.2104192

World Health Organization. (2017). *The WHO STEPwise approach to noncommunicable disease risk factor surveillance*. http://www.who.int/chp/steps/STEPS_Manual.pdf?ua=1

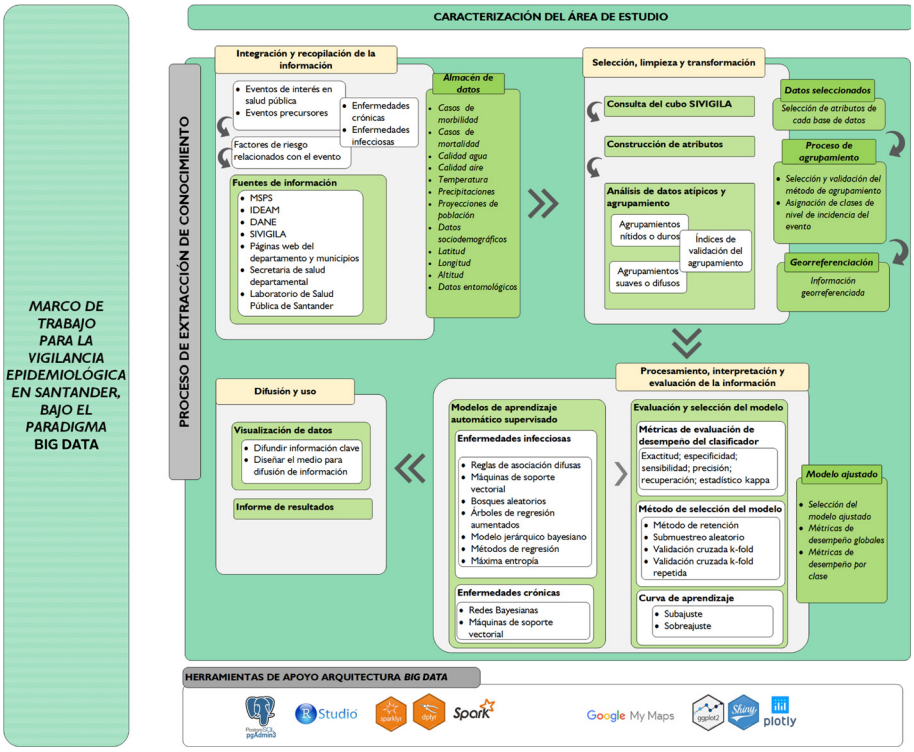
Wu, Y., Lee, G., Fu, X. J. y Hung, T. (2008). Detect climatic factors contributing to dengue outbreak based on wavelet, support vector machines and genetic algorithm. En *World Congress on Engineering 2008 Vols IIi* (vol. 1, pp. 303-307).

Xie, X. L. y Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 13*(8), 841-847. https://doi.org/10.1109/34.85677

Yang, S., Santillana, M. y Kou, S. C. (2015). Accurate estimation of influenza epidemics using Google search data via ARGO. *Proceedings of the National Academy of Sciences*, *112*(47), 14473-14478. https://doi.org/10.1073/pnas.1515373112

Anexos

Anexo 1. Marco de trabajo



Fuente: elaboración propia.

Anexo 2. Método de análisis y autores asociados a eventos

Autor	Evento	Método
Hay et al. (2013)	Dengue	Árbol de regresión aumentado (*Boosted Regresion Tree*)
Wu et al. (2008)	Dengue	Análisis de ondículas Máquinas de vectores de soporte Algoritmo genético Regresión lineal SVR
Rodríguez-Morales et al. (2017)	Zika	Regresión de Poisson múltiple y simple

Autor	Evento	Método
Ginsberg et al. (2009)	Influenza	Modelo lineal utilizando *log-odds*
Sinka et al. (2012)	Malaria	Árbol de regresión aumentado (*Boosted Regresion Tree*)
Bhatt et al. (2013)	Dengue	Árbol de regresión aumentado (*Boosted Regresion Tree*) Modelo jerárquico bayesiano
Jones et al. (2008)	Enfermedades infecciosas emergentes	Regresión logística multivariada
López et al. (2014)	Influenza	Regresión ponderada geográficamente (GWR, por sus siglas en inglés)
Buczak et al. (2012)	Dengue	Reglas de asociación difusas Regresión logística
Quintero-Herrera et al. (2015)	Dengue	Regresión de Poisson no lineal

Fuente: elaboración propia.

Anexo 3. Resultados de los mejores algoritmos de agrupamiento

Año	Método	Grupos	Índice_ PC	Índice_ PE	Índice_ XB	Índice_S
2007	K-medias (Hartigan, Forgy, Lloyd y MacQueen)	2	1,0000	0	0,6679	0,6143
	K-medias difuso	*2*	*0,7879*	*0,3337*	*0,3145*	*0,8103*
	Gustafson-Kessel	2	0,6036	0,5766	Inf	−0,0135
2008	K-medias (Hartigan, Forgy, Lloyd y MacQueen)	2	1,0000	0	0,6079	0,5819
	K-medias difuso	*2*	*0,8192*	*0,2923*	*0,1938*	*0,8329*
	Gustafson-Kessel	2	0,6149	0,5624	Inf	0,0384
2009	K-medias (Forgy, Lloyd y MacQueen)	5	1,0000	0	0,2760	0,6558
	K-medias difuso	*2*	*0,8518*	*0,2464*	*0,1353*	*0,8748*
	Gustafson-Kessel	2	0,6149	0,5624	Inf	−0,0013

Año	Método	Grupos	Índice_ PC	Índice_ PE	Índice_ XB	Índice_S
2010	K-medias (Hartigan, Forgy, Lloyd y MacQueen)	2	1,0000	0	0,4718	1,1693
	K-medias difuso	*3*	*0,6343*	*0,6431*	*0,7560*	*0,7983*
	Gustafson-Kessel	2	0,6036	0,5766	Inf	0,0003
2011	K-medias (Hartigan)	10	1,0000	0	0,3680	1,3884
	K-medias difuso	*2*	*0,6927*	*0,4596*	*0,8042*	*0,7115*
	Gustafson-Kessel	2	0,6149	0,5624	Inf	−0,0073
2012	K-medias (Forgy, Lloyd y MacQueen)	5	1,0000	0	0,4187	1,9887
	K-medias difuso	*2*	*0,7510*	*0,3871*	*0,3527*	*0,7128*
	Gustafson-Kessel	2	0,6149	0,5624	Inf	0,0037
2013	K-medias (Hartigan, Forgy, Lloyd y MacQueen)	2	1,0000	0	1,2724	0,5113
	K-medias difuso	*4*	*0,5696*	*0,8093*	*1,4414*	*0,7835*
	Gustafson-Kessel	2	0,6149	0,5624	Inf	−0,0028
2014	K-medias (Forgy, Lloyd y MacQueen)	3	1,0000	0	0,8598	0,4576
	K-medias difuso	*2*	*0,7501*	*0,3914*	*0,3519*	*0,7353*
	Gustafson-Kessel	2	0,6149	0,5624	Inf	−0,0355
2015	K-medias (Forgy, Lloyd y MacQueen)	2	1,0000	0	0,6754	0,5970
	K-medias difuso	*2*	*0,7442*	*0,3951*	*0,4363*	*0,7770*
	Gustafson-Kessel	2	0,6149	0,5624	Inf	0,0044
2016	K-medias (Hartigan-Wong)	2	1,0000	0	1,4725	0,4828
	K-medias difuso	*2*	*0,7683*	*0,3669*	*0,3228*	*0,7910*
	Gustafson-Kessel	2	0,6149	0,5624	Inf	0,0166

Fuente: elaboración propia.

Anexo 4. Perfil de los niveles de incidencia del dengue en Santander, periodo 2007-2016

Nivel de incidencia	Descripción
Muy baja	Nivel de incidencia promedio entre 0 y 54,304 casos por 100 000 habitantes. Presencia de casos de dengue en personas que residen en el área cabecera. La distribución de la población en estos municipios es mayor en el área rural y centro poblado que en el área cabecera. Mayor incidencia en personas del sexo masculino, en personas afiliadas al régimen de salud subsidiado, contributivo, especial y las no afiliadas. Las entidades prestadoras del servicio son en su mayoría públicas. Personas en el ciclo vital de la juventud. En el rango de edad de 5 a 29 años y de 35 a 39 años se presenta la mayor proporción de casos. Agrupa municipios con una altitud promedio de 1655 m s. n. m.
Baja	El nivel de incidencia promedio de dengue se encuentra entre 54,304 y 124,146 casos por 100 000 habitantes. Las personas afectadas son residentes del área cabecera, aunque también se destaca la presencia de casos en centros poblados. Hay mayor incidencia en los rangos de edad de 0 a 4 años y de 10 a 29 años, así como en el sexo masculino. Personas en las etapas del ciclo vital juventud o adultez. Se destaca la presencia de dengue en mujeres gestantes. Las personas están afiliadas al régimen subsidiado, contributivo o excepción. Las entidades prestadoras del servicio son de carácter público. Algunos casos son confirmados por laboratorio. Se agrupan municipios que se caracterizan por una altitud promedio de 1522,36 m s. n. m.
Media	Nivel de incidencia promedio entre 124,146 y 400,23 casos por 100 000 habitantes. Personas son residentes del área cabecera. Mayor presencia en el sexo masculino. La distribución de la población es mayor en el área cabecera que en las otras áreas. Se resalta la incidencia de dengue en madres comunitarias. El rango de edad de 5 a 34 años hay mayor presencia de casos. Personas en las etapas del ciclo vital juventud o adultez. Los regímenes de afiliación a salud predominantes son el subsidiado, contributivo y no afiliado. Se destacan las entidades públicas prestadoras del servicio. Se agrupan municipios con una altitud promedio de 1178,54 m s. n. m.
Alta	Nivel de incidencia promedio de dengue se encuentra entre 400,23 y 661,16 casos por 100 000 habitantes. Se destaca la presencia de casos en el área cabecera, en personas del sexo masculino y femenino, y en mujeres gestantes. El régimen de afiliación de las personas es subsidiado o contributivo, y las entidades prestadoras del servicio son públicas. Personas en las etapas del ciclo vital juventud o adultez. Mayor presencia de casos en los rangos de edad de 5 a 24 años y de 30 a 34 años. En cuanto a la distribución de la población, la mayoría se encuentra en el área cabecera, y los municipios que se agrupan se caracterizan con un promedio de altitud de 1276,9 m s. n. m.

Nivel de incidencia	Descripción
Muy alta	Nivel de incidencia promedio de dengue entre 661,16 y 1391,89 casos por 100 000 habitantes. Presencia de casos en personas que residen en el área cabecera, personas del sexo masculino y femenino, y en mujeres gestantes. En el rango de edad de 0 a 24 años se da la mayoría de los casos. Personas en las etapas del ciclo vital juventud o adultez. Los regímenes se afiliación a salud son el subsidiado o contributivo. Las entidades prestadoras del servicio se caracterizan por ser públicas, privadas o mixtas. Se resalta la confirmación de casos por nexo epidemiológico. En cuanto a la distribución de la población es mayor en el área cabecera que en las otras áreas. La altitud de los municipios se caracteriza por el promedio de 1177,70 m s. n. m.

Fuente: elaboración propia.

Capítulo 8. Metodología para la identificación y seguimiento del riesgo en salud de una cohorte poblacional atendida bajo el modelo Sermás

Edwin Alonso Álvarez Tobón,[1]
Paula Andrea Velásquez Restrepo,[2]
Marta Cecilia Ramírez Orrego[3] *y*
Alma Karina Rodríguez Quintero[4]

[1] Desarrollo organizacional
Institución prestadora de servicios
de salud, IPS Universitaria
Medellín, Colombia
edwin.alvarezt@ipsuniversitaria.com.co

[2] Desarrollo organizacional
Institución prestadora de servicios
de salud, IPS Universitaria
Medellín, Colombia
paula.velasquezr@ipsuniversitaria.com.co

[3] Dirección general
Institución prestadora de servicios
de salud, IPS Universitaria
Medellín, Colombia
marta.ramirezo@ipsuniversitaria.com.co

[4] Desarrollo organizacional
Institución prestadora de servicios
de salud, IPS Universitaria
Medellín, Colombia
alma.rodriguezq@ipsuniversitaria.com.co

Resumen

Los sistemas de salud son parte de la respuesta de la sociedad a las necesidades de un mayor bienestar de una población. Por consiguiente, deben coincidir la situación de salud de la población y la forma en que se estructura estos sistemas para responder socialmente a esa situación particular. Debido a esto, existe un creciente interés de las instituciones de salud en torno a la reformulación de la prestación de dichos servicios, con un enfoque poblacional e individualizado. En este capítulo se presenta una metodología adoptada por la IPS Universitaria en su

modelo de atención Sermás, que permite identificar, clasificar, definir grupos y estratificar, según el riesgo, cualquier cohorte poblacional asignada para su atención. Este enfoque metodológico tomó como base una revisión bibliográfica de los modelos descriptivo y predictivo que apoyan la gestión de riesgo en salud. Después, la metodología propuesta se aplica a una cohorte de pacientes, buscando compartir los resultados obtenidos.
Palabras clave: riesgo en salud, caracterización, segmentación, estratificación del riesgo, cogestión del riesgo

Introducción

Como consecuencia del descenso de la tasa de fecundidad y de la reducción de las enfermedades mortales de la niñez, la mortalidad materna y de personas mayores, ha habido un aumento significativo de la esperanza de vida en muchos países, lo que también causa un incremento de la población madura y longeva (Organización Mundial de la Salud, 2017). Entre el año 2000 y el 2050, se estima que la proporción de la población mundial adulta mayor (60 y más años de edad) se duplicará y pasará del 11 % a un 22 %. Los cambios demográficos registrados en Colombia son similares a los observados en el resto del mundo. En el 2012, la población adulta mayor colombiana ascendió al 10,8 % de la población total y para el 2050 se proyecta que será el 23 %. Así se convertirá en uno de los países que mayor tasa de envejecimiento tendrá en América Latina (Fedesarrollo y Fundación Saldarriaga Concha, 2015; Ministerio de Salud y Protección Social [Minsalud], 2013a).

El proceso de envejecimiento de la población señala que habrá un aumento relativo de las enfermedades crónicas no trasmisibles (ECNT), porque afectan más a los segmentos poblacionales de mayor edad. En las Américas, las ECNT son responsables de siete de cada diez muertes entre las personas de 70 años y más. Adicionalmente, la prevalencia de dos o más enfermedades crónicas concurrentes se incrementa con la edad (Organización Panamericana de la Salud, 2012).

El aumento de la demanda de servicios en salud de las personas con ECNT constituye un desafío para los sistemas de salud de muchos países, y Colombia no es la excepción. En el país, los altos costos que representan las enfermedades crónicas están reflejados por la complejidad de su diagnóstico y de su tratamiento, ya que se caracterizan por ser prolongados y complejos, lo que afecta a la población económicamente activa. También representan una severa carga financiera para los sistemas de salud, dado que su diagnóstico e intervención es tardío, por lo cual no se obtiene un beneficio significativo para la salud del paciente y de la familia (Gallardo Solarte et al., 2016). En este contexto, la estrategia de Atención Primaria en Salud (APS) debe constituirse en una solución factible para mejorar el abordaje de las enfermedades crónicas (Londoño Agudelo, 2017).

El Sistema General de Seguridad Social en Salud (SGSSS) de Colombia, definido por la Ley 100 de 1993, ha permitido el aumento en la cobertura y acceso a los servicios de salud a toda la población. Durante su vigencia el sistema ha realizado diversas reformas, entre las cuales se destacan la Ley 1122 de 2007, la Ley 1438 de 2015 y la Ley 1751 de 2015, conocida como Ley Estatutaria, que regula el derecho fundamental a la salud y, en consecuencia, le da origen a la Resolución 429 de 2016, donde se adopta la Política de Atención Integral en Salud (PAIS), cuyo objetivo es orientar el sistema de salud hacia la generación de nuevas y mejores prácticas para atender las necesidades de la población. Dicha política tiene un marco operacional, el Modelo Integral de Atención en Salud (MIAS), definido en la misma Resolución 429, modificado por el Modelo de Acción Integral Territorial (Maite), mediante la Resolución 2626 del 2019.

La Maite conserva el marco estratégico de la PAIS, fundamentado en la APS, con enfoque de salud familiar y comunitario, la gestión integral del riesgo y el enfoque diferencial para los distintos territorios y poblaciones. Esta define ocho líneas de acción, siendo la primera el aseguramiento con acciones específicas a cargo de las entidades administradoras de planes de beneficio en salud (EAPB) para, entre otras, gestionar el riesgo financiero, el riesgo en salud (GIRS) y la interacción con los demás agentes del sistema. También determina herramientas para la identificación de los grupos de riesgo y sus características, la implementación y adaptación de rutas integrales de atención en salud (RIAS), el seguimiento de cohortes, la aplicación de modelos predictivos de riesgo y la efectividad de los procesos de servicio (Resolución 2626 de 2019).

La puesta en práctica de la PAIS exige grandes retos y profundos cambios en los actores del sistema de salud. En el departamento de Antioquia, y más específicamente en la ciudad de Medellín, la Institución Prestadora de Servicios de Salud de Alta Complejidad (IPS Universitaria) definió dentro del marco de esta política su nuevo modelo de atención llamado Sermás, consecuente con la situación de salud, la dinámica del sistema, las políticas sectoriales y su direccionamiento estratégico. El modelo Sermás está centrado en el ser humano, personalizado y con enfoque diferencial; considera la atención integral, integrada y continua de la población usuaria de la IPS Universitaria, mediante la gestión del riesgo en salud y a través de los ámbitos ambulatorios, hospitalarios y domiciliarios. De igual manera, propone una sinergia entre el prestador primario y el complementario que conforma la red integrada de servicios de salud, que permite fortalecer la pertinencia y oportunidad en las acciones de cuidado y mantenimiento de la salud, detección de riesgos y diagnóstico oportuno, curación, rehabilitación y paliación (IPS Universitaria, 2017).

En el SGSSS, la prestación del servicio de salud usualmente se da mediante un acuerdo contractual entre las aseguradoras, como son las EAPB o las direcciones territoriales, y por quienes prestan el servicio, como son las instituciones

prestadoras de servicios de salud (IPS). Los mecanismos de pago de la ejecución del contrato puede darse de dos formas: retrospectivo, cuando es después de la prestación del servicio, y está relacionado con el pago de la atención de salud individual (actividad), por cada persona atendida, liquidada por el valor establecido contractualmente, y prospectivo, cuando se realiza antes de la prestación del servicio, y está relacionado con el pago de atenciones de salud, por una cohorte poblacional determinada, liquidada por el producto del valor establecido por la frecuencia de las atenciones o la cantidad de población (González Fuentes, 2017).

El modelo Sermás se acompaña de un mecanismo de pago prospectivo que puede ser pago global prospectivo, pago global por resultados, pago por grupo relacionado de diagnóstico, entre otros. En esta modalidad de contratación, las EAPB y los prestadores comparten el riego primario y el riesgo técnico en la atención de una población, lo cual ha sido denominado *cogestión del riesgo en salud* (Minsalud, 2018). El riesgo primario está determinado por la población afiliada por la EAPB y corresponde a la variación en el número de condiciones médicas que ocurren en un periodo y en algunos casos en su severidad (Castaño Yepes, 2014), y el riesgo técnico está determinado por el prestador y corresponde a la variación en el consumo de recursos en el proceso de atención, por factores relacionados mayormente con decisiones y conductas asumidas por el prestador en el tratamiento de una condición médica, no sustentada por la evidencia científica (Castaño Yepes, 2014).

Así que, para reorientar la atención basada en la gestión del riesgo, según lo propuesto por la PAIS y consecuentemente por el modelo Sermás, es necesario identificar y clasificar en grupos de riesgo a la población asignada y, luego, estratificarla, de tal modo que apoye la definición de las estrategias de atención. Por esto, se planteó el objetivo de establecer una metodología de identificación, clasificación en grupos de riesgo y estratificación de los grupos según el nivel de riesgo para los usuarios del modelo Sermás, de tal forma que apoye las actividades de planeación (brindando elementos para el direccionamiento estratégico-táctico), una adecuada gestión clínica, la consecución de mejores resultados en salud y el uso eficiente de los recursos. Ello es relevante para los usuarios y el sistema de salud, pues es una propuesta innovadora, con la posibilidad de que los resultados brinden a otros actores del sistema de salud un marco metodológico alternativo para mejorar su planeación y la eficiencia de sus actividades.

Se llevó a cabo una revisión de la literatura que permitiera seleccionar métodos para construir una metodología que diera una ruta de inicio a los actores del sistema sobre cómo empezar la implementación de la gestión del riesgo en salud. A partir de este trabajo, se seleccionaron algunos métodos que, con las limitaciones presentes, se pudieran adoptar y se identificaron retos y trabajos futuros que se pueden desarrollar.

Las fuentes de información primarias y secundarias. Para abordar las fuentes de información del sector de la salud en Colombia que permitan afrontar los retos metodológicos que implica la gestión del riesgo, se partió del trabajo de Bernal-Acevedo y Forero-Camacho (2011), el cual caracterizó y evaluó el sistema de información del sector, y de algunos documentos de Minsalud (2016). De acuerdo con esto, las fuentes de información se pueden clasificar en dos grupos, según la población objetivo: a) información basada en la población (censos, registro de hechos vitales y las encuestas poblacionales) y b) información basada en los servicios de salud (sistemas clásicos de vigilancia de las enfermedades, registro de las actividades de prestación de servicios de salud, etc.). La guía metodológica del Minsalud también hace referencia a plantillas e instructivos e, igualmente, referencia la disponibilidad de indicadores en el sitio web del Sistema Integrado de Información de la Protección Social (Sispro).

La identificación y clasificación de los usuarios según el riesgo. Para esto, Minsalud sugiere la metodología de caracterización poblacional, comprendida como la identificación y reconocimiento de la población con unas características en común en un momento determinado (Resolución 1536 de 2015; Minsalud, 2018).

En términos de caracterización de la población, Colombia tiene unos lineamientos metodológicos definidos técnicamente por la Dirección de Epidemiología y Demografía (2017) del Minsalud de Colombia, establecidos en la *Guía conceptual y metodológica para la caracterización de la población afiliada a las entidades administradoras de planes de beneficio en salud* (*EAPB*). Según la guía, una caracterización de una población usuaria debe estar compuesta por cinco componentes: territorial y demográfico, mortalidad, morbilidad, determinantes sociales en salud y priorización.

Para caracterizar la población se emplean diversos métodos de análisis comúnmente apoyados por medidas estadísticas, como las frecuencias absolutas e indicadores cuantitativos usuales en el análisis epidemiológico (Fajardo Gutiérrez, 2017).

Otro tipo de metodologías encontradas para identificar riesgos en salud en una población se apoya en la modelación predictiva, basada en modelos que buscan establecer relaciones entre conjuntos de variables para predecir resultados futuros, utilizando fórmulas y métodos estadísticos. La mayoría de las metodologías predictivas utilizan modelos de regresión, aunque cada vez más se estudian métodos basados en inteligencia artificial (Vuik et al., 2016).

La definición de los grupos de riesgo. Para esto existen diferentes enfoques metodológicos entre los que encontramos la segmentación.

La segmentación, definida como el proceso por medio del cual se lleva a cabo la separación de elementos en grupos homogéneos en el interior de ellos y heterogéneos entre ellos, se fundamenta en el reconocimiento de diferencias

significativas en sus características (Superintendencia Nacional de Salud, 2016). Aplicando este concepto a la gestión de riesgo en salud, la segmentación permite identificar colectivos homogéneos de pacientes y los patrones característicos de necesidad y comportamiento de cada grupo (Gobierno de Navarra-Departamento de Salud, 2013). El análisis de segmentación de la población basado en datos se ha utilizado ampliamente en diferentes contextos clínicos y tiene un gran potencial en la gestión de la salud de la población (Yan et al., 2018).

Existen distintas metodologías para segmentar o agrupar pacientes. Se encuentran modelos descriptivos que contemplan *modelos basados en criterio clínico*, apoyados en la formación, el conocimiento y la experiencia del personal de la salud para identificar pacientes de riesgo que pueden beneficiarse de una intervención temprana, y *modelos basados en reglas*, umbrales de determinados parámetros o en criterios prestablecidos de decisión, que describen a un paciente de alto riesgo y no modelos estadísticos (Gobierno de Navarra-Departamento de Salud, 2013; Osakidetza et al., 2015).

Para la realización de la segmentación usualmente se emplean técnicas estadísticas como tabulaciones cruzadas, análisis factorial, análisis de conglomerados (clúster), redes neuronales, análisis *log-lineal*, entre otros (López, 2013).

La estratificación del riesgo en salud. El Minsalud concibe el riesgo en salud como la "probabilidad de ocurrencia de un evento no deseado, evitable y negativo para la salud del individuo, que puede ser también el empeoramiento de una condición previa o la necesidad de requerir más consumo de bienes y servicios que hubiera podido evitarse" (Resolución 1841 de 2013).

Adicionalmente, la estratificación se define como la identificación o la agrupación de pacientes según la clasificación del riesgo o la severidad. Sirve así para definir de manera anticipada intervenciones que se adapten a sus necesidades de cuidados sanitarios futuros (Gobierno de Navarra-Departamento de Salud, 2013).

Para la estratificación existen sistemas de ajuste de riesgos o clasificación de pacientes, entre los que se encuentran: los *grupos clínicos ajustados* (*adjusted clinical groups*), que asignan a cada persona en una categoría exclusiva, basándose en criterios clínicos con el objetivo de predecir el costo en salud, los gastos farmacéuticos y los riesgos de hospitalización (Orueta Mendia et al., 2014); los *grupos de costes basados en diagnósticos* (*diagnostic cost groups*), que utilizan información de todos los diagnósticos y prescripciones y las características demográficas, para predecir los costos de salud en un año posterior al año de predicción (Ellis et al., 1996); el *grupo de riesgo clínico* (*clinical risk groups*), que clasifica a los individuos según los niveles de gravedad de su estado de salud y en función de su estado de cronicidad (Sicras-Mainar et al., 2006), y los *grupos relacionados por el diagnóstico* (*GRD*), que permiten relacionar los distintos tipos de pacientes tratados en un hospital, con el coste que representa su asistencia (Cortés et al.,

2016). Con relación a los GRD, su uso está recomendado por la Organización Mundial de la Salud y varios gobiernos latinoamericanos están explorando su implementación (Cortés et al., 2016). Sin embargo, para esto es fundamental contar con sistemas de información que tengan una alta calidad del dato, a fin de garantizar la precisión de las clasificaciones y, por ende, las decisiones que se tomen. En el caso colombiano se ha evidenciado una baja calidad en los registros individuales de prestación de servicios en salud (RIPS) para el sistema GRD y la ausencia de políticas que la fomenten y propicien esquemas comparables de gestión del riesgo de salud (Cortés et al., 2016).

Estos modelos consiguen explicar una parte importante de la variabilidad en la utilización de servicios de salud que realizará una población y ofrecen, para cada individuo, una estimación prospectiva del volumen de recursos de salud que requerirá el año siguiente.

De igual manera, se usan modelos descriptivos, basados en umbrales, que implican definir criterios para clasificar una cohorte, por ejemplo, en términos de uso de recursos, morbilidad o la edad. Esto se puede identificar en la estrategia del Departamento de Salud del Gobierno de Navarra (2013) que, mediante criterios de ponderación y umbrales de clasificación, calcula variables o indicadores de estratificación del riesgo: descontrol, uso de recursos y severidad para los segmentos de la población crónica obtenida según el tipo y cantidad de enfermedades.

El descontrol se refiere a la descompensación de una enfermedad crónica, que demanda atención de eventos inesperados, teniendo en cuenta la frecuencia de consultas de emergencia y hospitalizaciones no programadas. El uso de recursos permite medir el nivel de gravedad del estado del paciente por la utilización de servicios de salud, considerando hospitalizaciones, servicios de cirugía y consultas externas. La severidad es el resultado de la combinación del descontrol con el uso de recursos (Gobierno de Navarra-Departamento de Salud, 2013).

Priorización, reenfoque y seguimiento a los resultados dentro de un modelo de atención. Se debe resaltar que los resultados de la estratificación de los grupos de riesgo priorizados son indispensables para la definición, la organización, la implementación y el monitoreo de las estrategias de atención del modelo (Minsalud, 2018).

Consecuente con lo anterior, el documento de trabajo de la gestión integral del riesgo en salud de Minsalud (2018), en el caso de los grupos de riesgo crónicos, resalta la necesidad de la estratificación del nivel de riesgo de los usuarios, la cual permite identificar las necesidades de atención de la población basándose en su perfil de comorbilidad y priorizarlos de acuerdo con su severidad, a fin de definir estrategias de atención y focalizar recursos al nivel de riesgo identificado.

Un enfoque de esto es el modelo de Kaiser Permanente (Minsalud, 2011; Osakidetza, 2015), que por cada nivel de riesgo define estrategias de atención como la gestión de casos (Morales-Asencio, 2014; Pérez Giménez, 2017), la ges-

tión de la enfermedad por cohortes (Bunch-Barrera et al., 2016), la educación para la autogestión de la salud-factores protectores (Bonal Ruiz et al., 2017) y la gestión de demanda inducida (Resolución 412 de 2000).

Cuando nos referimos al monitoreo de los riesgos, hay que retomar la definición de riesgo en salud del Minsalud (2018) ya mencionada. Para esta, el evento no deseado es la ocurrencia de la enfermedad o su evolución desfavorable y sus causas son los diferentes factores asociados que implica el conocimiento de sus principales determinantes y los mecanismos a través de los cuales se genera. Para lo cual la ingeniería ha avanzado en diferentes modelos para predecir la aparición de una enfermedad, a través de la búsqueda de biomarcadores y el desarrollo de herramientas de soporte al diagnóstico médico basadas en el *machine learning* (Ahmad et al., 2015). Estas aplicaciones tienen como común denominador la necesidad de contar con herramientas capaces de analizar información proveniente de múltiples fuentes, tales como historias clínicas electrónicas, exámenes de laboratorio, imágenes médicas y otras ayudas diagnósticas, medicación prescrita, información de salud pública e, incluso, análisis genéticos y metabólicos.

Adicionalmente, se puede realizar un seguimiento y monitoreo de la gestión del riesgo en salud mediante un análisis del comportamiento en salud de la población, es decir, se puede evaluar el empeoramiento de una condición previa o la necesidad de requerir más consumo de bienes y servicios que hubieran podido evitarse, como lo estipula también la definición de riesgo en salud. Si retomamos el modelo del Gobierno de Navarra (2013), se puede llevar a cabo mediante el seguimiento a los indicadores de descontrol, uso de recursos y severidad.

Tras la revisión bibliográfica, se obtuvo un panorama básico de métodos seleccionables para la construcción de un enfoque metodológico que diera una ruta de inicio en la identificación del riesgo en salud, clasificación de grupos de riesgo y estratificación del nivel de riesgo, como elemento fundamental en la GIRS, posible para la IPS Universitaria y los actores del sistema. Dicho enfoque metodológico se desarrolló teniendo en cuenta las limitaciones presentes, pero al mismo tiempo sentará las bases para vislumbrar nuevos retos y trabajos futuros para la academia y los actores del SGSSS colombiano. Así que, en este capítulo, a partir de la experiencia —un caso aplicado a población adulta mayor y con patologías crónicas—, se busca compartir el enfoque metodológico desarrollado y los resultados obtenidos.

Métodos

Este apartado detalla el enfoque metodológico adoptado por la IPS Universitaria para la identificación de los usuarios según el riesgo, la clasificación en grupos de riesgo, la estratificación según el nivel riesgo de una población asignada y la priorización, reenfoque y seguimiento a resultados dentro de un modelo de atención,

como herramienta y fuente de información de apoyo al proceso de planeación para la GIRS de la población, de acuerdo con las necesidades en salud y el nivel de riesgo identificado. Adicionalmente, con base en la revisión bibliográfica realizada, se presentan como propuesta algunos indicadores para hacer seguimiento a la gestión, en periodos de cuatro meses, que permitan a los administradores del sistema hacer control y retroalimentación a las acciones planteadas.

Fuentes de información

Teniendo en cuenta que dentro de un modelo de cogestión del riesgo la gestión del riesgo en salud es una responsabilidad y un trabajo conjunto entre EAPB e IPS (clínicas y hospitales), la primera etapa que se desarrolla dentro de este enfoque metodológico es solicitar a la EAPB bases de datos de, por lo menos, el último año de las siguientes fuentes de información: a) base de datos única de afiliados; b) reportes obligatorios: RIPS, eventos de notificación obligatoria al Sistema de Vigilancia en Salud Pública, Cuenta de Alto Costo y actividades de protección específica, detección temprana y aplicación de las guías de atención integral para las enfermedades de interés en salud pública de obligatorio cumplimiento; c) registros institucionales en *software* propios, que dan cuenta de la demanda de tipo de servicio, especialidad, diagnóstico y el uso de recursos, y d) estudios o encuestas de la población usuaria o su grupo familiar.

Así mismo, se usan las bases de datos institucionales de la IPS, correspondientes a la población de interés, que permitan identificar información sociodemográfica, diagnósticos de morbilidad y características de los servicios en las atenciones realizadas; esto en los ámbitos ambulatorio, hospitalario y domiciliario.

Acorde con los atributos de calidad internacionales, se realizaron actividades que propicien niveles óptimos en la calidad de la información, como son a) *pertinencia y relevancia*: se extrae por usuario, de las bases de datos disponibles, solo las variables mínimamente necesarias, como las *sociodemográficas* (nombre, tipo de identificación, número de identificación, fecha de nacimiento y género), *diagnósticos de morbilidad* (todos los diagnósticos en códigos CIE-10, principales y relacionados, registrados en los diferentes servicios durante el periodo de análisis) y *servicios y frecuencia de atención en el periodo de análisis* (frecuencia de consultas por cada servicio, total de días de estancia por tipo de cama, frecuencia de ayudas diagnósticas de imagenología y laboratorio clínico); b) *precisión o exactitud*: se incluyen datos de por lo menos un año; c) *oportunidad y puntualidad*: uso de los datos anuales más recientes; d) *accesibilidad*: recepción de los datos para custodia y manipulación, mediante acuerdos de confidencialidad; e) *interoperabilidad*: se solicita en estructura de archivos de texto o de base de datos en formato Microsoft Excel, compatible con las herramientas informáticas institucionales empleadas para la manipulación de bases de datos, y f) *coheren-*

cia: previo al análisis se identifica la correspondencia del dato con la variable en términos del tipo, nivel de medición, rango de valores posibles, o la ausencia del dato; y tras el hallazgo de no coherencia, según sea el caso, se realiza imputación del dato o exclusión de variables.

De las fuentes de información se extraen las variables necesarias para obtener una base de datos única de usuarios asignados; con datos sociodemográficos; de demanda periódica de los servicios; la morbilidad en código CIE-10, según la lista de carga global de enfermedad modificada por el Minsalud (2013b); las patologías crónicas de riesgo cardiovascular, priorizadas por Sermás, y una serie de variables calculadas, a partir de la demanda de servicios y de los diagnósticos, útiles para la definición de los niveles de riesgo en salud.

Identificación y clasificación de los usuarios según el riesgo

Se consideró pertinente identificar el riesgo de los usuarios pertenecientes a la cohorte poblacional asignada al modelo Sermás, usando la *Guía conceptual y metodológica para la caracterización poblacional de las entidades administradoras de planes de beneficios de salud (EAPB)* (Dirección de Epidemiología y Demografía, 2017), pues es un lineamiento sugerido para la GIRS, estandarizado para las EAPB y avalado por los entes rectores de salud nacionales.

Es importante resaltar que la viabilidad de las fuentes de información determinará el alcance y la profundidad de la identificación del riesgo, permitiendo o restringiendo el desarrollo de los componentes de la metodología de caracterización. Sin embargo, dentro del enfoque metodológico propuesto por la IPS Universitaria, se adaptan dos de los cinco componentes de la guía conceptual propuesta por Minsalud, los cuales se describen a continuación:

Caracterización demográfica. Para el modelo Sermás se determinó describir la estructura poblacional, empleando la variable género y la distribución de edad por quinquenios, con inicio en la edad cero y cuyo último grupo es el de 80 y más años, siendo este último el único de rango diferente. Esto permite construir la pirámide poblacional. Con respecto a los indicadores demográficos, se consideró pertinente calcular la edad promedio y la razón hombre-mujer. En el enfoque diferencial se determinó identificar, para el análisis de morbilidad, a los usuarios por ciclos vitales: 0-5 años (primera infancia), 6-11 años (infancia), 12-18 años (adolescencia), 19-26 años (juventud), 27-59 años (adultez), 60 y más años (persona mayor) (Minsalud, 2014).

Caracterización de la morbilidad. Con base en la normatividad expedida por Minsalud se definió el análisis de una serie de indicadores, generados por EAPB y la entidad territorial (ET), los cuales se subdividen en tres lineamientos (Dirección de Epidemiología y Demografía, 2017):

a. Principales causas de morbilidad atendida: haciendo uso de la lista de carga global de enfermedad modificada por Minsalud (2013b), mediante indicadores de porcentaje de población atendida.
b. Morbilidad trazadora: sucesos de interés para la salud pública del territorio que sirven para medir el bienestar de sus pobladores (Departamento de Caldas, Observatorio de Salud, s. f.). Principalmente a través de indicadores de tasas de prevalencia e incidencia.
c. Patologías y eventos de alto consumo: Minsalud, mediante el Decreto 2699 de 2007, estableció la Cuenta de Alto Costo, la cual tiene definidas las patologías de alto costo para el sistema de salud, mediante indicadores de número de personas y costos, por patología de alto consumo.

En el caso de los dos primeros, se referencia al Sispro como la fuente de consulta de los indicadores y, en el tercero, es la EAPB, con sus registros de costos, la que debe elaborar los indicadores planteados (Dirección de Epidemiología y Demografía, 2017).

Es de especial importancia la morbilidad trazadora, dado que contiene las tasas de prevalencia e incidencia de las ECNT —diabetes mellitus (DM), hipertensión arterial (HTA), enfermedad pulmonar obstructiva crónica, enfermedad renal crónica (ERC) y patologías cardiovasculares—. Aunque no se analizan las patologías de interés para el territorio, en algunas, como la DM o la HTA, se coincide con estas. De igual modo, la comparación con las prevalencias de la EAPB está sujeta a la disponibilidad y a la correspondencia del Sispro en sus indicadores de caracterización.

En lo referente a las patologías y eventos de alto consumo, debido a las limitantes de acceso de información relacionada con la distribución de costos de los usuarios asignados, por parte de las EAPB, se optó por analizar a los usuarios con patologías consideradas por la Cuenta de Alto Costo. Aquí se debe clarificar que dicha cuenta tiene estipulado un reporte periódico, por parte de las EAPB, de patologías donde no todas son de alto costo, pues algunas corresponden a diagnósticos que las consideran como precursoras de patologías de alto costo. Por ejemplo, la HTA es precursora de la ERC. Así que se determinó para esta metodología como alto costo el análisis de la tasa de prevalencia, según el código CIE-10, de neoplasias malignas, VIH/sida, artritis reumatoide y hemofilia. Es importante aclarar que, si bien la ERC en uno de sus estadios es considerada enfermedad de alto costo, no se agrupó dentro del grupo de enfermedades de alto costo, sino que se desagregó junto con las patologías priorizadas por el modelo de atención Sermás.

Definición y clasificación de los grupos de riesgo

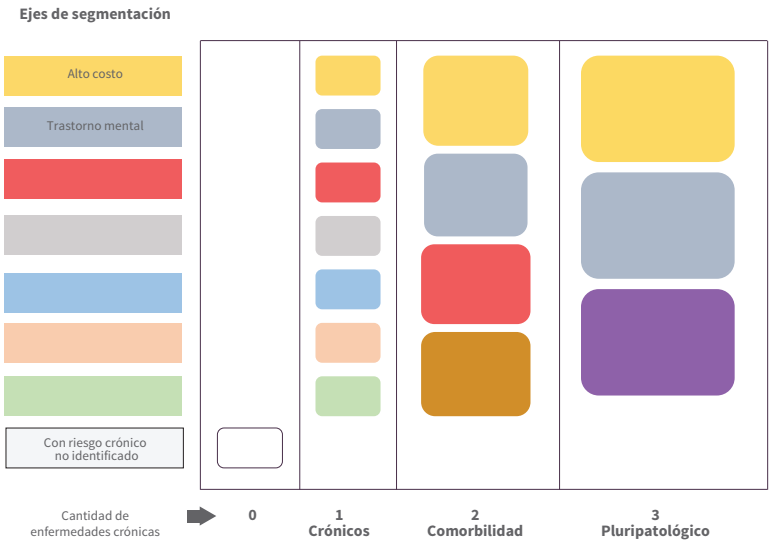
Para esto se plantea clasificar a los usuarios usando la metodología de segmentación basada en el criterio clínico de patologías.

La identificación de los riesgos, en los ítems de caracterización de las causas de atención, las enfermedades priorizadas por Sermás y las relacionadas con la Cuenta de Alto Costo definen las patologías, según su codificación en el CIE-10, empleadas como ejes de segmentación. Adicionalmente, se define otro criterio de clasificación, "cantidad de patologías crónicas", a partir del número de enfermedades, ejes de segmentación, presentes por usuarios: a) crónico (una patología), b) comorbilidad (presencia simultánea de dos patologías) y c) pluripatológico (presencia simultánea de tres o más patologías).

Teniendo en cuenta los dos criterios de segmentación, estos se combinaron a fin de obtener los diferentes segmentos o grupos de riesgo. Así, también, hay otros aspectos para considerar: a) los usuarios que, con la caracterización no se les identifica patología crónica, podrán ser agrupados en un solo segmento o en varios teniendo en cuenta la relevancia en la demanda de atención, pues deberán ser objeto de las estrategias de promoción de la salud y prevención de la enfermedad; b) las patologías, como el trastorno mental, que exigen estrategias adicionales al manejo fisiológico y de alto costo por su impacto, que pueden considerarse segmentos transversales, visibles en línea de comorbilidad y pluripatológicos, y c) en la combinación de los ejes de segmentación (comorbilidad y pluripatológicos), la cantidad de segmentos deberá ser la menos posible, recomendación tomada del modelo del Departamento de Salud de Navarra que inicialmente los clasificó en 24 segmentos y optó por un menor número, luego de analizar la utilidad clínica, de manera que su implementación resultara más operativa (Gobierno de Navarra-Departamento de Salud, 2013).

Consecuente con los criterios de segmentación, los grupos de riesgo se muestran en un gráfico de dos ejes (figura 1): en la parte horizontal, se ubican las categorías que indican el número de patologías crónicas, y en el lado vertical, se relacionan los ejes de segmentación. Los segmentos son localizados en las intersecciones entre las categorías de ambos ejes, representados por un rectángulo con un color personalizado y en su interior se registra la cantidad de usuarios del segmento. En el caso de los segmentos de más de una patología, se registra adicionalmente el nombre del segmento.

Figura 1. Clasificación de los grupos de riesgo priorizados para el modelo Sermás



Fuente: elaboración propia.

En la distribución de los costos de atención en salud se ha identificado que toda la población —así esté clasificada en un grupo específico de enfermedad, con comportamiento y consumo de recursos definidos— no presenta un riesgo homogéneo, es decir, existen subgrupos dentro de una misma población que, por sus características, especialmente por su estado de salud, tienen mayor probabilidad de consumir servicios (Dirección de Epidemiología y Demografía, 2017). De modo que el enfoque metodológico establecido por la IPS Universitaria plantea una estratificación del riesgo de los grupos identificados que permita planear e implementar estrategias diferenciales de GIRS y una gestión clínica con criterios de calidad: eficiente, pertinente y oportuna.

Estratificación del riesgo

Con el uso de un modelo descriptivo, se estratificaron los grupos de riesgo, y para esto se calcularon las variables "descontrol" y "uso de recursos", las cuales originaron el nivel de riesgo determinado como *severidad*.

Descontrol. Utilizando las variables número de urgencias y hospitalizaciones no programadas de la base de datos consolidada, se tomaron las frecuencias de dichas variables y se multiplicaron por el peso ponderado definido. A través de la metodología de análisis de expertos, grupo conformado por personal asistencial y administrativo de la IPS Universitaria, se validaron y aprobaron los pe-

sos ponderados y los umbrales definidos. De ahí que para el periodo de análisis propuesto, la ausencia de eventos en ambas variables se considera "control"; una sola urgencia, "bajo descontrol"; dos urgencias, "moderado descontrol", y más de dos urgencias o el suceso de hospitalizaciones no programadas, "alto descontrol".

Uso de recursos. A través de la metodología de análisis de expertos, se les asigna pesos ponderados a los servicios y se definen umbrales de clasificación, teniendo en cuenta los análisis de costos por servicio, el tiempo que abarca el periodo de análisis y las referencias bibliográficas sobre el consumo de recursos de pacientes crónicos (Agency for Clinical Innovation, 2014; Dirección de Epidemiología y Demografía, 2017). Con lo anterior, el uso de recursos se obtiene de comparar con los umbrales la sumatoria de la multiplicación de los pesos ponderados por la frecuencia de las variables definidas: número de consultas médicas generales y especializadas, las cirugías y los días de estancia hospitalaria por tipo de cama, de la base de datos consolidada. Los medicamentos, los insumos y las ayudas diagnósticas no se incluyeron, pues en gran medida están implícitos en la atención de cirugías, estancias hospitalarias y consultas (tabla 1).

Tabla 1. Pesos ponderados de las variables y umbrales de clasificación del uso de recursos de los usuarios del modelo Sermás

Pesos ponderados	Umbrales
Consulta externa general: 1	Bajo uso de recursos: 0-10
Consulta externa especializada: 2	
Estancia día cama hospitalaria (incluye urgencias >24 h): 5	Moderado uso de recursos: 11-25
Estancia día unidad de cuidados especiales: 10	
Estancia día unidad de cuidados intensivos: 15	Alto uso de recursos: 26 y más
Cirugía: 20	

Fuente: elaboración propia.

Las variables de descontrol y uso de recursos son el insumo para estratificar los grupos de riesgo; pero individualmente permiten el seguimiento periódico en apoyo a la GIRS, sobre todo en la gestión clínica.

Severidad. Se emplea una matriz que ubica verticalmente las tres categorías del uso de recursos y horizontalmente las cuatro de descontrol. Las celdas de intersección están sombreadas con tres colores y cada uno corresponde a un nivel de severidad: baja, moderada y alta (figura 2).

Figura 2. Estimación de la severidad en los usuarios atendidos bajo el modelo Sermás en la IPS Universitaria, según descontrol y uso de recursos

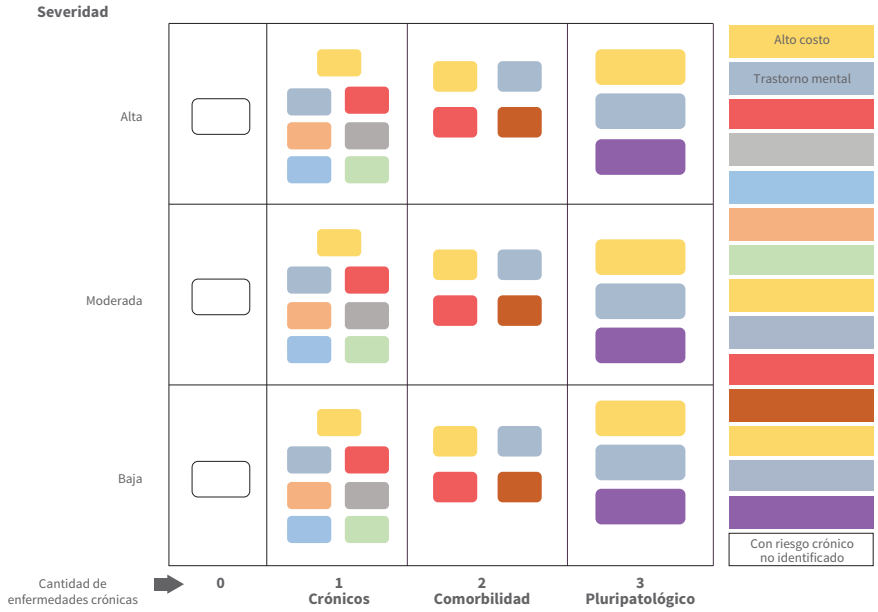


Fuente: elaboración propia.

Con el nivel de severidad se establece la estratificación de los usuarios, y este es el factor que subdivide los grupos de riesgo identificados; de modo que las estrategias de atención se ajusten no solo a la(s) patología(s) crónica(s), sino también a la condición médica y el uso de recursos. Adicionalmente, se decidió incluir en la estratificación el grupo de usuarios "con riesgo crónico no identificado", ya que permiten identificar resultados importantes en las variables de descontrol, uso de recursos y, por ende, en la severidad.

Se decidió emplear un esquema similar al usado en la clasificación de grupos de riesgo. El eje horizontal indica el número de patologías crónicas y el eje vertical relaciona las tres categorías del nivel de severidad. Los segmentos estratificados son ubicados en las intersecciones entre las categorías de ambos ejes, representados por un rectángulo con un color personalizado y en su interior se registra la cantidad de usuarios del segmento. Esta representación gráfica se acompaña de las convenciones de colores correspondientes a los segmentos (figura 3).

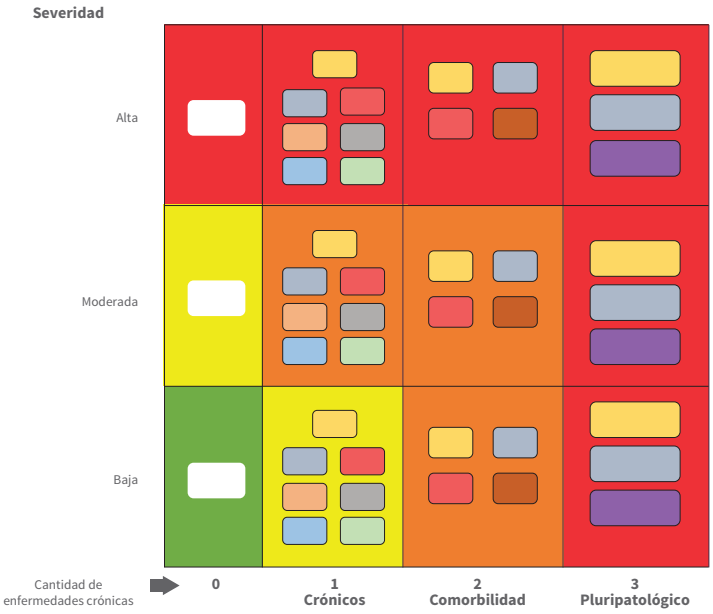
Figura 3. Estatificación de los grupos de riesgo priorizados para el modelo Sermás



Fuente: elaboración propia.

Como ya se mencionó, estratificar los grupos de riesgo es el insumo para identificar la pertinencia, según Kaiser Permanente, y relacionarlo con el tipo de estrategias de atención y la planeación en los niveles estratégico y táctico. Pero no es el único elemento, ya que se dispone de variables como el género, el ciclo vital y el lugar de residencia, que posibilitan una mejor planeación, con enfoque diferencial e integralidad; esto desde la aplicación de la metodología definida y con las fuentes de información disponible. Por otra parte, la planeación operativa de estos grupos exige más elementos para su abordaje, y allí la estratificación de los grupos de riesgo es un elemento de priorización para la fase inicial de la gestión clínica, definida por el modelo Sermás. De ahí que, ubicados en la figura 4 y haciendo un símil con los diagramas de calor empleados en la gestión del riesgo, los grupos definidos en la parte superior y en la columna de pluripatológicos son los de mayor prioridad; seguidos por los comórbidos no severos y crónicos con severidad moderada.

Figura 4. Representación de la priorización de los grupos de riesgo estatificados del modelo Sermás, mediante un mapa de calor



Fuente: elaboración propia.

Priorización, reenfoque y seguimiento de resultados dentro de un modelo de atención

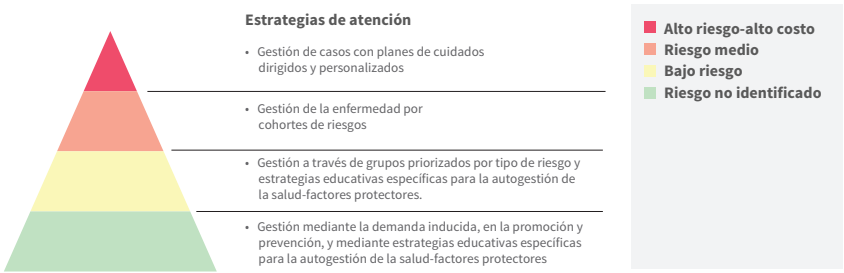
Acorde con los principios de atención diferencial y cogestión del riesgo, el modelo de atención Sermás de la IPS Universitaria plantea un reenfoque que permita diferenciar y priorizar los grupos de riesgo identificados durante la estratificación de la población asignada (figura 4) y establecer estrategias de atención acordes con las características y necesidades de dicha población. Lo anterior, a partir de la adopción de lo planteado en el modelo Kaiser Permanente, y el ajuste a las rutas de atención diferencial dependiendo del grupo, del estado de salud y del nivel de riesgo de los usuarios que acceden a través de las rutas a los diferentes ámbitos de atención dispuestos como el ambulatorio, hospitalario y domiciliario. Adicionalmente, se realiza la clasificación de la población en cuatro niveles de riesgo según el nivel de complejidad del paciente crónico (figura 4). Cada nivel de riesgo precisa una oferta de servicios diferente, ajustada a las características y necesidades de cada estrato. Es decir, se hace una reasignación de los recursos disponibles con el fin de maximizar el resultado final, el cual consiste en optimización de recur-

sos hospitalarios, mayor beneficio para el paciente crónico al que se le ayuda a controlar su enfermedad, menor número de complicaciones y disminución de la probabilidad de sufrir eventos adversos derivados de la hospitalización (Cotta et al., 2001). A continuación, se describen cada uno de los niveles de la pirámide de Kaiser Permanente:

En la base de la pirámide nos encontramos con el grupo poblacional sano o con factores de riesgo no identificados, es decir, pacientes que se encuentran dentro de la base de datos de la EAPB y cuya única información son datos básicos de sexo, edad, teléfono, dirección, lo cual no permite ubicarlos en algún estrato o valorar su riesgo:

- *Primer nivel*: población con riesgo bajo. Está compuesto por pacientes con riesgos materializados con una enfermedad crónica e infecciosa (tuberculosis) y que su severidad dio moderada.
- *Segundo nivel*: población con riesgo medio. Está compuesto por pacientes que tienen dos enfermedades crónicas o una sola enfermedad con complicaciones por otras condiciones como trastornos mentales, que requieren la atención en distintos ámbitos asistenciales.
- *Tercer nivel*: población con riesgo alto. Está compuesto por pacientes de alta complejidad, que tienen tres o más diagnósticos de patologías crónicas o lesión de órgano blanco; además, tienen determinantes sociales o están sin estos, que inciden en su condición de morbilidad. También pueden estar pacientes con una sola enfermedad crónica, pero que tuvieron alto nivel de descontrol de la enfermedad o uso de recursos. En esta clasificación se incluyen los eventos en salud catalogados como enfermedades catastróficas (alto costo). En la figura 5 se presenta para cada nivel la estrategia de atención que se va a desarrollar.

Figura 5. Grupos distribuidos por riesgo y estratégicas de atención



Fuente: adoptado de Kaiser Permanente por la IPS Universitaria.

A continuación, se definen las estrategias de atención:

- *Gestión de casos*: se elabora un plan de cuidados individualizado para cada paciente, en coordinación con el médico tratante y sustentado en registros electrónicos a los que tiene acceso el resto del equipo asistencial. Estos planes incorporan recomendaciones y evidencias procedentes de guías de práctica clínica.
- *Gestión de la enfermedad por cohortes*: la gestión de enfermedades es un sistema de intervenciones en salud coordinadas y cuyo objetivo es mejorar la salud en un grupo poblacional. Ello incluye la comunicación con la población y la promoción de los autocuidados por parte de los pacientes. Proporciona el seguimiento eficiente y costoefectivo de grupos poblacionales con patologías que comparten factores de riesgo relacionados.
- *Educación para la autogestión de la salud-factores protectores*: requiere enfoques pedagógicos diferenciales por cada nivel de riesgos, orientados a fortalecer las capacidades para el cuidado de la salud y la vida, generando cambios positivos evidenciables en los estilos de vida que impliquen atenuar los factores de riesgo y fortalecer los factores protectores prioritariamente en nutrición saludable, salud mental y actividad física.
- *Gestión de demanda inducida*: organiza, incentiva y orienta a la población hacia la utilización de los servicios de protección específica y detección temprana y la adhesión a los programas de control. Esta estrategia fue definida por la IPS Universitaria para el abordaje de la población con factores de riesgo aún sin identificar.

De acuerdo con los resultados de la estratificación del riesgo y la definición de las estrategias de atención según los niveles de riesgo identificados, se procedió a planificar el proceso de atención, definiendo el ámbito de atención (domiciliario o ambulatorio) a la población clasificada como alto riesgo-alto costo, riesgo medio y bajo riesgo.

Para el seguimiento del resultado y la evaluación de la gestión del riesgo se utilizaron los indicadores de descontrol, uso de recursos y severidad; estos nos permiten identificar desviaciones en el modelo y, a partir de su análisis, tomar acciones de mitigación y ajustar el modelo.

Según el periodo definido, se realiza un informe con los resultados de la medición de estos indicadores y se organiza una base de datos donde se identifican los pacientes y sus resultados actuales e históricos de estos indicadores. Con esta información se analizan los casos de alta severidad y se desarrollan acciones de mitigación de riesgos alineados al modelo de atención, que permitan impactar positivamente el indicador. Para estos se plantea un estudio de casos, mediante la evaluación de la historia clínica de cada paciente (auditoría clínica), se identifican las causas de los resultados no deseados en estos indicadores y se definen acciones correctivas y estrategias de atención alineadas al modelo que hubiera podido evitar estos resultados.

Resultados

Una vez definido el enfoque metodológico que se va a desarrollar en cada uno de los componentes, se aplica la metodología para la identificación y el seguimiento del riesgo en salud de una cohorte poblacional atendida bajo el modelo de cogestión del riesgo Sermás, a la cohorte poblacional perteneciente a una EAPB y cuya característica principal es que cumple los criterios de pluripatología. El número promedio mensual de pacientes asignados a la cohorte es de 4830 (± 80). A continuación, se presentan los resultados obtenidos del desarrollo de esta experiencia.

Gestión de los datos empleados para el desarrollo del modelo con la población asignada

Se tuvo acceso a la base de datos única de afiliados, que mensualmente comparte la EAPB con la IPS Universitaria desde abril del 2017, en su compromiso de reportar la información de los beneficiarios de Sermás. Así mismo, se determinó usar la información de dicha población registrada en el portal de aplicaciones de la IPS Universitaria, correspondiente a la demanda atendida entre abril del 2017 y noviembre del 2018: registros derivados de las consultas en los servicios de urgencias, internación hospitalaria, consulta externa, cirugía, ayudas diagnósticas y laboratorio clínico. De ambas fuentes de información se extrajeron las variables establecidas en la metodología y se obtuvo una sola base de datos, sometida a los procedimientos de gestión del dato, especificados en el enfoque metodológico, acorde con los criterios establecidos.

Identificación del riesgo

El enfoque metodológico se aplica sobre la cohorte poblacional de 4811 usuarios vigentes a noviembre del 2017, con los componentes mínimos establecidos de caracterización demográfica y caracterización de la morbilidad; en este último componente, se realizó un análisis diferencial de la morbilidad, por ciclos vitales y género.

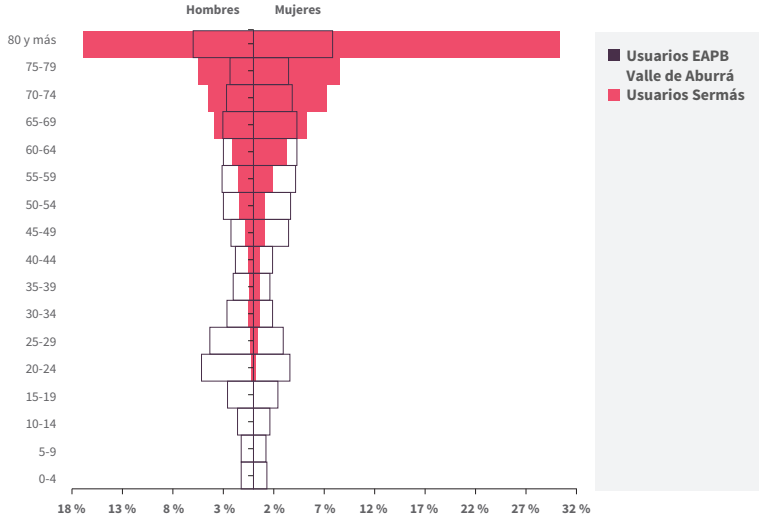
Caracterización demográfica

La cohorte poblacional asignada tiene en promedio 76 años de edad (± 15), las mujeres (con el 61 %) predominan sobre los hombres. El grupo etario más representativo es el de 80 o más años, con 2,78 (47,3 %) personas, entre estas 1466 mujeres.

Por ciclos vitales, se identifica la ausencia de menores de 20 años; hay 30 jóvenes, que corresponden al 0,6 %; 567 adultos (11,8 %), y 4214 (87,6 %) personas mayores. En los jóvenes y adultos la distribución por género es equitativa, y en las personas mayores 2630 (62 %) son mujeres. Al comparar la cohorte asig-

nada con relación a los afiliados de la EAPB residentes en Medellín y municipios aledaños (Valle de Aburrá), se identificaron grandes diferencias en las pirámides poblacionales, entre estas: a) la proporción de personas mayores en la EAPB es del 33,2 % y en la cohorte asignada es del 82,2 %; b) la razón hombre-mujer de la EAPB es mayor, pues por cada mujer hay 0,93 hombres, mientras que en los usuarios del modelo Sermás es de 0,64. Ello sugiere que en su estructura demográfica, la cohorte asignada (1,2 % de los afiliados a la EAPB) corresponde a una población de personas mayores, pues la diferencia sustancial en los últimos quinquenios y en la distribución por género sugiere que son usuarios de alto riesgo en salud, probablemente complejos, y que requieren un enfoque en la atención diferencial. Esto si se considera que la presencia de múltiples enfermedades es muy frecuente en edades avanzadas y que el número de enfermedades aumenta con la edad (Abizanda Soler et al., 2010) (figura 6).

Figura 6. Comparativa de pirámides poblacionales de los usuarios asignados para ser atendidos en la IPS Universitaria con el modelo Sermás y la de todos los afiliados a la EAPB residentes en el Valle de Aburrá, noviembre de 2017



Fuente: elaboración propia.

Caracterización de la morbilidad

Causas de morbilidad atendida. Tras ocho meses de atención, se identificaron, por grandes grupos de causas de morbilidad, las atenciones de los usuarios asignados. Los diagnósticos registrados en la prestación de los servicios indican que el 90,4 % de los usuarios demandó atención por enfermedades no transmisibles (ENT), seguidos con un 14,2 % de las condiciones transmisibles y nutricionales, y con un 9,4 % por lesiones. Es de anotar que estas causas no son excluyentes. En Colombia, las ENT representan el 82 % de la demanda de servicios en salud anual (periodo mayor al medido) de atenciones en el adulto mayor (Dirección de Epidemiología y Demografía, 2018). Es importante resaltar las otras dos categorías, pues tanto la malnutrición, incluido el sobrepeso, como las limitaciones de funcionalidad son condiciones en alto riesgo de sufrir por parte de la población adulta mayor (Arango y Ruiz, 2011). Por lo tanto, para el modelo Sermás y en el abordaje integral del usuario, intervenir complementariamente el componente nutricional es necesario en la gestión de los riesgos en salud. La funcionalidad es un parámetro indispensable para definir el ámbito de atención, que puede ser el domiciliario o el ambulatorio.

Por ciclos vitales y género, en porcentajes, las ENT son la causa de mayor atención, y se destacan en las personas mayores, sobre todo en los hombres, con el 92 %. Los jóvenes, especialmente las mujeres, constituyen en porcentaje el grupo más atendido por condiciones transmisibles y nutricionales, y los hombres adultos se destacan en las atenciones por lesiones. Llama la atención el porcentaje de diagnósticos correspondientes a síntomas, signos y hallazgos anormales clínicos y de laboratorio no clasificados en otra parte, ya que les resta precisión en la identificación de los riesgos en los usuarios, pues en algunos casos son registrados cuando no se tiene claridad en el diagnóstico, y ello propicia errores en la calidad del dato (tabla 2).

Tabla 2. Causas de morbilidad atendida en la IPS Universitaria de los usuarios asignados para ser atendidos con el modelo Sermás, según las grandes causas, noviembre de 2017

Grandes grupos	Juventud (en porcentaje)		Adultez (en porcentaje)		Persona mayor (en porcentaje)	
	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
Condiciones transmisibles y nutricionales	19	7	11	10	15	13

Grandes grupos	Juventud (en porcentaje)		Adultez (en porcentaje)		Persona mayor (en porcentaje)	
	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
Enfermedades no transmisibles	75	71	87	89	90	92
Lesiones	0	0	9	12	9	10
Síntomas, signos y hallazgos anormales clínicos y de laboratorio, no clasificados en otra parte	6	29	18	17	24	23

Fuente: elaboración propia.

Los grandes grupos de causas de morbilidad están conformados por subgrupos. En el caso de las ENT, que constituyen el mayor porcentaje de atención, se identificaron los usuarios atendidos por los subgrupos, según ciclo vital y género, presentados en un esquema de mapa de calor, con los resultados subdivididos en cuatro categorías empleando medidas de posición como los cuartiles. Se encontró que las enfermedades cardiovasculares son las más representativas, con el 66,6 % de los usuarios atendidos; seguidas de las condiciones neuropsiquiátricas, con el 26,7 %, y la DM, con el 22,3 %. Por ciclo vital, las personas mayores son las más atendidas en los subgrupos ya mencionados; adicionalmente, por género, se resalta el 26 % de usuarios atendidos por enfermedades genitourinarias en los hombres y el 20 % de mujeres con enfermedades músculo-esqueléticas. Con respecto a los usuarios en adultez, las principales causas de atención son las mismas de la gran causa ENT. Se destaca que en las mujeres el porcentaje de atención es mayor para las enfermedades cardiovasculares y la DM; esta última es superior al promedio de todos los usuarios. En lo referente a los usuarios jóvenes asignados, el porcentaje de hombres atendidos por causa de condiciones neuropsiquiátricas es el mayor (43 %) de todos los ciclos vitales. También se destacan el 19 % de mujeres atendidas por enfermedades de la piel y el 13 %, por enfermedades respiratorias (tabla 3).

Tabla 3. Causas de morbilidad atendida de los usuarios asignados para ser atendidos con el modelo Sermás en la IPS Universitaria, por subgrupos de la gran causa ENT, noviembre de 2017

Subgrupos enfermedades no transmisibles	Juventud (en porcentaje)		Adultez (en porcentaje)		Persona mayor (en porcentaje)	
	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
Enfermedades cardiovasculares	6	14	56	49	67	72
Condiciones neuropsiquiátricas	25	43	34	39	27	23
Diabetes mellitus	19	14	27	16	23	21
Enfermedades respiratorias	13	0	9	11	20	21
Enfermedades genitourinarias	6	7	12	18	15	26
Enfermedades músculo-esqueléticas	6	0	19	10	20	12
Trastornos endocrinos	0	7	14	8	14	8
Enfermedades digestivas	6	0	6	7	10	10
Enfermedades de la piel	19	0	10	8	8	8
Enfermedades de los órganos de los sentidos	0	0	3	5	6	6
Neoplasias malignas	0	0	2	3	3	5
Otras neoplasias	0	0	4	1	2	2
Anomalías congénitas	13	0	2	1	0	0
Condiciones orales	0	14	1	0	0	0

Fuente: elaboración propia.

Patologías priorizadas por el modelo Sermás. Se estimó la prevalencia de dichas enfermedades en los usuarios asignados, con análisis por ciclo vital y género, presentado a través de una tabla con mapa de calor. Se halló que por cada 100 usuarios asignados, hay 84 con HTA, 57 con DM, 46 casos con ERC, 43 con EPOC y 42 con trastornos mentales. Adicional a la HTA, se identificaron las otras enfermedades priorizadas, correspondientes al sistema circulatorio, donde la más relevante es la insuficiencia cardiaca, con 23 casos entre 100 usuarios; seguida

por las valvulopatías y las arritmias, con 19 casos. En lo relacionado con el ciclo vital, en los tres grupos la HTA y la DM son las causas de mayor prevalencia, exceptuando los jóvenes, en quienes la segunda causa es el trastorno mental, con 33 casos por cada 100 usuarios jóvenes. Con relación al género, usualmente los hombres registran las mayores prevalencias; sin embargo, en los ciclos vitales persona mayor y adultez, las mujeres registran mayor prevalencia, con 44 casos por cada 100 usuarios. En el grupo de adultez, la prevalencia de mujeres con HTA es un poco mayor, 72 casos por cada 100 mujeres adultas; mientras que en los hombres es de 67. Entre los jóvenes, los hombres tienen prevalencias mayores a las mujeres, y no se identificó prevalencia de enfermedades priorizadas del sistema circulatorio diferentes a la HTA (tabla 4).

Tabla 4. Prevalencia de las patologías priorizadas por el modelo Sermás por cada 100 usuarios asignados a la IPS Universitaria para ser atendidos con el modelo Sermás, según ciclo vital y género, noviembre de 2017

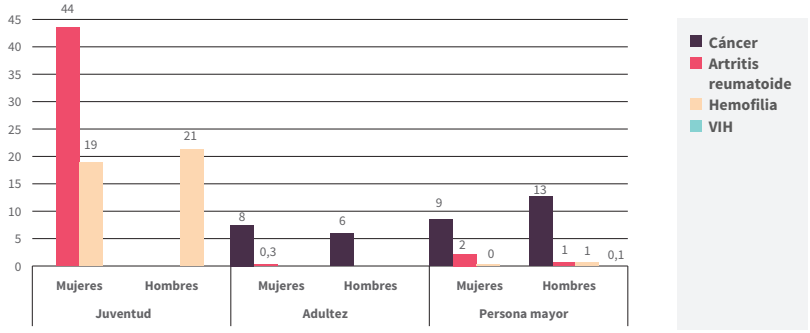
Patologías priorizadas por Sermás	Juventud		Adultez		Persona mayor	
	Mujeres	Hombres	Mujeres	Hombres	Mujeres	Hombres
Hipertensión arterial	31	36	72	67	84	89
Diabetes mellitus	25	36	47	48	56	61
Enfermedad renal crónica	19	21	20	32	44	58
Enfermedad pulmonar obstructiva crónica	6	14	19	23	44	49
Trastorno mental	25	43	44	37	44	40
Insuficiencia cardiaca	0	0	10	15	21	30
Valvulopatías y arritmias	0	0	9	11	18	24
Cardiopatías	0	7	7	12	15	24
Accidente cerebrovascular	0	0	10	10	17	20

Fuente: elaboración propia.

Hay que señalar que la metodología contempla el análisis comparativo, que permite establecer el nivel de riesgo de la prevalencia hallada con una población de referencia, que es la de la EAPB. Sin embargo, el Sispro o las demás fuentes de información no suministran datos de prevalencia de la mayoría de patologías priorizadas y con las especificidades de los usuarios Sermás. En razón a lo anterior, teniendo en cuenta el ciclo vital, se compararon resultados de investigaciones identificando diferencias significativas, por cada 100 personas. En los usuarios Sermás hay: a) 30 casos más de HTA, b) 46 casos más de DM (Arango y Ruiz, 2011) y c) 25 casos más de EPOC (Bernal Pinilla, 2015).

Patologías de la Cuenta de Alto Costo. Con metodología similar a la empleada en el análisis y presentación de las patologías priorizadas, se caracterizaron las cuatro patologías de la Cuenta de Alto Costo definidas en este componente. El cáncer es la de mayor prevalencia, con 10 casos por cada 100 usuarios de la cohorte; le sigue la hemofilia y la artritis reumatoide, con un caso en ambas. El análisis por ciclo vital, en general, muestra que el cáncer es la patología de mayor prevalencia, con excepción de los jóvenes, en quienes la artritis reumatoide en las mujeres es la más relevante, seguida por la hemofilia en ambos géneros. Lo anterior confirma el alto riesgo en salud de los usuarios asignados (figura 7).

Figura 7. Prevalencia de patologías de la Cuenta de Alto Costo, por cada 100 usuarios asignados a la IPS Universitaria para ser atendidos con el modelo Sermás, según ciclo vital y género, noviembre de 2017



Fuente: elaboración propia.

En definitiva, la situación de salud de la cohorte da cuenta del reto que asumió la IPS Universitaria, especialmente en usuarios con patologías crónicas como la DM y la HTA, consideradas precursoras de enfermedades de alto costo como la ERC (Cuenta de Alto Costo, 2017). Tras lo anterior, y fiel a la idea de que "es un error pensar que la presencia de una enfermedad en la vejez significa que la persona ya no es saludable" (Organización Mundial de la Salud, 2015),

la IPS Universitaria ha asumido el reto de la gestión del riesgo en salud de esta cohorte poblacional, partiendo del conocimiento adecuado de los usuarios, como componente de la gestión del riesgo, ajustado a los lineamientos teóricos y metodológicos de modelo Sermás.

Clasificación en grupos de riesgo

Con la identificación del riesgo, mediante la metodología de caracterización, se establecieron los criterios diagnósticos o clínicos de clasificación de la población usuaria asignada. Acorde con las tres perspectivas de riesgo definidas: causas de consulta, enfermedades priorizadas o de interés específico para el modelo Sermás y las patologías de alto costo.

Así, pues, las causas de consulta indicaron que las ENT son las de mayor demanda, entre las cuales se destacan los subgrupos que contienen las principales enfermedades crónicas: enfermedades cardiovasculares, condiciones neuropsiquiátricas, diabetes mellitus, enfermedades respiratorias y genitourinarias. También se observó que, de manera importante, las patologías priorizadas por Sermás, de riesgo cardiovascular, están presentes en la población usuaria, y en lo correspondiente a las enfermedades de la Cuenta de Alto Costo, la prevalencia establecida, especialmente de cáncer, indica la necesidad de ser consideradas.

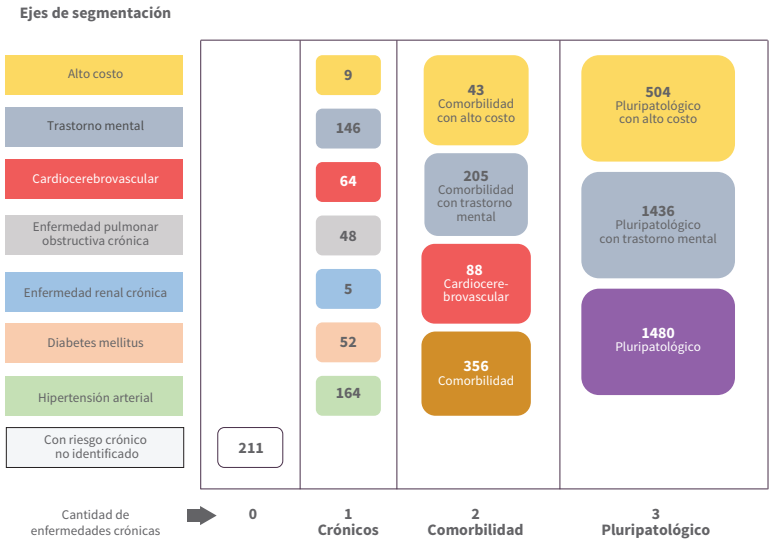
Teniendo claridad de los ejes de segmentación (primer criterio clínico) y la cantidad de patologías crónicas por usuario (segundo criterio clínico), como elementos en la clasificación de los grupos de riesgo priorizados, según las metodologías de segmentación basadas en criterios clínicos, se definieron siete diagnósticos como ejes de segmentación: a) HTA, b) DM, c) EPOC, d) ERC (en estadios 3A, 4B y 5), e) cardiocerebrovascular, f) trastorno mental y g) alto costo. Se consideró importante incluir estos dos últimos como ejes trasversales, debido al impacto en el estado de salud de los usuarios y en los costos de atención.

Es importante resaltar que el eje de segmentación cardiocerebrovascular está constituido por usuarios con alguna de las patologías cardiovasculares priorizadas por el modelo Sermás: insuficiencia cardiaca, accidente cerebrovascular, cardiopatía isquémica, valvulopatías y arritmias. En el caso de los segmentos de comorbilidades, se acordó un grupo en el cual ambas fuesen cardiovasculares, otro con las diferentes patologías del riesgo cardiovascular y los dos restantes en que se identifican las patologías de alto costo y trastorno mental que, por las razones ya mencionadas, implican su visualización para las estrategias de atención y consideraciones contractuales. Algo similar se propone en los pacientes pluripatológicos, al establecer tres segmentos: un grupo con tres o más diagnósticos crónicos, otro donde mínimo uno de ellos es alto costo y el restante con usuarios que entre sus múltiples morbilidades tienen el trastorno mental (figura 8).

En definitiva, se obtuvieron 15 segmentos o grupos de riesgo priorizados, en los que se clasifica la población usuaria asignada de la siguiente manera: 211

usuarios en el grupo con riesgo no identificado; 488 con una sola enfermedad crónica en 7 grupos (entre las cuales se destaca el de HTA, con 164 personas; seguido por el de trastorno mental, con 146, y el cardiocerebrovascular, con 64 personas); 692 usuarios con dos patologías crónicas, reunidos en 4 grupos (43 en el de comorbilidad con alguna de alto costo y 205 en comorbilidad que incluye el trastorno mental); 3420 usuarios pluripatológicos, de los cuales 504 están en el grupo alto costo; 1436 en el que considera el trastorno mental, y 1480 en el grupo donde se combinan los diferentes ejes de segmentación no transversales (figura 8).

Figura 8. Clasificación en grupos de riesgo priorizados de los usuarios asignados para ser atendidos con el modelo Sermás en la IPS Universitaria, noviembre de 2017



Fuente: elaboración propia.

Estratificación del riesgo

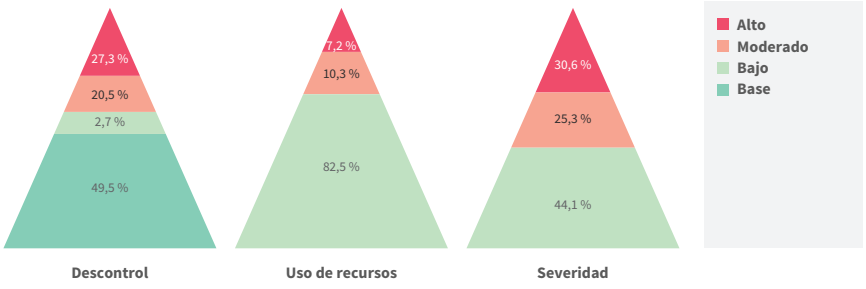
El nivel de riesgo de los grupos hallados se estableció con el grado de severidad de cada usuario de los diferentes grupos, calculado con los indicadores, descontrol y uso de recursos, ambos obtenidos a partir de la utilización de los servicios de salud.

Descontrol. En los usuarios, vigentes a noviembre del 2017, el 49,5 % se clasificó en control de su enfermedad y el otro 50,5 % presentó un grado de enfermedad no controlada, distribuida así: 2,7 % bajo, 20,5 % moderado y 27,3 % con un alto nivel de enfermedad no controlada (figura 9).

Uso de recursos. El 82,5 % de los usuarios del modelo Sermás se encontraba en bajo uso de recursos, el 10,3 % tenía un uso moderado y el 7,2 % restante se ubicó en alto consumo de recursos (figura 9). En el caso de los usuarios con alto consumo de recursos, el 68 % de la puntuación que los colocó en ese rango corresponde a las estancias de internación en cama hospitalaria normal; seguido por un 14 % de utilización del servicio de cirugía y un 7 % con estancia de internación en la unidad de cuidados intensivos.

Severidad. La combinación de los niveles de descontrol y de uso de recursos determinaron que el 44,1 % de los usuarios pertenece al rango de baja severidad; el 25,3 %, a moderada, y el 30,6 %, a alto (figura 9).

Figura 9. Estratificación del riesgo o nivel de severidad de los usuarios asignados para ser atendidos con el modelo Sermás, según las variables de descontrol y uso de recursos, noviembre de 2017

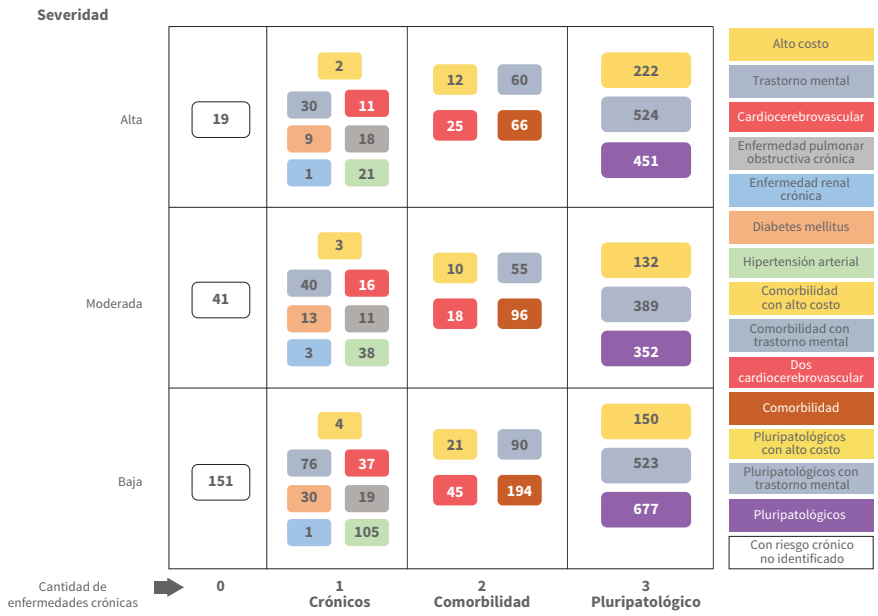


Fuente: elaboración propia.

Estratificación de los grupos de riesgo. Al aplicar el resultado de la severidad a los segmentos identificados, se establecieron los niveles de riesgo en salud que en el interior de cada segmento los diferencian y que se requieren para priorizar las estrategias de atención del modelo Sermás.

De ahí que todos los grupos de riesgo, incluyendo los de "riesgo no identificado", ubicaron usuarios en los tres niveles de severidad, por cantidad de enfermedades; en los grupos crónicos y comórbidos es mayor el número de usuarios en los niveles de severidad baja y estos decrecen al aumentar la severidad. Llaman la atención los grupos de usuarios que, producto del alto uso de recursos o descontrol, no deberían presentar alta severidad, como los 19 usuarios con riesgo no crónico o los 21 con HTA. Ello sugiere un seguimiento específico, ante la presencia de patologías crónicas ocultas o la demanda anormal de servicios; de igual modo, con los usuarios de baja severidad pertenecientes a grupos clínicamente considerados de alto riesgo, como los 175 usuarios de alto costo, 523 pluripatológicos con trastorno mental y 677 pluripatológicos (figura 10).

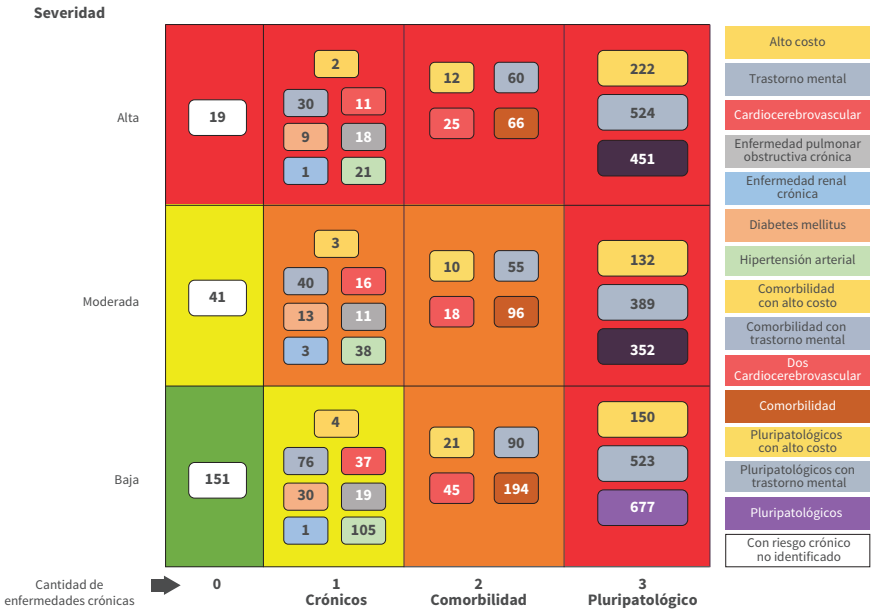
Figura 10. Desagregación de los grupos de riesgo, de acuerdo con el nivel de severidad, de los usuarios asignados para ser atendidos con el modelo Sermás en la IPS Universitaria, noviembre de 2017



Fuente: elaboración propia.

Se le aplicó la escala de calor a la estratificación, teniendo como nivel alto (rojo) a los pacientes que dieron alta severidad y a los pruripatológicos, por su condición de múltiples patologías; como nivel de riesgo medio (naranja), a los pacientes con severidad moderada, crónicos o comórbidos; comonivel de riesgo bajo (amarillo) a los pacientes no crónicos o con riesgo no identificado que tuvieron un nivel de severidad moderado, y los pacientes crónicos con severidad baja, y, finalmente, como riesgo no identificado (verde), a los pacientes con baja severidad y riesgo crónico no identificado (figura 11).

Figura 11. Desagregación de los grupos de riesgo, de acuerdo con el nivel de severidad, de los usuarios asignados para ser atendidos con el modelo Sermás en la IPS Universitaria, noviembre de 2017

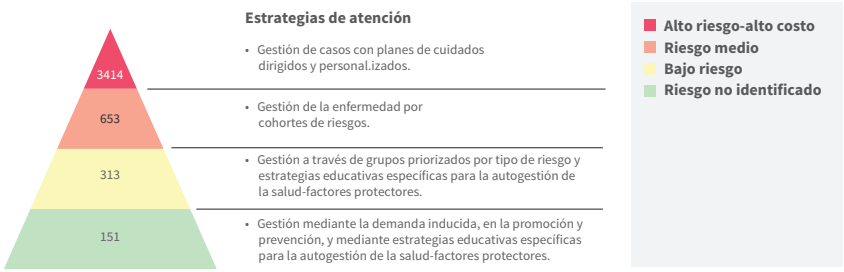


Fuente: elaboración propia.

Priorización y reenfoque dentro de un modelo de atención

Teniendo en cuenta los resultados de la estratificación del riesgo y de acuerdo con el modelo poblacional de Kaiser Permanente, se espera que en riesgo bajo esté el 80 % de la población, que el 15 % sea de riesgo mediano y que el 5 % restante esté en alto riesgo (Osakidetza, 2015). Así mismo, la metodología de Navarra expone en sus resultados valores porcentuales para tres niveles de la variable de clasificación de severidad: baja (68 %), mediana (22 %) y alta (10 %) (Gobierno de Navarra-Departamento de Salud, 2013). No obstante, en ninguna de las referencias sus rangos pueden emplearse como parámetros de comparación de los resultados obtenidos en la estratificación de la cohorte, pues corresponden a clasificaciones poblacionales aleatorias y no de grupos específicos. Por tal motivo, las metas y los parámetros de comparación para los resultados del descontrol, uso de recursos y severidad son una construcción institucional para cada cohorte durante el desarrollo del modelo Sermás (figura 12).

Figura 12. Desagregación de los grupos de riesgo, de acuerdo con el nivel de severidad, de los usuarios asignados para ser atendidos con el modelo Sermás en la IPS Universitaria



Fuente: elaboración propia.

Seguimiento de resultados y gestión en salud

Es importante mencionar que, individualmente, tanto el descontrol como el uso de recursos son empleados en la institución como indicadores para el seguimiento y priorización para la intervención de los usuarios. Así es como, a través de estos, se han identificado los usuarios que de modo recurrente son calificados como altamente descontrolados o con alto consumo de recursos, quienes —de acuerdo con las estrategias del modelo— son sujetos de atención caso a caso.

Los resultados presentados se obtuvieron después de desarrollar durante más de un año las estrategias del modelo de atención médica para cada uno de los estratos identificados de la población asignada. Las mediciones se llevaron a cabo cada cuatro meses, para los componentes de clasificación de riesgo (severidad), determinados por la falta de control de la enfermedad y el uso de recursos.

Descontrol. A noviembre del 2018, las cuatro mediciones llevadas a cabo revelaron una disminución porcentual de usuarios Sermás en alto descontrol, pasando del 27,3 % en noviembre del 2017 al 11,3 % en el mismo mes del 2018. Del mismo modo, se señala que el esfuerzo de tener el máximo porcentaje de usuarios en control se refleja especialmente en los meses de marzo y noviembre del 2018. Además, se debe indicar que, por las complejas condiciones de salud, tal vez es inevitable no hallar alto descontrol; así mismo, al no hallarse registros de este tipo de mediciones en poblaciones similares, la experiencia será la que defina las metas de descontrol (figura 13).

Figura 13. Comparación de cuatro mediciones del nivel de descontrol en los usuarios asignados para ser atendidos con el modelo Sermás en la IPS Universitaria

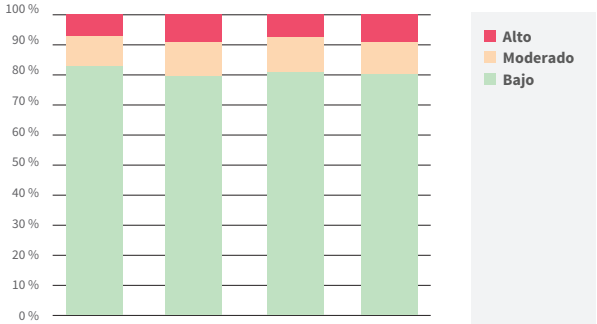


	Noviembre 2017	Marzo 2018	Julio 2018	Noviembre 2018
Alto	27,3 %	13,4 %	11,9 %	11,3 %
Moderado	20,5 %	1,4 %	5,3 %	3,3 %
Bajo	2,7 %	2,6 %	13,1 %	4,8 %
Control	49,5 %	82,6 %	69,6 %	80,7 %

Fuente: elaboración propia.

Uso de recursos. No se han identificado diferencias significativas (en porcentajes) en el consumo de recursos en las cuatro mediciones de este indicador. Vale la pena mencionar que gestionar el descontrol implica minimizar la frecuencia de urgencias y hospitalizaciones no programadas, lo cual lleva a disminuir recursos; pero los resultados han mostrado que hay una dinámica esperada de servicios en este tipo de grupos de riesgo, donde aproximadamente en el 80 % de los usuarios es baja y en el 11 % es moderada. Por tal razón, los análisis y la gestión clínica se han centrado sobre el 9 % con alto uso de recursos, y se han encontrado pacientes que en varias mediciones quedaron clasificados como altos consumidores de recursos: 209 usuarios en dos mediciones, 58 en tres y 18 en las cuatro (figura 14).

Figura 14. Comparación de cuatro mediciones del nivel de uso de recursos en los usuarios asignados para ser atendidos con el modelo Sermás en la IPS Universitaria

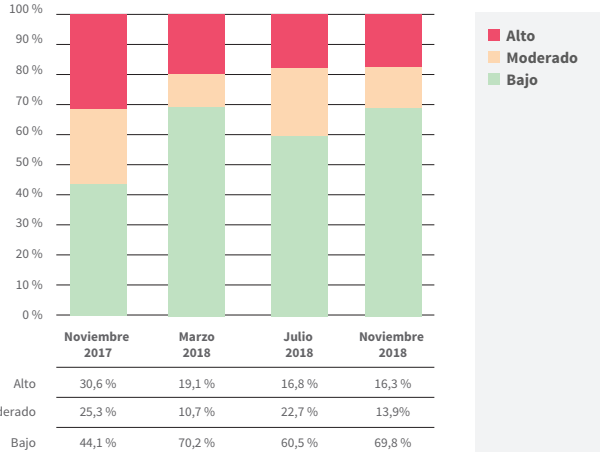


	Noviembre 2017	Marzo 2018	Julio 2018	Noviembre 2018
Alto	7,2 %	9,2 %	7,5 %	9,0 %
Moderado	10,3 %	11,4 %	12,2 %	11,0 %
Bajo	82,5 %	79,3 %	80,3 %	79,5 %

Fuente: elaboración propia.

Severidad. A noviembre del 2018, el 70 % de los usuarios mostró baja severidad; el 14 %, moderada, y 16 %, severa. Comparada con las demás mediciones, entre agosto y noviembre del 2018, porcentualmente la cohorte presentó menor severidad. Teniendo en cuenta el comportamiento de las variables con las que se calcula este indicador, visualmente la tendencia de disminución del alto descontrol se refleja en la clasificación de severidad, ante la relativa estabilidad del uso de recursos. Llama la atención que la baja severidad en ningún periodo ha superado el 70 %, y considerando el riesgo en salud identificado en esta población, el reto está en disminuir a los usuarios de alta severidad a moderada. De igual modo, el seguimiento periódico ha identificado a los usuarios recurrentes en la alta severidad, como insumo para la reorientación de las estrategias de atención, diferenciados por el indicador determinante: descontrol o uso de recursos (figura 15).

Figura 15. Comparación de cuatro mediciones del nivel de severidad en los usuarios asignados para ser atendidos con el modelo Sermás en la IPS Universitaria



	Noviembre 2017	Marzo 2018	Julio 2018	Noviembre 2018
Alto	30,6 %	19,1 %	16,8 %	16,3 %
Moderado	25,3 %	10,7 %	22,7 %	13,9%
Bajo	44,1 %	70,2 %	60,5 %	69,8 %

Fuente: elaboración propia.

Análisis y discusión

El análisis de una población basada en información estadística, epidemiológica, demográfica, etc., para una IPS, es una metodología que da respuesta a lo presentado, ya que respalda las decisiones basadas en evidencia sobre la salud de la población y permite la formulación de estrategias que considere tanto la atención como la prevención.

En la identificación y clasificación de los usuarios según el riesgo, la caracterización de la población mediante la aplicación de herramientas que contienen los lineamientos conceptuales y metodológicos permitió catalogar las características, las necesidades y los problemas de la población asignada, y ello sirvió como punto de partida para la gestión del riesgo en salud y para conocer el estado actual de la salud de la población objeto. Es importante tener en cuenta que el proceso de caracterización poblacional debe respetar una estructura ordenada y lógica, donde la información puede provenir de numerosas fuentes, y son la accesibilidad a dichas fuentes y la calidad de la información limitantes para poder desarrollar una caracterización poblacional completa, como lo estipula la guía, y tener en cuenta variables importantes, como los determinantes en salud, los cuales permitirían en cierta medida predecir el riesgo de sufrir una enfermedad y anticiparse a la gestión en salud de esa población.

La estratificación permite estimar sus niveles de severidad y conocer sus patrones de descontrol y uso de recursos. En ese sentido, la estratificación de segmentos por grupo de riesgo ha posibilitado priorizar la intervención individual, grupal, oportuna, pertinente y eficiente de los usuarios clasificados con descontrol alto y moderado, de modo que se ha mejorado el riesgo de eventos evitables y negativos. Así mismo, permite el análisis diferencial y asertivo de las causas, las soluciones y la posterior priorización en grupos de riesgo para mitigar el riesgo del alto consumo de recursos.

Es común hallar textos donde se refieren a los conceptos de segmentación y estratificación de la población como un mismo procedimiento o como sinónimos. Pero otros los desarrollan como dos componentes diferentes dentro de la actividad de identificación y clasificación de pacientes. Tal es el caso del Gobierno de Navarra (2013), en el que la segmentación se realiza con base en la presencia de varias enfermedades crónicas y la estratificación consiste en clasificar por niveles de riesgo o severidad los segmentos identificados según el descontrol de la enfermedad y el uso de servicios de salud.

En la actualidad, con el aumento de la demanda de los servicios de salud y los altos costos que estos requieren, se debe analizar la población por medio de la identificación, la clasificación, la segmentación y la estratificación de sus riesgos. Ello permite dirigirse de manera más eficiente y efectiva a los estratos identificados por medio de estrategias adaptadas a cada uno de estos.

En los diferentes componentes del enfoque metodológico propuesto (caracterización, segmentación y estratificación) se usa información de diferentes fuentes, que se transforma aplicando técnicas de análisis cualitativas y cuantitativas de variados niveles de complejidad. Así, en la medida en que esto se promueva, dará lugar a la retroalimentación y evidenciará la necesidad de posibles ajustes para el mejoramiento y fortalecimiento de cada una de las partes del proceso.

Desde este trabajo, se vislumbran grandes adelantos y cómo en la actualidad compañías y grupos de investigación alrededor del mundo han identificado el potencial de la información en salud para proveer diferentes tipos de análisis que permitan mejorar la gestión de los sistemas de salud, al hallarse modelos para predecir la aparición de una enfermedad, a través de la búsqueda de biomarcadores y el desarrollo de herramientas de soporte al diagnóstico médico basadas en el *machine learning* (Ahmad et al., 2015). Ellos han impactado diferentes especialidades de la práctica médica y siguen siendo materia de investigación. Estos modelos se han utilizado, por ejemplo, para la detección temprana y caracterización de la enfermedad de Alzheimer en escenarios clínicos utilizando conceptos de bioperfil y *clustering* (Escudero et al., 2011) o el uso de redes neuronales artificiales para la identificación de enfermedades pulmonares (Er et al., 2010). Con lo anterior se ha establecido cómo la atención de los pacientes, a través de conceptos como medicina personalizada de precisión, usa la información de los genes, las proteínas

y el ambiente de una persona para prevenir, diagnosticar y tratar una enfermedad, brindado un tratamiento individualizado, dirigido a las necesidades del paciente en función de sus propias características genéticas, biomarcadores, fenotípicas o psicosociales (Jameson y Longo, 2015).

Pero los avances en modelación no solo se han centrado en identificar los riesgos y las necesidades de los pacientes; también han apoyado la gestión de los recursos del sistema, donde a través del modelamiento predictivo se ha logrado predecir el número de días que un paciente estará hospitalizado o la probabilidad de readmisión, e incluso en algunos casos buscar similitudes entre pacientes para seleccionar el tratamiento de mayor eficacia de acuerdo con el historial clínico. Igualmente, se utilizan para la detección temprana del inicio de la enfermedad e individualizar la atención de los pacientes (Ng et al., 2015). Los modelos predictivos para identificación, segmentación, estratificación de riesgo están compuestos por el algoritmo matemático que calcula el riesgo para cada paciente. Por lo general, el algoritmo se basa en modelos de regresión múltiple, aunque a veces se utilizan redes neurales o árboles de decisión (Cárdenas et al., 2018).

De este trabajo es importante resaltar que, independientemente de la metodología que se use en la práctica, los conjuntos de datos de salud vinculados deben estar disponibles. Todos los estudios dependen en gran medida de información detallada del paciente, a menudo obtenidos de diferentes fuentes y conjuntos de datos vinculados a él. Esto pone de relieve una restricción importante, ya que la disponibilidad de conjuntos de datos de atención médica vinculados sigue siendo limitada en muchos países, incluyendo el nuestro, en el cual se ha evidenciado una baja calidad en los RIPS y ausencia de políticas que la fomenten y propicien esquemas comparables de gestión del riesgo de salud (Cortés et al., 2016).

Dado lo anterior, como apoyo al SGSSS en Colombia, Minsalud busca optimizar el uso y el aprovechamiento de las fuentes de información que son obligatorias por normatividad legal, dedicando sus esfuerzos a superar las deficiencias que afecten su cobertura, oportunidad o calidad, pero no sustituyéndolos o reemplazándolos por sistemas paralelos no obligatorios (Dirección de Epidemiología y Demografía, 2013). Por esta razón, el proceso de caracterización poblacional propende a emplear la información rutinaria captada por norma legal y aquella que suministran los estudios y encuestas poblacionales. Para lograr esto, la Oficina de Tecnología de la Información y Comunicación y la Dirección de Epidemiología y Demografía han venido desarrollando el proceso de integración de registros y fuentes externas de datos con información y estadísticas de salud al Sistema de Gestión de Datos del Sistema de Información para la Protección Social (Sispro). Estos esfuerzos son un primer paso que promueve retos para los diferentes actores del sistema y entes gubernamentales frente a temas de estandarización, integración de fuentes y disponibilidad de información con calidad.

Como oportunidad de mejora y como estudios futuros, se plantea la necesidad de definir los criterios y utilidades de la aportación que realizan los modelos predictivos en la estratificación y segmentación de pacientes crónicos con diferentes niveles de necesidades; y a la vez definir qué eventos necesitamos predecir. Así mismo, se deberá contar con herramientas avanzadas de *business intelligence*, minería de datos y *big data*, que no solo nos permitirán evaluar los resultados, sino que nos ayudarán a predecir riesgos, segmentar y estratificar la población e identificar áreas de mejora.

Conclusiones

Este estudio demostró cómo el análisis de la población basado en información de una entidad prestadora de servicios de salud se utiliza de forma amplia y puede proporcionar una descripción de las necesidades de atención de la población objetivo para apoyar las políticas de intervención de dicha población. Adicionalmente, se evidenció que para el análisis de resultados de este tipo de estudio es importante contar con criterio estadístico y experiencia clínica.

Los resultados evidenciados indican que la metodología aporta elementos necesarios para identificar, comparar y analizar la dinámica del riesgo en salud de otras cohortes poblacionales homogéneas asignadas para atender bajo el modelo Sermás de la IPS Universitaria, en los componentes de la atención, los resultados en salud y el uso eficiente de recursos.

Por medio de la gestión de riesgo en salud se logra identificar pacientes que puedan beneficiarse de intervenciones para prevenir eventos no deseados y permite una mejor planificación y gestión eficiente de los recursos. Además, los profesionales en salud pueden utilizar la información de estos resultados para adoptar protocolos específicos para la atención de los pacientes.

Debemos avanzar hacia el análisis inteligente de la información y la construcción de modelos de segmentación, estratificación y predicción de comportamiento, y modelos que nos permitan proyectar el uso de los recursos que podría exigir una población asignada. Con base en esto, debemos estimar el costo de la gestión de la salud y realizar un análisis de riesgo financiero.

Referencias

Abizanda Soler, P., Paterna Mellinas, G., Martínez Sánchez, E. y López Jiménez, E. (2010). Evaluación de la comorbilidad en la población anciana: utilidad y validez de los instrumentos de medida. *Revista Española de Geriatría y Gerontología*, *45*(4), 219-228.

Agency for Clinical Innovation. (2014). *Risk stratification: A discussion paper for NSW health's approach to risk stratification.* https://www.aci.health.nsw.gov.

au/__data/assets/pdf_file/0015/253005/Risk_Stratification_discussion_paper.PDF

Ahmad, P., Qamar, S. y Rizvi, S. Q. (2015). Techniques of data mining in healthcare: A review. *International Journal of Computer Applications*, *120*(15), 38-50. https://doi.org/10.5120/21307-4126

Arango, V. E. y Ruiz, I. C. (2011). *Diagnóstico de los adultos mayores de Colombia*. Secretaría Distrital de Planeación. http://www.sdp.gov.co/sites/default/files/diag_adul_mayor.pdf

Bernal-Acevedo, O. y Forero-Camacho, J. C. (2011). Information systems in health sector in Colombia. *Revista Gerencia y Politicas de Salud*, *10*(21), 85-100.

Bernal Pinilla, J. (2015). *Carga de la enfermedad pulmonar obstructiva crónica en Colombia* [tesis de maestría]. Pontificia Universidad Javeriana. https://repository.javeriana.edu.co/bitstream/handle/10554/19457/BernalPinillaJimena2015.pdf?sequence=1&isAllowed=y

Bonal Ruiz, R., López Vásquez, N., Vargas, P., Meoño Martín, T. y Brañas Coelho, R. W. (2017). Apoyo al automanejo de condiciones crónicas: un desafio de los sistemas de salud de América Latina. *Revista Finlay*, *7*(4), 268-277. http://scielo.sld.cu/pdf/rf/v7n4/rf06407.pdf

Bunch-Barrera, A., Tamer-David, L. M., Ardila-Celis, F., Laganis-Valcarcel, S., Castaño-Rodríguez, R., Vesga-Gualdrón, J. I., López Jaramillo, P., Muñoz, F., Sánchez-Pedraza, R. y Sanabria-Arenas, R. M. (2016). Impacto de un modelo de gestión de enfermedad en una población con tratamiento de diálisis en Colombia. *Revista de La Facultad de Medicina*, *64*(4), 695-700. http://dx.doi.org/10.15446/revfacmed.v64n4.54556

Cárdenas, C., González, S., Nahuel, R., Herrera, P., Ferrada, L. y Celis, D. (2018). Diseño de un modelo predictivo de pesquisa cardiovascular utilizando árboles de decisión: propensión de pacientes a presentar diabetes tipo 2, hipertensión arterial o dislipidemia. Estudio piloto, comuna de Quellón, Chiloé. *Revista Chilena de Cardiología*, *37*(2), 126-133.

Castaño Yepes, R. A. (2014). *Mecanismos de pago en salud: anatomía, fisiología y fisiopatología*. Ecoe.

Cortés, A., Gorbanev, I., Agudelo-Londoño, S. y Yepes, F. J. (2016). Grupos relacionados por el diagnóstico: experiencia en tres hospitales de alta complejidad en Colombia. *Universitas Médica*, *57*(2), 171-181. https://doi.org/10.11144/Javeriana.umed57-2.grde

Cotta, R. M. M., Suárez-Varela, M., Llopis González, A., Cotta Filho, J. S., Real, E. R. y Ricós, J. A. (2001). La hospitalización domiciliaria: antecedentes, situación actual y perspectivas. *Revista Panamericana de Salud Pública*, *10*, 45-55. http://www.scielo.edu.uy/pdf/rpsp/v10n1/5850.pdf

Cuenta de Alto Costo (CAC). (2017). *Situación de la enfermedad renal crónica, la hipertensión arterial y la diabetes mellitus en Colombia 2017*. Fondo Colombiano

de Enfermedades de Alto Costo. https://cuentadealtocosto.org/site/publicaciones/situacion-de-la-enfermedad-renal-cronica-la-hipertension-arterial-y-la-diabetes-mellitus-en-colombia-2017/

Decreto 2699. (2007, 13 de julio). Por la cual se establecen algunas normas relacionadas con el Sistema General de Seguridad Social en Salud y se dictan otras disposiciones. Ministerio de Protección Social. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/Decreto-2699-de-2007.pdf

Departamento de Caldas, Observatorio de Salud. (s. f.). *Eventos trazadores*.

Dirección de Epidemiología y Demografía. (2013). *Guía metodológica para registros, observatorios, sistemas de seguimiento y salas situacionales nacionales en salud.* Ministerio de Salud y Protección Social.

Dirección de Epidemiología y Demografía. (2017). *Guía conceptual y metodológica para la caracterización de la población afiliada a las entidades administradoras de planes de beneficios de salud (EAPB)*. Ministerio de Salud y Protección Social. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/GCFI/Guia_caracterizacion_EAPB%20version%2021072014.pdf

Dirección de Epidemiología y Demografía. (2018). *Análisis de situación de salud (ASIS) Colombia, 2017*. Ministerio de Salud y Protección Social. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/PSP/asis-nacional-2017.pdf

Ellis, R. P., Pope, G. C., Iezzoni, L. I., Ayanian, J. Z., Bates, D. W., Burstin, H. y Ash, A. S. (1996). Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financing Review*, *17*(3), 101.

Er, O., Yumusak, N. y Temurtas, F. (2010). Chest diseases diagnosis using artificial neural networks. *Expert Systems with Applications*, *37*(12), 7648-7655. https://doi.org/10.1016/j.eswa.2010.04.078

Escudero, J., Zajicek, J. P. y Ifeachor, E. (2011). Early detection and characterization of Alzheimer's disease in clinical scenarios using Bioprofile concepts and K-means. En *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 6470-6473). https://doi.org/10.1109/IEMBS.2011.6091597

Fajardo Gutiérrez, A. (2017). Medición en epidemiología: prevalencia, incidencia, riesgo, medidas de impacto. *Revista Alergia México, 64*(1), 109-120. http://dx.doi.org/10.29262/ram.v64i1.252

Fedesarrollo y Fundación Saldarriaga Concha. (2015). *Misión Colombia Envejece: cifras, retos y recomendaciones.* https://www.fedesarrollo.org.co/sites/default/files/envejece.pdf

Gallardo Solarte, K., Benavides Acosta, F. P. y Rosales Jiménez, R. (2016). Costos de la enfermedad crónica no transmisible: la realidad colombiana. *Revista Ciencias de la Salud*, *14*(1), 103-114. https://doi.org/10.12804/revsalud14.01.2016.09

Gobierno de Navarra-Departamento de Salud. (2013). *Estrategia Navarra de atención integrada a pacientes crónicos y pluripatológicos*. http://www.navarra.es/home_es/Temas/Portal+de+la+Salud/Profesionales/Documentacion+y+publicaciones/Planes+y+programas/Estrategia+navarra+de+atencion+integrada+a+pacientes+cronicos+y+pluripatologicos.htm

González Fuentes, A. A. (2017). *Comparación de sistemas de pago convenido en los costos de la prestación de servicios ambulatorios y hospitalarios* [tesis de maestría]. Universidad del Rosario. http://repository.urosario.edu.co/handle/10336/13963

IPS Universitaria. (2017). *Modelo de Atención Sermás*. https://ipsuniversitaria.com.co/ser-mas/

Jameson, J. L. y Longo, D. L. (2015). Precision medicine—personalized, problematic, and promising. *Obstetrical & Gynecological Survey*, *70*(10), 612-614.

Londoño Agudelo, E. (2017). Las enfermedades crónicas y la ineludible transformación de los sistemas de salud en América Latina. *Revista Cubana de Salud Pública*, *43*(1), 68-74.

López, C. P. (2013). *Técnicas de segmentación: conceptos, herramientas y aplicaciones*. Alfaomega.

Ministerio de Salud y Protección Social (Minsalud). (2011). *Guía de instrucción para el diseño de programas de gestión de enfermedades crónicas*. https://www.minsalud.gov.co/sites/rid/1/Guía diseño de programas de gestión de enfermedades cronicas.pdf

Ministerio de Salud y Protección Social (Minsalud). (2013a). *Envejecimiento demográfico: Colombia 1951-2020 dinámica demográfica y estructuras poblacionales*. Imprenta Nacional de Colombia. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/PS/Envejecimiento-demografico-Colombia-1951-2020.pdf

Ministerio de Salud y Protección Social (Minsalud). (2013b). *Lista de tabulación de la morbilidad*. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/INEC/INTOR/10._Lista_Morbilidad.pdf

Ministerio de Salud y Protección Social (Minsalud). (2014). *Guía conceptual y metodológica para la construcción del ASIS de las entidades territoriales*. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/PSP/Guia%20ASIS%2028112013.pdf

Ministerio de Salud y Protección Social (Minsalud). (2016). *Guía metodológica: Sistema de Evaluación del Desempeño del Sistema de Salud de Colombia bajo Criterios de la Organización para la Cooperación y el Desarrollo Económico - OCDE*. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VS/ED/GCFI/guia-ross-ocde.pdf

Ministerio de Salud y Protección Social (Minsalud). (2018). *Gestión integral del riesgo en salud: perspectiva desde el aseguramiento en el contexto de la Política de*

Atención Integral en Salud. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VP/DOA/girs-prespectiva-desde-aseguramiento.pdf

Morales-Asencio, J. M. (2014). Gestión de casos y cronicidad compleja: conceptos, modelos, evidencias e incertidumbres. *Enfermería Clínica*, *24*(1), 23-34.

Ng, K., Sun, J., Hu, J. y Wang, F. (2015). Personalized predictive modeling and risk factor identification using patient similarity. *AMIA Summits on Translational Science Proceedings*, *2015*, 132-136.

Organización Mundial de la Salud. (2015). *Informe mundial sobre el envejecimiento y la salud*. https://apps.who.int/iris/bitstream/handle/10665/186466/9789240694873_spa.pdf?sequence=1

Organización Mundial de la Salud. (2017). *10 datos sobre la situación sanitaria mundial*. https://www.who.int/features/factfiles/global_burden/es/

Organización Panamericana de la Salud. (2012). *Envejecimiento saludable y enfermedades no transmisibles*. https://www.paho.org/hq/dmdocuments/2012/Fact-Sheet-Seniors-NCDs-Spanish.pdf

Orueta Mendia, J. F., García-Álvarez, A., Alonso-Morán, E. y Nuño-Solinis, R. (2014). Desarrollo de un modelo de predicción de riesgo de hospitalizaciones no programadas en el País Vasco. *Revista Española de Salud Pública*, *88*, 251-260.

Osakidetza, Centro de Investigación en Cronicidad, Gobierno Vasco. (2015). *Estratificación del riesgo: una herramienta para responder mejor a las necesidades de salud de las personas y de la población*. https://ec.europa.eu/eip/ageing/sites/eipaha/files/results_attachments/estratificacion_del_riesgo.pdf

Pérez Giménez, G. (2017). *Impacto de un programa de gestión de casos a pacientes crónicos complejos* [tesis de doctorado]. Universitat de Barcelona. https://www.tdx.cat/bitstream/handle/10803/457349/GPG_TESIS.pdf?sequence=1&isAllowed=y

Resolución 0412 (2000, 25 de febrero). Por la cual se establecen las actividades, procedimientos e intervenciones de demanda inducida y obligatorio cumplimiento y se adoptan las normas técnicas y guías de atención para el desarrollo de las acciones de protección específica y detección temprana y la atención de enfermedades de interés en salud pública. Ministerio de Salud y Protección Social. https://docs.supersalud.gov.co/PortalWeb/Juridica/OtraNormativa/R0412000.pdf

Resolución 1536 (2015, 11 de mayo). Por la cual se establecen disposiciones sobre el proceso de planeación integral para la salud. Ministerio de Salud y Protección Social. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/resolucion-1536-de-2015.pdf

Resolución 1841 (2013, 28 de mayo). Por la cual se adopta el Plan Decenal de Salud Pública 2012-2021. Ministerio de Salud y Protección Social. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/resolucion-1841-de-2013.pdf

Resolución 2626 (2019, 27 de septiembre). Por la cual se modifica la Política de Atención Integral en Salud -PAIS- y se adopta el Modelo de Acción Integral Territorial-MAITE. Ministerio de Salud y Protección Social. https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/DE/DIJ/resolucion-2626-de-2019.pdf

Sicras-Mainar, A., Serrat-Tarrés, J., Navarro-Artieda, R. y Llopart-López, J. R. (2006). Posibilidades de los Grupos Clínicos Ajustados (Ajusted Clinical Groups-ACG's) en el ajuste de riesgos de pago capitativo. *Revista Española de Salud Pública*, *80*(1), 55-65.

Superintendencia Nacional de Salud (Supersalud). (2016). *Circular externa 009, por la cual se imparten instrucciones relativas al sistema de administración del riesgo de lavado de activos y la financiación del terrorismo (SARLAFT)*. https://www.uiaf.gov.co/caracterizacion_usuarios/perfiles/reportantes/superintendencia_nacional_salud/28400

Vuik, S. I., Mayer, E. y Darzi, A. (2016). A quantitative evidence base for population health: Applying utilization-based cluster analysis to segment a patient population. *Population Health Metrics*, *14*(44), 1-9.

Yan, S., Kwan, Y., Tan, C., Thumboo, J. y Low, L. (2018). A systematic review of the clinical application of data-driven population segmentation analysis. *BMC Medical Research Methodology*, *18*(121), 1-12.

Colaboradores

Revisores

Andrés Felipe Gutiérrez
Facultad de Administración, Universidad de los Andes, Bogotá, Colombia

Carlos Arturo Manrique
Escuela Internacional de Ciencias Económicas y Administrativas, Universidad de La Sabana, Chía, Colombia

Carlos Franco
Escuela de Administración, Universidad del Rosario, Bogotá, Colombia

Carolina Castañeda
Departamento de Economía y Empresa, Universidad Pompeu Fabra, Barcelona, España

Clara Inés Pardo Martínez
Escuela de Administración, Universidad del Rosario, Bogotá, Colombia

Consuelo Peña Aponte
Secretaría Distrital de Salud, Bogotá, Colombia

Eliana María González-Neira
Departamento de Ingeniería Industrial, Pontificia Universidad Javeriana, Bogotá, Colombia

Elyn L. Solano-Charris
Escuela Internacional de Ciencias Económicas y Administrativas, Universidad de La Sabana, Chía, Colombia

Fabián Andrés Castaño
Departamento de Ingeniería Civil e Industrial, Pontificia Universidad Javeriana, Cali, Colombia

Helena María Cancelado Carretero
Departamento de Ingeniería Industrial, Universidad Icesi, Cali, Colombia

José Fernando Jiménez
Departamento de Ingeniería Industrial, Pontificia Universidad Javeriana, Bogotá, Colombia

Juan Carlos Rivera Agudelo
Departamento de Ciencias Matemáticas, Universidad Eafit, Medellín, Colombia

Juan Guillermo Villegas
Departamento de Ingeniería Industrial, Universidad de Antioquia, Medellín, Colombia

Julio César Castellanos
Hospital Universitario San Ignacio, Bogotá, Colombia

Paula Alejandra Escudero Marín
Departamento de Ciencias Matemáticas, Universidad Eafit, Medellín, Colombia

Paula Andrea Velázquez
Desarrollo Organizacional, Institución Prestadora de Servicios de Salud IPS Universitaria

Rolando José Acosta Amado
Departamento de Ingeniería Industrial, Universidad Icesi, Cali, Colombia

Stevenson Bolívar Atuesta
Departamento de Ingeniería Industrial, Pontificia Universidad Javeriana, Bogotá, Colombia

Yaneth Cecilia Girón
Secretaría Distrital de Salud, Bogotá, Colombia

Yoni Fernando Ceballos
Departamento de Ingeniería Industrial, Universidad de Antioquia, Medellín, Colombia

Comité editorial

William J. Guerrero
Facultad de Ingeniería, Universidad de La Sabana, Chía, Colombia

Juan Sebastián Jaén-Posada
Departamento de Ingeniería Industrial, Universidad de Antioquia, Medellín, Colombia

Daniel R. Suárez
Departamento de Ingeniería Industrial, Pontificia Universidad Javeriana, Bogotá, Colombia