



**AJUSTE POR RIESGO PARA EL CÁLCULO DE LA UPC EN COLOMBIA:
AJUSTE DESDE LAS VARIABLES DE ESTADO DE SALUD PARA LAS
ASEGURADORAS COLOMBIANAS**

Jeisson Fernando Lancheros Jiménez

**Universidad Santo Tomás
Maestría en Ciencias Económicas
Bogotá D.C., Colombia
2019**

**AJUSTE POR RIESGO PARA EL CÁLCULO DE LA UPC EN COLOMBIA:
AJUSTE DESDE LAS VARIABLES DE ESTADO DE SALUD PARA LAS
ASEGURADORAS COLOMBIANAS**

Jeisson Fernando Lancheros Jiménez

Tesis presentada como requisito para optar por el título de:
Magíster en ciencias económicas

Director: Martín Romero PhD

**Universidad Santo Tomás
Maestría en Ciencias Económicas
Bogotá D.C., Colombia
2019**

Dedicatoria

A mi familia.

Resumen

El ajuste por riesgo es un elemento fundamental en el diseño de políticas públicas en salud. Actualmente, el Ministerio de Salud y Protección Social de Colombia estima los valores de la unidad de pago por capitación a través de un modelo de ajuste por riesgo, que únicamente tiene en cuenta variables como la ubicación geográfica, edad y sexo. El objetivo de este documento es demostrar que un modelo de ajuste por riesgo que tenga en cuenta las variables de estado de salud puede asignar de manera más eficiente los recursos destinados a las aseguradoras del régimen contributivo a través de la unidad de pago por capitación. Por medio de una revisión de la actual metodología de cálculo y de la comparación con un modelo propuesto con grupos de riesgo que indican el estado de salud de un individuo, haciendo uso de diversos indicadores de precisión, predicción y análisis de sensibilidad sobre una muestra de una base datos de suficiencia de la unidad de pago por capitación, este documento estima, en el contexto de una aseguradora del régimen contributivo, los valores que mejoran la asignación de recursos para cada grupo de riesgo.

Los valores estimados por el modelo propuesto tienen un mejor desempeño en los indicadores seleccionados, y son estables ante variaciones, como lo demuestran los análisis de sensibilidad. Igualmente, garantizan que el gasto sea menor al ingreso que se recibiría con la implementación del modelo de ajuste por riesgo propuesto.

Palabras clave: ajuste por riesgo, unidad de pago por capitación, grupos de riesgo.

Clasificación JEL: análisis de mercados de la salud (I11); aseguramiento en salud pública y privada (I13); política gubernamental, regulación, salud pública (I18); evaluación de modelos, validación y selección (C52); métodos de pronóstico y predicción, métodos de simulación (C53).

Tabla de contenido

INTRODUCCIÓN	7
CAPÍTULO 1. FUNDAMENTACION TEÓRICA.....	9
1.1 Marco teórico.....	9
1.2 Cálculo de la UPC	11
1.2.1 Fundamentos teóricos	12
1.2.1.1 Método de prima pura:.....	14
1.2.1.2 Método de razón de pérdida:	15
1.3 Teoría de la prima y del seguro	19
CAPÍTULO 2. DETERMINANTES DEL DESEMPEÑO DE LA UPC	21
2.1 Alternativas al cálculo actual	24
CAPÍTULO 3. MODELO DE AJUSTE POR RIESGO PROPUESTO	29
CAPÍTULO 4. RESULTADOS DEL MODELO.....	44
CONCLUSIONES.....	54
ANEXOS	57
Anexo A. Regresión modelo A: costo sin transformar	57
Anexo B. Regresión modelo B: costo logarítmico.....	62
Anexo C. Regresión modelo C: costo logarítmico ponderado	67

Índice de tablas

Tabla 1. Grupos etarios	17
Tabla 2. Grupos de riesgo: pacientes no crónicos	35
Tabla 3. Grupos de riesgo: pacientes crónicos	35
Tabla 4. Grupos de riesgo: pacientes multicrónicos.....	36
Tabla 5. Indicadores modelo A.....	45
Tabla 6. Indicadores modelo B.....	46
Tabla 7. Indicadores modelo C.....	46
Tabla 8. Indicadores modelo C: regresión por cuantiles	47
Tabla 9. Valores estimados para cada grupo de riesgo	48
Tabla 10. Distribución por UPC bajo modelo actual	50
Tabla 11. Distribución por UPC con el modelo propuesto: NAIVE	51
Tabla 12. Distribución por UPC con el modelo propuesto: SMEARING	51

INTRODUCCIÓN

Con el objetivo de consolidar un sistema de protección social que sea más eficiente y que pueda proveer de servicios y asistencia a la población, Colombia ha avanzado en temas de seguridad social. El sistema de Seguridad Social en Salud, creado a partir de la ley 100 de 1993, es el mecanismo estatal donde se canalizan las políticas públicas relacionadas con la salud de la población en general, garantizando que los recursos recaudados, destinados para tal fin, suplan las necesidades, siempre crecientes, de la población. Sin embargo, dentro de este sistema hay una distinción clara: la prestación de los servicios de salud está separada de la gestión de los recursos y el aseguramiento en salud. Estos recursos, y su correcta asignación a toda la población, son el objeto de este documento.

Actualmente, el sistema de salud está compuesto por tres agentes: el Estado, las aseguradoras y los prestadores. Los roles de financiación, regulación y sanción están a cargo del Estado, mientras que la prestación se garantiza a través de las aseguradoras, de carácter privado, y se realiza en las instalaciones pertenecientes a los prestadores, que constituyen el principal contacto de la población con el sistema de salud. Las actividades de los prestadores son financiadas a través de los recursos que las aseguradoras pagan por concepto de atención de determinada población, a través de diferentes mecanismos de contratación y modelos de atención, y son estas aseguradoras las que administran los recursos que el Estado destina para garantizar la atención de la población del país, a través de la unidad de pago por capitación (UPC).

Dadas las características heterogéneas de la población de cualquier país, la asignación de recursos a través de una prima, como lo es la UPC, es un elemento fundamental tanto en el financiamiento de las actividades de los prestadores como en la sostenibilidad del ejercicio de las aseguradoras. Lo que redundará en la calidad de la atención que reciben las personas que tienen contacto con el sistema de salud.

El cálculo de una prima es sensible a la metodología que se escoja para su estimación. Lo que hace que tanto las variables escogidas como los modelos de estimación, y sus indicadores de predicción y explicación, sean relevantes al momento de definir una política encaminada en la eficiencia en el uso de recursos públicos destinados a la atención en salud de la población. Teniendo en cuenta lo anteriormente expuesto, la pregunta central de este estudio es ¿ una metodología diferente de ajuste por riesgo puede mejorar la distribución de los recursos de la UPC para la población de las aseguradoras colombianas?. Como resultado de este razonamiento, el objetivo general de ese estudio es demostrar que el ajuste por riesgo que incluye variables de estado de salud puede mejorar la asignación de recursos en las aseguradoras colombianas.

Este documento hace un estudio de la metodología de cálculo de la UPC en Colombia a partir de documentos oficiales y de literatura especializada respecto al ajuste por riesgo que se realiza en la actualidad, y el que se podría realizar. Posteriormente, se propone una metodología de cálculo, usando una base de datos de una población representativa de un asegurador del régimen contributivo en Colombia. Finalmente, se presentan los resultados y se realizan recomendaciones de política, discusión de los resultados y conclusiones generales.

CAPÍTULO 1. FUNDAMENTACION TEÓRICA

1.1 Marco teórico

Uno de los primeros autores en desarrollar los conceptos detallados de la economía de la salud fue Keneth Arrow, con la publicación del trabajo *“Uncertainty and the Welfare economics of Medical Care”* (Arrow, 1963). En este trabajo, el autor presenta un análisis sobre los servicios de salud en el marco de los conceptos de la economía del bienestar, que también fueron planteados inicialmente por Arrow. El análisis del mercado de los servicios de salud busca *“contrastar la realidad de los mercados con la teoría sobre competencia perfecta y plantear opciones de política o mecanismos de no mercado en pro de una mejor asignación de recursos”* (Jairo & Kristian, 2016).

Uno de los objetivos de Arrow en el desarrollo de esta obra en particular, fue comprobar si el mercado es un mecanismo suficientemente eficiente para lograr la asignación de recursos, y dado el caso de que el mecanismo de mercado no garantizara condiciones de equilibrio general, comprobar si existe algún mecanismo de política pública, que permita asegurar que se maximice la utilidad de la sociedad, en este caso percibida como el bienestar que trae el acceso a los bienes y servicios que oferta el mercado de salud.

Estimar una prima que asegure un equilibrio competitivo para garantizar el acceso a un bien público, a pesar de las externalidades y asimetrías de información propias del sector salud, es un ejercicio en economía del bienestar, no solo por que se busca maximizar la utilidad percibida por los usuarios del sistema de salud, sino por que se busca lograr un equilibrio financiero que asegure la sostenibilidad del sistema a través del tiempo. Dicha prima es un mecanismo financiero y un mecanismo de bienestar, que se puede materializar a través del ajuste por riesgo.

El Ministerio de Salud y Protección Social (Minsalud) estima anualmente la prima para el financiamiento del plan obligatorio de beneficios en salud, con base en las recomendaciones que se presentan en el *Estudio de suficiencia y de los mecanismos de ajuste por riesgo que determinan la unidad de pago por capitación (UPC) para garantizar el plan obligatorio de salud*, realizado por la Comisión Asesora de Beneficios, Costos, Tarifas y Condiciones de Operación del Aseguramiento en Salud (Ministerio de Salud y Protección Social, 2016a), que está a cargo de «la asesoría en los temas referentes a la formulación de la política de beneficios, costos y tarifas en aseguramiento en salud» (Ministerio de Salud y Protección Social, 2012).

Este estudio analiza el régimen contributivo, sin tener en cuenta el régimen subsidiado, a pesar de que el plan obligatorio de salud (POS) ha sido unificado, con el objetivo de que la UPC sea suficiente para financiar el Plan de Beneficios en Salud (PB) con base en los mecanismos de ajuste por riesgo. El objetivo del cálculo de la prima es reflejar el gasto en salud de acuerdo con las variables demográficas seleccionadas, distribuyendo esta UPC en las entidades promotoras de salud (EPS) del país.

Sin embargo, las fuentes de financiación, los costos en salud, la operación del sistema y los costos administrativos se comportan de manera diferente para cada entidad aseguradora, de la misma forma que sucede entre el régimen contributivo y el régimen subsidiado. El cálculo del Minsalud utiliza los métodos de la Casualty Actuarial Society, donde se incorpora una revisión de la ecuación fundamental del seguro, ajustes por *Incurred but not reported* (IBNR) e *Incurred but not enough reported* (IBNER), *trending*, primas comerciales, ajuste por compensaciones causadas y no realizadas, y ajuste por riesgo. No obstante, los valores estimados, para algunos casos, no logran el objetivo de ser suficientes para la financiación del plan de beneficios debido a la diferencia en el comportamiento de cada EPS. Según Restrepo & López,

la financiación de un sistema de salud representa un proceso complejo y fundamental para cada gobierno, que debe cumplir tres funciones específicas e interrelacionadas: la primera es recaudar los fondos suficientes para la atención en salud; la segunda es mancomunar estos fondos para dispersar los riesgos financieros entre los afiliados y, por último, utilizar dichos fondos para la prestación de los servicios de salud. (Restrepo & López, 2012, pág 2)

Los procesos de financiación de estos sistemas se afectan por dos situaciones específicas: inicialmente, por los problemas económicos existentes en cada país y por una demanda creciente de mejores servicios y una lista cada vez mayor de nuevas tecnologías y medicamentos que demandan los usuarios (OMS, 2010).

El objetivo de este documento es demostrar que la prima asignada a las EPS (mediante la asignación por UPC) no ha sido suficiente, demostrando que la metodología de cálculo de la prima por UPC presenta una insuficiencia estructural, generando una ineficiencia en la asignación de los recursos destinados a los aseguradores del país.

Dado que uno de los objetivos del cálculo de la UPC es la asignación suficiente para cada persona dentro del sistema de salud, el ajuste por riesgo es una metodología que asegura que cada afiliado reciba recursos por UPC según su perfil de riesgo, de tal forma que aquellos que suponen un mayor costo, con base en información clínica e histórica, reciban un ingreso acorde a sus necesidades.

1.2 Cálculo de la UPC

Para el cálculo de la UPC se han tomado como base conceptos de la teoría actuarial, la teoría económica e indicaciones médicas, que a su vez fueron validados por la literatura internacional, como lo son la teoría de la prima, teoría del seguro (Perez, 2016), eficiencia y equidad en la seguridad social, teoría de los bienes públicos, las variables macroeconómicas y su impacto sobre el presupuesto

nacional (Flórez Pinilla, 2010), y aspectos médicos (Chicaiza Becerra & Cabedo Semper, 2007). Esto quiere decir que se han tomado parámetros técnicos adecuados que permitan soportar dicho cálculo. No obstante, el cálculo de esta prima, desde su creación, ha tenido ajustes de acuerdo con los acontecimientos que ha sufrido la economía y que han tenido efectos sobre el equilibrio financiero del Sistema General de Seguridad Social en Salud (SGSSS).

Es por esto que los gobiernos de turno, mediante el Minsalud, han buscado alternativas que permitan, mediante el cálculo de la UPC, generar mayor equidad y eficiencia dentro del sistema. Sin embargo, existen algunos aspectos que se deben considerar dentro de la forma en que se hace el cálculo, con el fin de que esta pueda ser acorde a las realidades y necesidades de cada asegurador.

1.2.1 Fundamentos teóricos

Para todos los periodos analizados se encuentra que el Minsalud realiza un estudio de corte transversal, tomando como principal fuente de datos los servicios registrados en la base de datos de Suficiencia de la UPC de dos años atrás (es decir, para el periodo de aplicación 2018 se analiza la del periodo de experiencia 2016, que es analizada en el periodo de 2017) (Ministerio de Salud y Protección Social, 2018) para proyectar la prima del siguiente año, observando los ingresos y gastos reportados por las EPS a nivel nacional.

Los estudios tienen en cuenta la información poblacional (edad, sexo, municipio, zonas de residencia, ingreso base de cotización (IBC) y tipo de afiliado, entre otras), a través de la base de datos de afiliaciones para cada régimen; información de servicios prestados a cargo del POS a partir de las bases de datos de registros de todas las prestaciones en salud para el régimen correspondiente, que fueron recibidas por los afiliados, registrando así los costos transaccionales para la EPS, el usuario y el sistema de salud; igualmente, se utilizan los ingresos y egresos de cada EPS como forma de validación de la información reportada. De esta manera

se construye la información para el cálculo de la prima con base en los costos causados el año anterior.

Esta metodología es común a todos los años analizados, teniendo en cuenta que para el año 2006 aún existía información incompleta, lo que no permitía una estimación correcta de los costos. «En el régimen contributivo las fuentes de información exploradas para establecer la suficiencia de la UPC presentan resultados muy diversos que impiden concluir si la UPC definida para el año 2004 es suficiente o insuficiente» (Ministerio de Salud y Protección Social, 2006).

Respecto a las variables utilizadas, para cada requerimiento se evidencia que la información disponible permite identificar tanto las características demográficas del afiliado como sus particularidades geográficas y de consumo de servicios (procedimientos, medicamentos, insumos), bien sean pertenecientes al plan obligatorio de servicios o excluidos de este. Esta información proviene de las bases de datos de prestaciones reportadas por cada aseguradora, el Fondo de Salud y Garantía (FOSYGA) (ahora Administradora de los Recursos del Sistema General de Seguridad Social en Salud, ADRES), para el régimen contributivo, y los estados financieros reportados a la Superintendencia Nacional de Salud (SNS), así como la población nacional que genera el Departamento Administrativo Nacional de Estadística (DANE) para el periodo de estudio correspondiente.

Una vez validada la información, se realiza la descripción de la población objetivo, teniendo en cuenta el bono demográfico¹, la morbilidad atendida durante el año, la descripción demográfica y la distribución por grupos de edad y sexo, así como la

¹ La Cepal define el Bono Demográfico como «un periodo en que la proporción de personas en edades potencialmente productivas crece de manera sostenida en relación con la de personas en edades potencialmente inactivas (...) se genera una situación particularmente favorable para el desarrollo, ya que aumentan las posibilidades de ahorro y de inversión en el crecimiento económico» (2008).

descripción reportada por los aseguradores y la distribución poblacional a nivel de departamentos de la población del año correspondiente al estudio.

Posterior a la descripción de la población, la metodología del Minsalud prosigue a realizar la evaluación de la suficiencia de la UPC. Según el Estudio de Suficiencia del 2015, «El objetivo central del análisis actuarial de tarifas, es fijar un nivel de prima que sea suficiente para cubrir las obligaciones en salud a cargo de las aseguradoras, y generar con ellas el nivel esperado de utilidad» (Ministerio de Salud y Protección Social, 2015).

Con base en las ecuaciones fundamentales de los seguros, la UPC se calcula a partir de los métodos de prima pura y razón de pérdida, a saber:

1.2.1.1 Método de prima pura:

$$P_I = L + E_L + (E_F + V \times P_I) + (Q_I \times P_I) \quad (1)$$

Despejando la prima suficiente:

$$P_I = \frac{L+E_L+E_F}{1-V-Q_I} \quad (2)$$

Donde²:

P_I = Prima indicada.

L = Siniestros.

² Para el régimen contributivo, de acuerdo con la Ley1438, del total de la prima, máximo el 10 % podrá destinarse a gastos administrativos y utilidad. Esto genera los siguientes parámetros de la ecuación básica del seguro para el régimen contributivo (Ministerio de Salud y Protección Social, 2013):

$$E_L = 0$$

$$E_F = 0$$

$$V + Q_I = 0.1$$

E_L = Gastos por ajustes de siniestros.

E_F = Gastos fijos.

V = Gastos variables como porcentaje de las primas.

Q_I = Nivel de *target* de utilidad en porcentaje de las primas.

1.2.1.2 Método de razón de pérdida:

$$P_I = P_c * (1 + Ind) \quad (3)$$

Despejando la solución óptima:

$$Ind = \frac{L+E_L+E_F}{P_c(1-V-Q_t)} \quad (4)$$

Donde:

P_I = Prima indicada.

P_c = Prima actual.

Ind = Indicación de tarifa.

De acuerdo con lo anterior, la metodología de cálculo de la UPC utiliza el método de razón de pérdida para encontrar el incremento necesario sobre la UPC actual, garantizando que la ecuación fundamental del seguro se encuentre balanceada. Esto permite que sea posible financiar la realidad futura de los afiliados y se garantiza la condición de balance global y balance a nivel de los grupos de riesgo escogidos (grupo etario, sexo y ubicación geográfica).

Finalmente, el cálculo de la prima pura de riesgo para el año correspondiente, para cada afiliado, «toma el costo observado al interior de cada grupo etario y geográfico, el cual es dividido por la población expuesta de ese grupo etario en el año de experiencia» (Ministerio de Salud y Protección Social, 2015) y se estima, para cada grupo de riesgo, de la siguiente forma (Ministerio de Salud y Protección Social, 2014):

$$\hat{\mu}_j = \left[\frac{\sum_{i=1}^{n_j} X_i}{Exp_{j(t)(1+\xi)}} \right] (1 + IBNR)(1 + \tau)(1 + f) \quad (5)$$

Donde:

$\hat{\mu}_j$ es la prima pura de riesgo j.

X_i es el costo de la atención en individuos del grupo de riesgo j.

IBNR Es el ajuste por siniestros incurridos, pero no reportados o no correctamente reservados.

τ es el *trending* por inflación.

f es el *trending* por frecuencia.

ξ es el factor de ajuste por compensaciones faltantes.

Esta fórmula es la que condensa el cálculo de la prima $\hat{\mu}_j$ para cada grupo de riesgo, seleccionada por el Minsalud, según los ponderadores de sexo, zona geográfica y grupo de edad.

De acuerdo con los lineamientos del Minsalud, el ajuste por riesgo que se realiza al incluir las variables de edad (grupo etario), ubicación geográfica y sexo, tiene como objetivo «identificar las tendencias de costos y corregir con base en ellas los incentivos de los aseguradores de buscar selección de afiliados y reducir la calidad de los servicios prestados» (Ministerio de Salud y Protección Social, 2016a). Es decir, el objetivo del ajuste por riesgo que se realiza actualmente es compensar consistentemente los costos agregados de una aseguradora al momento de atender a la población.

Además de los grupos de riesgo desagregado por sexo, el Minsalud (Ministerio de Salud y Protección Social, 2016a) incluye la categoría de ubicación geográfica (zona especial, zona alejada, zona normal y ciudades) y los siguientes grupos etarios:

Tabla 1. Grupos etarios

Grupo etario
<1 año
1 a 4 años
5 a 14 años
15 a 18 años hombres
15 a 18 años mujeres
19 a 44 años hombres
19 a 44 años mujeres
45 a 49 años
50 a 54 años
55 a 59 años
60 a 64 años
65 a 69 años
70 a 74 años
>75 años

Fuente: (Ministerio de Salud y Protección Social, 2016a).

Una vez definidos los componentes del cálculo de la UPC, es necesario explicar tanto la importancia de cada uno de estos como sus limitaciones:

- **IBNR:** este indicador incluye también el IBNER, y hace referencia a todos los procedimientos que las instituciones prestadoras de salud (IPS) llevan a cabo, pero que las EPS no tienen certeza que existen, ni de los procedimientos ni de los costos asociados a estos. Por ello, para que una atención médica se convierta efectivamente en una obligación cancelada, es necesario llevar a cabo una serie de etapas, las cuales pueden variar dependiendo de los procesos internos de cada EPS (Ministerio de Salud y Protección Social, 2008). Aquí radica una de las limitaciones, dado que cada EPS tiene diferentes mecanismos de recolección de información, los cuales varían entre sí generando, al calcular el promedio de IBNR, que se presenten desviaciones frente este, impulsando a aquellas que presentan desviaciones hacia arriba mejoren sus procesos de recolección de información y las que

presenten desviaciones hacia abajo les sea indiferente, ya que reciben más de lo que realmente tienen en cuentas por pagar.

Otro aspecto que se debe considerar es que, dado que se necesitan por lo menos tres periodos para calcular el valor del IBNR, estos valores deben ser deflactados para estimar el verdadero valor. Sin embargo, las metodologías actuales tienen en cuenta la inflación general causada y no una inflación del sector que permita medir el verdadero efecto del incremento de los precios de procedimientos, consultas y medicamentos. Con el objetivo de disminuir el valor del IBNR, el periodo de procesamiento de la información es en julio del año siguiente a la prestación de los servicios, y a la generación de la base de datos. Es importante aclarar que incluir la inflación general es un error metodológico, pues se debería incluir la inflación en salud para este cálculo. Sin embargo, el IPC en salud se refiere a los incrementos en los productos del sector salud que están incluidos en la canasta familiar y no necesariamente el incremento en los precios internos del mercado de la salud, que comprende elementos adicionales a los que están incluidos en la canasta familiar. Este error metodológico afecta al IBNR, pues el valor que se usa para deflactar no es el valor que comprende la inflación del mercado de la salud en general.

- **Inflación:** se refiere al incremento generalizado de los precios dentro de la economía, pues dado que la inflación se puede ver como un promedio de los incrementos de los diferentes sectores, este puede terminar siendo significativamente diferente de lo que en realidad se generó en el sector salud, lo cual afectaría a las EPS. Esto también ocurre cuando se deflacta el IBNR.
- **Frecuencia de uso:** se estima dividiendo el costo total de un K-esimo grupo, dividido sobre el número de expuestos.
- **Compensación:** se refiere al tiempo de duración de un afiliado dentro del periodo de un año. Es decir, si un paciente estuvo afiliado solo por 15 días, este tiempo debe ser compensado. Este está relacionado positivamente con el tiempo de exposición de los afiliados, el cual hace referencia al número de

años expuestos durante el periodo de experiencia desde el momento de su afiliación³.

1.3 Teoría de la prima y del seguro

El concepto de prima se refiere al pago que hace un individuo a modo de incentivo a otro individuo por la consecución u obtención de un bien o servicio. En el campo del aseguramiento una *prima de seguro* se refiere al pago efectuado por el asegurado como contraprestación por la cobertura de riesgo que un asegurador le da. Todo pago podrá llamarse prima, siempre y cuando cumpla con los siguientes principios:

Principio de utilidad esperada: consiste en que la utilidad generada por un contrato debe ser mayor o igual a lo que generaría si no se tomara dicho contrato. Lo anterior, teniendo en cuenta la aversión al riesgo que tienen los agentes aseguradores. Esto sería equivalente a que el asegurador fuera neutral o incluso averso al riesgo.

Principio de prima neta: consiste en que el valor de la prima cobrada debe ser mayor a las reclamaciones percibidas (costos). Es decir, los ingresos recibidos por el ejercicio de aseguramiento deben ser igual a los costos asumidos por dicho ejercicio.

Principio de la varianza: el principio de la varianza indica que la prima debe ser el valor esperado, incrementada en proporción a la varianza del riesgo.

Principio de desviación estándar: indica que la prima debe ser el valor esperado, incrementada en proporción a la desviación estándar del riesgo asumido (Ministerio de Salud y Protección Social, 2008).

Existen otros principios que debe cumplir un pago para ser considerado una prima, como el principio exponencial o el principio percentil, pero estos no se tendrán en

³ Ibid

cuenta en este análisis de suficiencia de la UPC, pues el mejor ajuste deseable, de acuerdo con los principios del cálculo, es aquel que minimiza los costos de aseguramiento, genera primas consistentes entre subgrupos y permite interpretaciones claras por parte de todos los actores (Ministerio de Salud y Protección Social, 2008). Por otro lado, el seguro, definido como un contrato por medio del cual un individuo cubre toda probabilidad de ocurrencia de un siniestro (exento de riesgo), el cual se salda mediante el pago de una prima, según lo indica la teoría del seguro. Los elementos reales del seguro son las personas físicas o jurídicas que lo integran, el asegurador, el tomador del seguro y el beneficiario de la póliza (Moreno, 2010).

CAPÍTULO 2. DETERMINANTES DEL DESEMPEÑO DE LA UPC

Para comprender los detalles relevantes del cálculo de la UPC, se realizó una revisión de todas las versiones del *Estudio de suficiencia y de los mecanismos de ajuste de riesgo para el cálculo de la unidad de pago de capitación para garantizar el plan de Beneficios en Salud* desde el año 2006 hasta el año 2016, publicadas por el Minsalud. Dada la naturaleza global del estudio, únicamente se analizó la metodología y los resultados para el régimen contributivo.

Todas las versiones del estudio coinciden en que se realiza una investigación multidisciplinaria que, por medio de un análisis a partir de las bases de datos disponibles usando metodología estadística, pretende determinar la suficiencia de la UPC con la que se debe financiar el POS (que a partir de 2012 es igual para ambos regímenes), determinando y evaluando los mecanismos de ajuste por riesgo que establecen el valor anual de la UPC para el régimen contributivo. De acuerdo con el estudio publicado en 2016,

los escenarios y análisis planteados buscan que la UPC sea suficiente para financiar el plan de beneficios y que los mecanismos de ajuste de riesgo estén reflejando el gasto en salud de acuerdo con las variables demográficas y geográficas seleccionadas. (Ministerio de Salud y Protección Social, 2016a)

El principio que justifica la utilización de una metodología que tenga en cuenta el ajuste por riesgo para el cálculo de la UPC, es que por medio de este se puede hacer una distribución más eficiente de los recursos. Es decir, de acuerdo con las necesidades de salud de cada afiliado. En la actual metodología del cálculo de la UPC se han incluido ajustes por riesgo en la asignación de recursos prospectiva. Según lo expuesto en el apartado anterior, esta metodología pretende reflejar el riesgo que asume cada EPS con la diferenciación de tarifas por los grupos de riesgo. Sin embargo, estos grupos ignoran variables de estados de salud que deberían ser incluidas en el cálculo de la prima.

A pesar de que uno de los mecanismos con el que se busca asegurar eficiencia por parte de las aseguradoras es por medio de este ajuste *ex ante* (se quiere que las EPS optimicen sus funciones de costo), aún existen ineficiencias generadas por un ajuste por riesgo incompleto. Alfonso, Riascos & Romero (Riascos, Alfonso, & Romero, 2012) afirman que el modelo utilizado por el Gobierno únicamente cuenta con menos de un 2 % de explicación de la variación del costo, por la construcción de los grupos de riesgo escogida. La literatura revisada demuestra que al realizar una mejor agrupación por riesgo, teniendo en cuenta las características epidemiológicas de la población, el poder predictivo de la metodología de cálculo de la UPC mejora, asignando con mayor eficiencia los recursos de las aseguradoras, lo que a su vez genera condiciones de estabilidad en los costos del sistema de salud (Riascos, Romero, & Serna, 2017).

Por su parte, en la sentencia T-760 de la Corte Constitucional, el organismo les ordena al Minsalud y a la Comisión de Regulación en Salud (CRES) adoptar medidas necesarias para superar las fallas de regulación en los planes de beneficios, y especifica que esta regulación también deberá incentivar que las EPS y las entidades territoriales garanticen a las personas el acceso a los servicios de salud, es decir, esta sentencia estuvo dirigida a garantizar el derecho a la salud de los colombianos y deberá, por esta razón, enfocarse en modificar todos aquellos elementos del diseño del SGSSS que se podrían traducirse en limitantes al derecho de la salud (CRES, 2010).

En el diseño de SGSSS, la UPC es uno de los principales instrumentos de regulación con que cuentan las entidades que dirigen el sistema. Por tal razón, se busca garantizar el derecho a la salud, por lo que es importante tomar acciones que mejoren la regulación por medio de una prima general, como lo es la UPC. La actual forma de cálculo tiene, entre otras cosas, el propósito de reducir los incentivos a la selección del riesgo por parte de las EPS. En estudios anteriores, como el de Riascos *et al.* (2012) (2013), se demostró que un buen ajuste de riesgo evita la

selección de riesgo y sus efectos. No contar con mecanismos adecuados de ajuste de riesgo de la UPC podría impedir la competencia por calidad y aumentaría la competencia por medio de reducción de costos, por lo que las aseguradoras tendrían incentivos para mantener grupos de población de bajo riesgo que generen menores costos.

Todo esto tiene como resultado que la selección de riesgo concentre pacientes de bajo riesgo en las aseguradoras que tengan la oportunidad de hacer esta selección y un aumento de los grupos de alto riesgos en otras EPS. Debido a que en la UPC no se reconocen estas diferencias, las finanzas de las aseguradoras pueden terminar siendo afectadas negativamente. De acuerdo con lo publicado por la CRES:

En la medida en que un ajuste de riesgo más profundo de la UPC le permitiría al asegurador soportar la carga financiera de una alta concentración de pacientes con un alto gasto en salud esperado (pacientes con enfermedades crónicas por ejemplo), se podría facilitar o promover la competencia por calidad, puesto que incluso en estos pacientes costosos existiría la posibilidad de tener cierto margen de rentabilidad, por ejemplo, mediante una buena atención (evitando complicaciones) y una buena gestión del riesgo en salud. Así las cosas, con el ajuste de riesgo se podría reducir uno de los elementos que podrían impedir el funcionamiento de la competencia por calidad, pues las entidades podrían llegar a estar interesadas en atraer más afiliados, incluso aquellos afiliados con un riesgo en salud elevado (si y solo si el ajuste de riesgo de la UPC incluye ese tipo de riesgo). (2010)

Teniendo en cuenta lo anterior, podría ser conveniente profundizar los mecanismos de ajustes de riesgo, transformarlos hacia un perfil, no solo demográfico, sino epidemiológico, que considere el estado de salud de los afiliados para calcular y reconocer los valores de la UPC a las EPS.

2.1 Alternativas al cálculo actual

Dado que la actual fórmula para estimar la UPC se basa en un cálculo estándar, es decir, por medio de una regresión lineal ponderada basada en tres agrupaciones de riesgo (edad, sexo y zona geográfica), se deja un porcentaje significativo de la varianza del costo anual de los servicios PB sin explicar. Disponer de un ajuste de riesgo adecuado permitiría que los ingresos de la entidad siempre reflejen el riesgo de la población afiliada⁴, lo cual justificaría cambiar la metodología de cálculo de los grupos de riesgo por una agrupación que incluya el detalle del estado de salud del afiliado. Además, este ajuste se basa en el hecho de que es posible incrementar el poder explicativo del modelo al agregar el objeto de aseguramiento, es decir, el estado de salud.

El trabajo de Alfonso, Riascos & Romero (2012) prueba que los modelos que utilizan grupos relacionados de diagnóstico, basados en la información reportada al Minsalud (discapacidad y hospitalización), generan una explicación de la varianza de los costos cercana al 14 %. Esto, en comparación con el valor del 1,27 % del modelo utilizado actualmente, es un argumento para aceptar la estimación de los grupos de riesgo que incluyen la información clínica como una alternativa más eficiente para el sistema de salud colombiano. De la misma forma, el modelo utilizado actualmente sobreestima el costo de los individuos con bajo costo, estimando un costo hasta 15 veces mayor (para el quintil más bajo del gasto), mientras que subestima el costo de los individuos que son más costosos, logrando predecir únicamente el 27 % del costo del gasto real para el quintil más alto del gasto. Como resultado, la asignación de recursos por este modelo hace que el *cream skinning*⁵ sea una práctica que resulte rentable para las EPS.

4 ibid

⁵ Se entiende como *cream skinning* la práctica de asegurar únicamente a los individuos que tienen un riesgo menor de siniestro, generando barreras de acceso a aquellos sobre los que tienen más probabilidades de representar mayores costos para la aseguradora. Esta práctica no implica negar el acceso al seguro, pues en caso colombiano está prohibido por ley. Pero sí fortalece prácticas

Un modelo que cumpla las condiciones deseadas, en términos de nivel de predicción del costo, debe asegurar que no existen incentivos para manipular la información que las aseguradoras reportan, además de ser fácilmente accesible para efectos de verificación o auditoría. Esto constituye uno de los retos de los modelos de ajuste por riesgo.

Por otro lado, los aspectos técnicos del cálculo de un modelo con un mejor ajuste por riesgo, de acuerdo con Schokkert & Van de Voorde (2004), se pueden trabajar en el marco de los modelos lineales, basados en la teoría de la elección social. Los estimadores lineales que se han encontrado en los estudios revisados cumplen parcialmente con las propiedades necesarias para explicar (Ministerio de Salud y Protección Social, 2013, 2015, 2016a), de una mejor manera, la varianza del costo en salud correspondiente al plan de beneficios. A pesar de que las variables que incluyen no son completamente inmunes a la manipulación (Riascos *et al.*, 2012), disminuyen considerablemente los riesgos de que las EPS realicen *cream skimming*, pues reduce los incentivos para crear barreras de acceso a los afiliados que presentan preexistencias.

El trabajo de Alfonso, Riascos & Romero (2012) modela el cálculo de la prima utilizando una base de datos de cerca de 24 millones de afiliados y variables de estado de salud del paciente (discapacidad, hospitalización, consulta al médico especialista, morbilidad y dos metodologías de grupos relacionados de diagnóstico). Además, compara los resultados con el modelo actual de cálculo de la UPC, midiendo el poder predictivo de cada modelo, principalmente utilizando el R^2 por medio de la varianza explicada por cada modelo.

Estos resultados demuestran que el actual modelo cuenta únicamente con un poder explicativo inferior al 2 %, mientras que los modelos que incluyen variables de

como demorar el trámite de pacientes sobre los que se conoce previamente el estado de salud, sin negar directamente el seguro.

estado de salud logran tener hasta un 15,41 % de poder explicativo, con resultados del 12,57 % y 14,90 % para los modelos que incluyen grupos relacionados con diagnóstico. Esto demuestra que los modelos que únicamente tienen en cuenta los factores demográficos son incapaces de explicar más del 3 % o 4 % del gasto en salud, es decir, la metodología de cálculo es estructuralmente insuficiente para asignar los recursos dentro de las EPS del país.

De la misma forma, los modelos permiten evaluar los incentivos de realizar *cream skimming* al ajustar la capitación a la luz de los diferentes modelos, analizando la capacidad de cada modelo de predecir los quintiles extremos frente al gasto real. Esto demuestra que la prima estimada para todos los modelos propuestos genera incentivos para la selección de riesgo. Sin embargo, los incentivos se reducen al utilizar modelos que incluyen grupos relacionados con diagnóstico que, según la literatura revisada, logran predecir hasta el 84 % del gasto de los pacientes más costosos (Riascos, Guerrero, Chernichovsky, 2013) (Riascos *et al.*, 2012).

La evidencia demuestra que, dada la falta de poder explicativo y predictivo del actual modelo con el que se calcula la UPC, es necesario desarrollar una nueva metodología para corregir los errores en la asignación de recursos que presenta el modelo actual. Para esto, los modelos que incluyen en su metodología de ajuste por riesgo variables del estado de salud presentan un mejor poder explicativo. Implícitamente, el Minsalud acepta que la asignación actual no es la más adecuada, al realizar un ajuste *ex post* asignando dinero a las aseguradoras a través de la cuenta de alto costo (Ministerio de Salud y Protección Social, 2016a) teniendo en cuenta enfermedades, principalmente para pacientes de alto costo. Este ajuste reconoce que, en casos particulares y de manera sistemática, es necesario realizar un ajuste financiero adicional para cubrir la totalidad de los gastos de grupos particulares de pacientes cuya asignación por UPC no es suficiente para las aseguradoras.

El documento de Schokkaert & Van de Voorde (2004) explica las razones por las que los estimadores insesgados, como los que se utilizan para el cálculo actual de la UPC, generan incentivos para que las aseguradoras no estén dispuestas a aceptar población sobre la que estiman que sus costos serán significativamente mayores a los valores que recibirán como compensación por aseguramiento poblacional. Ellis (2008) expone algunas soluciones para estos incentivos por medio de un ajuste por riesgo óptimo, comparando los métodos tradicionales de ajuste por riesgo, como es el caso de los estimadores de la UPC, con metodologías alternativas. Al tener en cuenta las características de cada afiliado de manera parcial, es posible incurrir en pérdidas para el asegurador o, en el mejor de los casos, tener una relación de aseguramiento que no genere utilidades para el asegurador.

En el modelo propuesto por Ellis, el objetivo del asegurador es maximizar el bienestar del asegurado, en lugar de solo tratar de evitar los costos por fuera de los rangos esperados o llegar a cubrir los gastos sin incurrir en pérdidas. En lugar de restringir a las EPS considerando un ajuste por riesgo con base en la fórmula tradicional, se propone pagar al asegurador en función de algunas características que difieren del costo esperado por el modelo, como los diagnósticos asociados a la atención.

Una solución a esta distorsión, por medio del ajuste por riesgo, consiste en pagar sistemáticamente un monto mayor por los afiliados sobre los que se presume un mayor costo futuro, y generar pagos menores de la población sobre la que se asume un menor costo de manera prospectiva. Por cada paciente que tenga un grupo de riesgo mayor, se le debe asignar al asegurador recursos a una razón mayor que el costo incremental de cada paciente asegurado (Ellis, 2008). Este pago adicional generará que el asegurador tenga más incentivos para gestionar de una mejor forma pacientes de alto riesgo y ofrecer servicios que mantengan sus costos controlados bajo rangos establecidos. Bajo el modelo actual no es posible incentivar a las EPS a eliminar las barreras de acceso a la población presumiblemente más costosa. Al

pagar valores significativamente mayores por la población más riesgosa se eliminaría el incentivo al *cream skimming*.

CAPÍTULO 3. MODELO DE AJUSTE POR RIESGO PROPUESTO

El objetivo del modelo es comprobar si la modificación del ajuste por riesgo en la metodología de cálculo de la UPC en el país puede generar una mejor explicación del gasto en salud de los servicios cubiertos por el plan de beneficios. La propuesta se centra en mejorar la forma en la que se realiza el ajuste *ex ante*, al enriquecer la información utilizada, teniendo en cuenta variables que tengan información de morbilidad y carga de enfermedad, utilizando, para la población del régimen contributivo, una metodología más sofisticada con la que se pronostique el gasto.

El algoritmo propuesto considera la construcción de grupos de riesgo adicionales basados en la carga de enfermedad, además de las variables que propone la metodología actual. El modelo debe demostrar que al usar el ajuste por morbilidad se mejora la predicción del gasto, tanto a nivel general como en los grupos más y menos costosos de la población. Trabajos como el de Riascos & Vargas (2014) sugieren que la diferencia entre el gasto predicho y el gasto observado es en promedio menor cuando se usan variables predictoras que incluyan el estado de salud. En general, el objetivo del ajuste por riesgo es asegurar que el gasto en salud que debe ser atribuido a la UPC no exceda las capacidades de pago, tanto de las aseguradoras, que cumplen la función de entidades que administran los recursos, como del total de recursos estatales.

Actualmente, el mecanismo de ajuste por riesgo consiste en una regresión lineal que contiene como variable regresada el gasto en salud, y como variables regresoras factores sociodemográficos como el sexo, grupos etarios, ubicación geográfica, y sus correspondientes interacciones. Los grupos etarios se definen de acuerdo con lo establecido por el Minsalud (2014), y la ubicación geográfica se define según el municipio de residencia, donde las áreas metropolitanas son el grupo urbano, los municipios cercanos a las áreas metropolitanas se clasifican como normales, y los municipios periféricos se clasifican como especiales.

Como indicadores de predicción, autores como Riascos, Romero & Serna (2017) sugieren el error cuadrático medio, el error medio absoluto, el R2 y la razón total del gasto predicho por el modelo frente al gasto total observado, para el total de la base de datos, como para los quintiles más altos y más bajos de costo. En el quintil más costoso se encuentran los individuos más costosos para el sistema.

El modelo de cálculo de la UPC que se desea estimar se basa en el trabajo de Riascos *et al.* (2017), quienes describen el modelo de gasto como la regresión que explica el gasto anual Y_{it} del afiliado i durante el año t , calculado como la suma del costo de todos los servicios durante un año completo, donde d_{it} es el número de días de afiliación. De esta forma, el gasto anual en salud de cada afiliado i está representado en la ecuación 6:

$$g_{it} = 360 \times \frac{Y_{it}}{d_{it}} \quad (6)$$

Por otro lado, la fórmula de ajuste por riesgo actual, utilizada por el Estado, estima, con base en los registros de un año t , la predicción del gasto anual para el año $t+2$ (Riascos *et al.*, 2017), representada en la ecuación 7:

$$\log(g_i) = \beta_0 + \sum_{j=1}^K \beta_j D_j + \varepsilon_i \quad (7)$$

Donde D es un conjunto de K variables dicotómicas que indican la presencia de ciertas características demográficas, como el sexo o el grupo etario, γ_i representa el vector de interacciones entre las i variables, y ε_i es un error aleatorio específico para cada afiliado. Los coeficientes del modelo se interpretan como el gasto adicional incurrido por un grupo de riesgo con la combinación de variables denotadas por D , comparados con la categoría base, sobre la que todas las variables categóricas toman el valor de cero, simultáneamente. La categoría base para estimar la prima tendrá una UPC definida en la ecuación 8:

$$UPC_0 = \hat{\beta}_0 \quad (8)$$

Que, de acuerdo con Riascos *et al.* (2017), es equivalente al valor global para la UPC definida por el Minsalud, en la ecuación 9:

$$UPC_t = 360 \times \frac{\sum_{i=1}^N Y_{it}}{\sum_{i=1}^N d_{it}} \quad (9)$$

Y la UPC para cada grupo de riesgo r está expresada en la ecuación 10:

$$UPC_r = UPC_0 \sum_{j=1}^K \hat{\beta}_j D_j \quad (10)$$

Que es equivalente a la expresión 11, para cualquier momento t :

$$UPC_{rt} = 360 \times \frac{\sum_{i \in r}^{N_r} Y_{it}}{\sum_{i \in r}^{N_r} d_{it}} \quad (11)$$

El Gobierno realiza una agrupación que controla por factores socio demográficos que están fuertemente correlacionados con el gasto en salud. Sin embargo, estudios como el de Riascos *et al.* (2013) (2012) (2017) (2014) demuestran que existen marcadas diferencias entre la UPC y el gasto observado en las bases de datos. Existe evidencia de que estrategias de categorización del riesgo implementadas por aseguradoras, como los trabajos de Castano & Zambrano (2006) y Gómez-Suárez (2007) demuestran. La diferencia entre el valor predicho y el valor causado para cada aseguradora, es la justificación para la implementación de una categorización más precisa de los grupos de riesgo, que la que existe actualmente.

Este estudio compara la fórmula actual del Gobierno colombiano frente a dos agrupaciones específicas de grupos de riesgo, que controlan la morbilidad,

agrupada en diferentes categorías de enfermedades crónicas de acuerdo con el trabajo de Romero (2016), teniendo en cuenta la recomendación clínica para la creación de grupos de riesgo. A diferencia de otros trabajos que utilizan técnicas estadísticas como el aprendizaje de máquinas, el modelo planteado parte de la recomendación de expertos médicos para la creación de diferentes agrupaciones de riesgo basadas en criterios que sirvan como parámetro para la identificación de cada uno de los afiliados dentro de la base de datos de un asegurador, incluyendo no solo a los que tienen algún tipo de enfermedad crónica, sino a aquellos afiliados que realizan consumos o tienen atenciones sin necesidad de tener alguna condición particular de salud.

A través del análisis de bases de datos de uso de recursos del sistema de salud (atenciones ambulatorias, hospitalarias, urgencias, servicios domiciliarios, consumo de medicamentos y uso de dispositivos médicos) para un periodo de un año, para una población objetivo, la metodología de agrupación de paciente desarrollada por (Romero, 2016), mediante algoritmos y reglas de identificación que consideran las variables demográficas, diagnósticos médicos registrados en la base de datos para cada momento de consumo, una clasificación de pacientes en categorías que son mutuamente excluyentes, priorizando la identificación de pacientes con enfermedades crónicas.

Para predecir el gasto anual de los afiliados de una aseguradora del régimen contributivo en Colombia, se toma un panel que tiene los registros individuales de quienes eran usuarios activos durante un tiempo determinado. Para la construcción de los modelos se tomaron las características demográficas mencionadas (sexo, edad, grupo etario definido por el Minsalud, ubicación geográfica en tres tipos de área, y sus interacciones).

La fuente de datos es una muestra representativa de 300.000 usuarios de la base de datos de la Suficiencia de la UPC para un asegurador representativo con más de

tres millones de afiliados, con presencia en todo el país, tanto a nivel geográfico como a nivel de incidencia y prevalencia de condiciones de salud.

La base de datos cuenta con más de ocho millones de registros de consumo para el año 2017, incluyendo información como fecha de nacimiento, edad, sexo, tipo de atención prestada (a través de códigos de Clasificación Única de Procedimientos en Salud, CUPS), información de estado de salud usando diagnósticos CIE-10 (World Health Organization, 2016) y el costo de cada atención, que es la variable a explicar.

Dado que el modelo debe explicar el gasto en salud, la variable dependiente será el gasto en salud atribuido a cada individuo en el tiempo. Las variables explicativas del modelo son los grupos de riesgo definidos.

En primer lugar, se va a estimar el modelo utilizado actualmente para el cálculo de la UPC (variables demográficas y ubicación geográfica), seguido de los modelos propuestos con los grupos de riesgo planteados, para las dos agrupaciones definidas. La primera regresión se considerará como el caso base, entendida como la fórmula actual del Gobierno. Para esto se utilizará una regresión lineal ponderada, que toma la siguiente forma, sugerida por Riascos *et al.* (2017):

$$y_{i,t+1} = \beta_0 + \sum_{j=1}^K \beta_k D_{ik,t} + \varepsilon_i \quad (12)$$

Donde $y_{i,t+1}$ es el gasto anual de un afiliado i en el año 2011, ponderado por los días de afiliación en el año ($360 (x_{i,t}/d_{i,t})$). $D_{ik,t}$ es un conjunto de variables categorías que indican la pertenencia a algún grupo de riesgo del individuo i . ε_i es el error aleatorio específico de cada afiliado i , que tiene una distribución normal. Para estimar el modelo lineal que usa el Minsalud, se utilizan mínimos cuadrados ponderados, siguiendo las fórmulas 13 y 14:

$$y_{i,t+1} = w_{i,t+1} Y_{i,t+1} \quad (13)$$

$$w_{i,t+1} = \frac{d_{i,t+1}}{360} \quad (14)$$

Una vez estimado el modelo de la ecuación 7, sujeto a las condiciones en 8 y 9, se estima un modelo que cambia los grupos de riesgo, de acuerdo con lo propuesto por Romero (2016), lo que da como resultado la ecuación 15:

$$y_{i,t+1} = \beta_0 + \sum_{j=1}^K \beta_k \varphi_{ik,t} + \varepsilon_i \quad (15)$$

Donde φ representa la agrupación propuesta de grupos de riesgo, que incluye de manera implícita los grupos propuestos por el Minsalud, como se detalla en las tablas Tabla 1. Grupos etarios Tabla 1 (pacientes no crónicos), Tabla 2 (pacientes crónicos) y Tabla 3 (pacientes multicrónicos):

Tabla 2. Grupos de riesgo: pacientes no crónicos

Código	Grupo de riesgo	Código	Grupo de riesgo	Código	Grupo de riesgo
1	Menor de un año hombre	7	Adolescentes hombre	13	Adulto hombre
2	Menor de un año mujer	8	Adolescentes mujer	14	Adulto mujer
3	Primera infancia hombre	9	Juventud hombre	15	Adulto mayor hombre
4	Primera infancia mujer	10	Juventud mujer	16	Adulto mayor mujer
5	Infancia hombre	11	Adulto joven hombre	17	Anciano hombre
6	Infancia mujer	12	Adulto joven mujer	18	Anciano mujer

Fuente: Romero (2016).

Tabla 3. Grupos de riesgo: pacientes crónicos

Código	Grupo de riesgo	Código	Grupo de riesgo	Código	Grupo de riesgo
19	Cáncer hombre	32	Neurológicas mujer	45	Demencias hombre
20	Cáncer mujer	33	Osteoarticular hombre	46	Demencias mujer
21	Cardiovascular hombre	34	Osteoarticular mujer	47	Complicaciones diabetes hombre
22	Cardiovascular mujer	35	Respiratoria crónica hombre	48	Complicaciones diabetes mujer
23	Diabetes hombre	36	Respiratoria crónica mujer	49	Hipertensión pulmonar hombre
24	Diabetes mujer	37	Salud mental hombre	50	Hipertensión pulmonar mujer
25	Gastrointestinal hombre	38	Salud mental mujer	51	Osteoarticular biológico hombre
26	Gastrointestinal mujer	39	Trasplantes hombre	52	Osteoarticular biológico mujer
27	Alto costo hombre	40	Trasplantes mujer	53	Gastrointestinal biológico hombre
28	Alto costo mujer	41	VIH hombre	54	Gastrointestinal biológico mujer

29	Interés en salud pública hombre	42	VIH mujer
30	Interés en salud pública mujer	43	Hipertensión hombre
31	Neurológicas hombre	44	Hipertensión mujer

Fuente: Romero (2016).

Tabla 4. Grupos de riesgo: pacientes multicrónicos

Código	Grupo de riesgo
57	Dos enfermedades crónicas
58	Tres enfermedades crónicas
59	Cuatro o más enfermedades crónicas
60	Diabetes & Cardiovascular

Fuente: Romero (2016).

Para comparar los modelos se utilizarán las siguientes medidas, usadas en los trabajos de Riascos *et al.* (2017) y Riascos & Vargas (2014):

- Error cuadrático medio (RMSE):

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N w_i (y_i - \hat{y}_i)^2} \quad (16)$$

- Error absoluto medio (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^N w_i |y_i - \hat{y}_i| \quad (17)$$

- Razón predictiva de gasto anualizado (PRA):

$$PRA = \frac{\sum_{i=1}^N \hat{y}_i}{\sum_{i=1}^N y_i} \quad (18)$$

- Razón predictiva de gasto no anualizado (PRNA):

$$PRNA = \frac{\sum_{i=1}^N w_i \hat{y}_i}{\sum_{i=1}^N w_i y_i} \quad (19)$$

- Pseudo R-cuadrado ($PseudoR^2$):

$$PseudoR^2 = 1 - \frac{\ln L(\widehat{Mod}_{Completo})}{\ln L(\widehat{Mod}_{Intercepto})} \quad (20)$$

Estos indicadores de predicción serán calculados para la totalidad de registros de la base de datos, así como para los quintiles más altos y más bajos del costo total

que está reflejado en la base de datos. El mejor modelo será el que tenga menores RMSE y MAE, un pseudo R-cuadrado mayor y razones predictivas cercanas a 1, tanto en la muestra completa como en los quintiles extremos del gasto de la muestra.

Dado que se propone comparar el desempeño de un modelo con variables de estado de salud frente al modelo actual, usando los grupos de riesgo de la metodología que actualmente se aplica, los dos grupos de modelos se van a estimar para la misma base de datos, y los indicadores se compararán con el objetivo de observar diferencias en el desempeño.

Para completar un análisis de sensibilidad de los parámetros estimados, y con la información de cada modelo econométrico (el propuesto y el actual), se realizarán estimaciones utilizando validación cruzada, siguiendo la metodología propuesta por Arlot & Celisse (2009) para comprobar cuál de los dos modelos es más estable, realizando iteraciones de los parámetros obtenidos y predicciones de los indicadores en simulación.

La metodología de validación cruzada (*cross-validation*) es una técnica que permite medir cómo los resultados de un análisis estadístico se pueden generalizar a una base de datos independiente. Se usa frecuentemente cuando el objetivo es predecir los valores de una variable, y se desea estimar qué tan preciso es un modelo en la práctica.

Angel & Fernandez (2015) proponen la validación cruzada como una alternativa que estima de manera más exacta y con menos supuestos, comparada con estadísticos como el Akaike Information Criterion (AIC) o el Bayesian Information Criterion (BIC).

Para la estimación de los indicadores se utilizará la técnica de *cross-validation K-fold*, de acuerdo con lo propuesto por Angel & Fernandez (2015) y Arlot & Celisse (2009), donde se divide al azar la base de datos en k grupos iguales. Se deja el k

grupo fuera de la estimación de la regresión, y se estima el modelo con las $k-1$ partes (combinadas) y se obtienen las predicciones de la k parte que se dejó fuera de la estimación. Las fórmulas para los indicadores usando este método son las siguientes:

$$CV_{RMSE} = \sum_{k=1}^k \frac{n_k}{n} RMSE_k \quad (21)$$

$$CV_{MAE} = \sum_{k=1}^k \frac{n_k}{n} MAE_k \quad (22)$$

$$CV_{PR2} = \sum_{k=1}^k \frac{n_k}{n} PseudoR_k^2 \quad (22)$$

Donde CV_i es la estimación de validación cruzada de cada indicador. k es el número de grupos en los que se ha dividido la base de datos y N es el número de registros de la base de datos. Por indicación de los autores se utilizará $k = 5$ (Angel & Fernandez, 2015).

Adicional a estas estimaciones, y para comprobar la hipótesis planteada por Riascos *et al.* (2012) (2017) y Riascos & Vargas (2014) acerca del poder predictivo de los modelos de ajuste por riesgo usados, se calcularán dos regresiones cuantílicas para los cuantiles correspondientes al 5 % y al 95 % del costo. Es decir, para el 5 % menos costoso y para el 5 % más costoso, transformando la ecuación 15, generando las siguientes regresiones para los modelos propuestos:

$$y_{i,t+1_{5\%}} = \beta_0 + \sum_{j=1}^K \beta_k \varphi_{ik,t} + \varepsilon_i \quad (23)$$

$$y_{i,t+1,95\%} = \beta_0 + \sum_{j=1}^K \beta_k \varphi_{ik,t} + \varepsilon_i \quad (24)$$

Para medir el desempeño de estos modelos para los grupos más y menos costosos se calcularán los siguientes indicadores, adaptando las ecuaciones 18 y 19 para cada caso:

$$PRA_{5\%} = \frac{\sum_{i=1}^N \hat{y}_i}{\sum_{i=1}^N y_i} \quad (25)$$

$$PRA_{95\%} = \frac{\sum_{i=1}^N \hat{y}_i}{\sum_{i=1}^N y_i} \quad (26)$$

$$PRNA_{5\%} = \frac{\sum_{i=1}^N w_i \hat{y}_i}{\sum_{i=1}^N w_i y_i} \quad (26)$$

$$PRNA_{95\%} = \frac{\sum_{i=1}^N w_i \hat{y}_i}{\sum_{i=1}^N w_i y_i} \quad (27)$$

Los tres grupos de variables a incluir en los modelos son:

- Variables usadas por el modelo actual (sin interacciones): sexo, grupo etario, ubicación geográfica.
- Variables usadas por el modelo actual (incluyendo interacciones): sexo, grupo etario, ubicación geográfica y las interacciones entre todas las variables, por pares.
- Variables del modelo de ajuste por riesgo propuesto: 60 grupos de riesgo planteados (tablas Tabla 1, Tabla 2 y Tabla 3).

La variable dependiente del modelo será incluida de dos formas:

- Costo directo de la base de datos.
- Logaritmo natural del costo registrado en la base de datos.

Si el modelo con mejor desempeño en los indicadores planteados es el que tiene la variable dependiente transformada, será necesario realizar una transformación matemática de los parámetros obtenidos en el modelo para que los resultados representen el costo asociado a cada grupo de la agrupación por riesgo.

Para lo anterior, se usarán tres transformaciones diferentes y se evaluará la que tenga mejor poder predictivo. Siguiendo la metodología propuesta por Barrientos & Firinguetti (1997) y Duan (1983), las fórmulas para la retransformación de los parámetros a valores de costo son las siguientes:

- Transformación RAW: esta transformación directa de los valores toma los valores obtenidos y los transforma usando su valor exponenciado directamente (Barrientos & Firinguetti, 1997):

$$\hat{E}(Y_0) = \exp(x_0' \hat{\beta}) \quad (28)$$

- Transformación NAIVE: esta transformación, de acuerdo con Duan (1983), considera, además de la transformación directa, la varianza de los errores, obteniendo:

$$\hat{E}(Y_0) = \exp\left(x_0' \hat{\beta} + \frac{\sigma^2}{2}\right) \quad (29)$$

- Transformación SMEARING: esta transformación, propuesta por Duan (1983), estima los valores con mayor precisión, incluso si los errores no tienen una distribución normal. La fórmula, según Barrientos & Firinguetti (1997), es:

$$\hat{E}(Y_0) = \frac{1}{n} \sum_{i=1}^n \exp(x_0' \hat{\beta} + e_i) \quad (30)$$

Una vez establecido el modelo con mejor desempeño en los indicadores, y hechas las transformaciones necesarias, se realizará una comprobación del poder predictivo global del modelo, para estimar en qué porcentaje los valores predichos en el modelo logran predecir el costo total reportado en la base de datos original, usando la siguiente fórmula:

$$Pred_{Total} = \frac{\sum_{i=1}^n Costo_{modelo}}{\sum_{i=1}^n Costo_{original}} \quad (31)$$

Al igual que para los indicadores, se realizará validación cruzada para este indicador, con un $k = 10$, para comprobar la sensibilidad del poder predictivo global de los parámetros del modelo ante cambios en la composición de la población, cuyos consumos están registrados en la base de datos.

Como validación final del modelo escogido, se genera una base de datos de prueba, con población generada aleatoriamente (50.000 usuarios). A esta población se le asigna el valor por UPC que le hubiera correspondido en el año 2017, incluyendo el valor asignado por usuario para actividades de promoción y prevención (PyP, correspondiente a 22.335,2 COP por usuario al año), de acuerdo a la Resolución 6411 de 2016 (Ministerio de Salud y Protección Social, 2016b), teniendo en cuenta su grupo etario, sexo y ubicación geográfica. A este ingreso (UPC + PyP) se le agregará un 3,5 %, correspondiente a los ingresos del asegurador por copagos y cuotas moderadoras, provenientes de los usuarios. Estos rubros constituirán el ingreso total percibido por la población aleatoria generada.

Además, a cada usuario generado de manera aleatoria se le asignará el costo anual correspondiente, tomando como referencia el comportamiento de los usuarios con

características similares en la muestra de la base de datos de suficiencia de la UPC sobre la que se calcularon los parámetros del modelo escogido.

Por último, se asignará el ingreso por UPC por grupo de riesgo correspondiente al modelo de ajuste por riesgo seleccionado. De esta forma, se generará una base de datos de prueba, donde es posible comparar el ingreso por UPC bajo el modelo actual, el ingreso por UPC con el modelo propuesto y el gasto de la población, asignando usuarios al azar con diferentes condiciones de salud, edad, sexo y ubicación geográfica. La simulación de usuarios, ingreso y costo se realizará con 1 000 iteraciones, para comprobar la estabilidad del modelo de ajuste por riesgo estimado, frente al modelo de ajuste por riesgo actual.

Para medir la diferencia entre el ingreso recibido por cada modelo y el gasto de cada usuario durante el año, se calculará la diferencia entre ingreso y gasto como porcentaje del ingreso correspondiente a UPC + PyP, de la siguiente forma:

$$Dif_i = \frac{IngresoUPC_i - GastoTotal_i}{IngresoUPC_i} \quad (32)$$

Donde i representa el modelo usado para calcular el ingreso (modelo actual o modelo propuesto).

CAPÍTULO 4. RESULTADOS DEL MODELO

Teniendo en cuenta la metodología descrita anteriormente, se presentarán los resultados en el orden en que se realizaron las estimaciones.

La base de datos con la que se realizaron las estimaciones es un registro anonimizado de consumo para el año 2017 de cerca de 300.000 usuarios seleccionados al azar, de una aseguradora representativa del régimen contributivo. La base de datos de la muestra seleccionada suma 8.361.569 registros. Las variables corresponden a la estructura reglamentaria para la base de datos de la suficiencia de la UPC, descrita por el Minsalud (2016c).

Para identificar el mejor modelo, ajustado a las ecuaciones 12 y 15, se probaron los grupos de riesgo del modelo actual (con y sin interacciones) y los 60 grupos de riesgo planteados, transformando el costo en tres modelos:

- Modelo A: $Y_i = Costo_i$ (costo sin transformar)
- Modelo B: $Y_i = Ln(Costo_i)$ (logaritmo natural del costo)
- Modelo C: $Y_i = Ln(w_i * Y_i)$ (logaritmo natural del costo ponderado por días de afiliación de cada usuario).

Los indicadores de predicción, una vez realizada la validación cruzada con $k=10$ para el cada modelo, son los siguientes (los errores estándar se reportan en paréntesis para los indicadores RMSE, MAE y pseudo R2):

Tabla 5. Indicadores modelo A

Modelo A: costo sin transformar	RMSE	MAE	Pseudo R2	PRNonAnn	PRAnn
Regresión grupos Minsalud	556,860*** (51,681)	42,633*** (225,3)	0.0000429** (0.000014)	1	1.0216002
Regresión grupos Minsalud interacciones	561,891*** (35,561)	42 624*** (147.6)	0,000103** (0.0000279)	1	1.0213866
Regresión grupos de riesgo	521,263*** (81,184)	41,030*** (258.2)	0.00370*** (0.000354)	0,78247255	1.164565

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Fuente: elaboración propia, 2019.

Los resultados del modelo A no muestran un buen poder predictivo para los indicadores planteados. Principalmente en los indicadores RMSE y MAE, donde los errores son elevados y pseudo R2, con un valor cercano a cero. Los tres grupos de indicadores son estadísticamente significativos. A pesar de que los valores no son los esperados para un modelo de predicción en ninguna de las tres transformaciones del costo, los parámetros estimados con el modelo propuesto de ajuste por riesgo tienen un menor error en el RMSE y MAE, y un mayor pseudo R2.

Los indicadores de predicción anualizados y no anualizados son similares en los dos modelos con el ajuste por riesgo actual, mientras que tienen un comportamiento contradictorio en el modelo de ajuste por riesgo propuesto.

Sin embargo, el modelo sin transformar el costo tiene errores significativamente altos, por lo que se descarta este modelo, dado el posible riesgo de sesgo en los parámetros, lo que podría llevar a subestimación o sobrestimación del costo con cualquier grupo de variables seleccionadas.

Tabla 6. Indicadores modelo B

Modelo B: costo logarítmico	RMSE	MAE	Pseudo R2	PRNonAnn	PRAnn
Regresión grupos Minsalud	1.746*** (0.00111)	1.486*** (0.000438)	0.0398*** (0.000232)	1	0.99293751
Regresión grupos Minsalud interacciones	1.745*** (0.000807)	1.485*** (0.000553)	0.0386*** (0.0000553)	1	0.99275541
Regresión grupos de riesgo	1.736*** (0.000906)	1.481*** (0.000518)	0.0572*** (0.000113)	0.81324661	1.1420559

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Fuente: elaboración propia, 2019.

El grupo de regresiones del modelo B reduce significativamente los errores de los indicadores RMSE y MAE. Además tiene valores mayores en el pseudo R2 que el modelo A. Los valores de predicción de gasto anualizado y no anualizado son similares en los tres modelos planteados. Sin embargo, el modelo con mayor pseudo R2 y menor RMSE y MAE es el de ajuste por riesgo propuesto, en comparación con los dos que usan las variables indicadas por el Minsalud. El modelo B tiene un mejor desempeño en los indicadores planteados

Tabla 7. Indicadores modelo C

Modelo C: costo logarítmico ponderado	RMSE	MAE	Pseudo R2	PRNonAnn	PRAnn
Regresión grupos Minsalud	1.748*** (0.000443)	1.484*** (0.000503)	0.0390*** (0.000387)	0.99582511	1
Regresión grupos Minsalud interacciones	1.747*** (0.000576)	1.483*** (0.000231)	0.0397*** (0.000231)	0.99543577	1
Regresión grupos de riesgo	1.738*** (0.000510)	1.480*** (0.000646)	0.0566*** (0.000370)	0.99543577	1

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Fuente: elaboración propia, 2019.

Los resultados del modelo C son similares a los del modelo B en los indicadores RMSE, MAE y pseudo R2. Sin embargo, los resultados de los indicadores de gasto anualizado y no anualizado son similares en las tres regresiones, lo que indica que la inclusión del ponderador de los días de afiliación no tiene un efecto significativo en los indicadores de error y explicación de varianza, pero sí estabiliza la predicción del costo total anualizado y no anualizado.

Dado lo anterior, se considera el modelo C como la mejor alternativa para realizar las regresiones por cuantiles. Para estas regresiones se calcularon únicamente los indicadores de gasto anualizado y no anualizado.

Tabla 8. Indicadores modelo C: regresión por cuantiles

Modelo C: costo logarítmico ponderado	5 %		95 %	
	PRNonAnn	PRAnn	PRNonAnn	PRAnn
Regresiones cuantílicas				
Regresión grupos Minsalud	0.66321945	1.4892346	1.24577	0.78601134
Regresión grupos Minsalud interacciones	0.66422367	1.4866269	1.2486665	0.78502035
Regresión grupos de riesgo	0.66422367	1.4866269	1.2486665	0.78502035

Fuente: elaboración propia, 2019.

Los resultados de las regresiones para los cuantiles 5 y 95 no muestran una diferencia significativa entre las tres regresiones, por lo que, dados los resultados, no se considera que existan diferencias significativas entre regresiones para uno u otro cuantil. Sin embargo, el modelo propuesto, junto con el modelo que incluye interacciones entre variables, son los que tienen mejor desempeño en los indicadores calculados.

Dado que el modelo C es una transformación del costo de cada registro en la base de datos, se obtienen dos alternativas de modelo para estimar el valor de cada

parámetro correspondiente a un grupo de riesgo, que corresponden a las ecuaciones 29 y 30. Los valores de la ecuación 28 no se tienen en cuenta, pues las estimaciones de los parámetros no analizan la varianza o los valores de las variables independientes:

Tabla 9. Valores estimados para cada grupo de riesgo

Código	Grupo de riesgo	Estimación NAIVE	Estimación SMEARING
1	Menor de un año hombre	\$1.248,099	\$1.318.820
2	Menor de un año mujer	\$1.597,611	\$1.688.137
3	Primera infancia hombre	\$464.119	\$490.417
4	Primera infancia mujer	\$597.558	\$631.417
5	Infancia hombre	\$276.858	\$292.546
6	Infancia mujer	\$384,680	\$406.477
7	Adolescentes hombre	\$215.622	\$227.840
8	Adolescentes mujer	\$321.244	\$339.446
9	Juventud hombre	\$355.269	\$375.400
10	Juventud mujer	\$552.376	\$583.675
11	Adulto joven hombre	\$333.113	\$351.988
12	Adulto joven mujer	\$553.874	\$585.259
13	Adulto hombre	\$311.076	\$328.703
14	Adulto mujer	\$516.466	\$545.730
15	Adulto mayor hombre	\$343.164	\$362.609
16	Adulto mayor mujer	\$500.662	\$529.031
17	Anciano hombre	\$511.099	\$540.059
18	Anciano mujer	\$536.843	\$567.262
19	Cáncer hombre	\$5.653.619	\$5.973.970
20	Cáncer mujer	\$6.134.992	\$6.482.619
21	Cardiovascular hombre	\$1.773.371	\$1.873.855
22	Cardiovascular mujer	\$2.399.213	\$2.535.160
23	Diabetes hombre	\$1.575.258	\$1.664.517
24	Diabetes mujer	\$2.094.602	\$2.213.289
25	Gastrointestinal hombre	\$2.160.737	\$2.283.171

26	Gastrointestinal mujer	\$2.903.269	\$3.067.778
27	Alto costo hombre	\$3.026.035	\$3.197.500
28	Alto costo mujer	\$4.022.526	\$4.250.456
29	Interés en salud pública hombre	\$3.076.176	\$3.250.482
30	Interés en salud pública mujer	\$4.142.706	\$4.377.445
31	Neurológicas hombre	\$2.622.529	\$2.771,130
32	Neurológicas mujer	\$3.479.423	\$3.676.578
33	Osteoarticular hombre	\$2.188.061	\$2.312.044
34	Osteoarticular mujer	\$2.960.281	\$3.128.020
35	Respiratoria Crónica hombre	\$2.234.341	\$2.360.946
36	Respiratoria Crónica mujer	\$3.018.884	\$3.189.943
37	Salud mental hombre	\$1.449.004	\$1.531.110
38	Salud mental mujer	\$2.088.970	\$2.207.338
39	Trasplantes hombre	\$10.369.822	\$10.957.409
40	Trasplantes mujer	\$13.605.323	\$14.376.244
41	VIH hombre	\$9.600.366	\$10.144.353
42	VIH mujer	\$9.397.765	\$9.930.272
43	Hipertensión hombre	\$1.134.307	\$1.198.581
44	Hipertensión mujer	\$1.528.096	\$1.614.682
45	Demencias hombre	\$3.889.164	\$4.109.536
46	Demencias mujer	\$5.713.241	\$6.036.971
47	Complicaciones diabetes hombre	\$2.858.738	\$3.020.723
48	Complicaciones diabetes mujer	\$2.982.510	\$3.151.509
49	Hipertensión pulmonar hombre	\$9.337.708	\$9.866.811
50	Hipertensión pulmonar mujer	\$11.597.084	\$12.254.212
51	Osteoarticular biológico hombre	\$2.592.620	\$2.739.525
52	Osteoarticular biológico mujer	\$3.375.674	\$3.566.951
53	Gastrointestinal biológico hombre	\$4.445.196	\$4.697.075
54	Gastrointestinal biológico mujer	\$3.224.044	\$3.406.729
55	Enfermedad renal hombre	\$3.035.434	\$3.207.432
56	Enfermedad renal mujer	\$3.251.208	\$3.435.432
57	Dos Enfermedades Crónicas	\$6.013.110	\$6.353.833
58	Tres Enfermedades Crónicas	\$8.937.259	\$9.443.673
59	Cuatro o más Enferm. Crónicas	\$18.615.021	\$19.669.807

60	Diabetes & Cardiovascular	\$3.302.616	\$3.489.753
----	---------------------------	-------------	-------------

Fuente: elaboración propia, 2019.

Estos valores son los utilizados para estimar la ecuación 31, tanto para la estimación NAIVE como para la estimación SMEARING. Lo que arroja como resultado los siguientes valores, con $k=10$:

$$Pred_{TotalNaive} = 1.0473388$$

$$Pred_{TotalSmearing} = 1.0786379$$

Estos resultados indican que el modelo con la transformación NAIVE tiene un error del 4,73 % frente al costo total reportado en la base de datos originalmente, mientras que el modelo con la transformación SMEARING tiene un error de 7,86 % frente al costo total de la base original.

Finalmente, se estiman los resultados de la ecuación 32, tomando como modelo de ajuste por riesgo las estimaciones NAIVE y SMEARING, en dos escenarios:

- Resultados sobre la base de datos:

Tabla 10. Distribución por UPC bajo modelo actual

Indicador	Valor
Ingreso UPC + PyP	\$45.847.624.100
Ingreso copagos y cuotas moderadoras	\$1.604.666.844
Ingreso total UPC	\$47.452.290.944
Gasto total	\$51.985.739.562
Diferencia ⁶	-\$6.138.115.462
Diferencia porcentual	-13.388%

Fuente: elaboración propia, 2019.

⁶ Tanto la diferencia como la diferencia porcentual se calculan frente al ingreso por UPC + PyP, pues son los valores que se pueden pronosticar en el modelo. Los modelos de ajuste por riesgo no contemplan la estimación de los ingresos por copagos y cuotas moderadoras, así que este pronóstico está fuera del alcance de esta investigación.

Como se puede observar, para la población total de la base de datos, a lo largo de un año, los ingresos por UPC + PyP no cubren los consumos de la población, lo que lleva a un déficit de 13,38 % bajo la metodología de ajuste por riesgo actual.

Tabla 11. Distribución por UPC con el modelo propuesto: NAIVE

Indicador	Valor
Ingreso UPC + PyP	\$50.032.687.764
Ingreso copagos y cuotas moderadoras	\$1.751.144.072
Ingreso total UPC	\$51.783.831.836
Gasto total	\$51.985.739.562
Diferencia ⁷	-\$1.953.051.797,82
Diferencia porcentual	-3,90 %

Fuente: elaboración propia, 2019.

Los resultados del modelo con la transformación NAIVE mejoran notablemente la diferencia entre el ingreso por UPC + PyP para la población asignada, sin embargo, aún se presenta un déficit frente al costo total de los usuarios en el año.

Tabla 12. Distribución por UPC con el modelo propuesto: SMEARING

Indicador	Valor
Ingreso UPC + PyP	\$52.867.700.034
Ingreso copagos y cuotas Moderadoras	\$1.850.369.501
Ingreso total UPC	\$54.718.069.535
Gasto total ⁸	\$51.985.739.562
Diferencia	\$881.960.471,78
Diferencia porcentual	1,67 %

^{7,8} Tanto la diferencia como la diferencia porcentual se calculan frente al ingreso por UPC + PyP, pues son los valores que se pueden pronosticar en el modelo. Los modelos de ajuste por riesgo no contemplan la estimación de los ingresos por copagos y cuotas moderadoras, así que este pronóstico está fuera del alcance de esta investigación.

Fuente: Elaboración propia, 2019.

Finalmente, el modelo en el que se usa la transformación SMEARING tiene un superávit de 1,67 % frente al ingreso. Esto significa que la transformación a los parámetros calculados hace que este sea el modelo más adecuado, pues le genera resultados más favorables al asegurador, lo que beneficia el acceso de la población asegurada.

- Resultados con un total de 1.000 iteraciones para una población de 50.000 usuarios generada al azar:

$$Dif_{UPCActual} = -8,715 \%$$

$$Dif_{UPCNaive} = -4,420 \%$$

$$Dif_{UPCSmearing} = 1,183 \%$$

Las iteraciones para estimar la diferencia entre el ingreso por UPC y el gasto total de la población durante un año completo indican la estabilidad del modelo ante cambios en la composición de la población, es decir, más o menos usuarios con enfermedades crónicas instauradas que representan tanto un ingreso como un costo mayor.

De acuerdo con estos resultados, el modelo que tiene un mejor poder predictivo y explicativo, como lo demuestran los indicadores calculados y los análisis de sensibilidad realizados, es el modelo con la transformación SMEARING.

Los valores calculados en la Tabla 8 para la estimación SMEARING son los valores por cada grupo de riesgo que generan una mejor asignación del recurso en salud, teniendo en cuenta el estado de salud de cada individuo dentro del sistema. Bajo este esquema de asignación de la UPC, una aseguradora podría contar un superávit

superior al 1 % de los ingresos recibidos, una vez terminado un año de atenciones y consumos para toda la población asegurada. Este excedente no tiene en cuenta los recaudos por copagos y cuotas moderadoras, lo que aumentaría el ingreso disponible de la aseguradora, dependiendo de sus niveles de recaudo, que para este modelo se asumieron en 3,5 %.

CONCLUSIONES

La literatura consultada aporta información que permite concluir que el ajuste por riesgo, incluyendo información sobre el estado de salud de grupos de enfermedades de interés, genera mejores resultados predictivos que los modelos que no tienen en cuenta estas variables, como lo es el modelo actual de cálculo de la UPC en Colombia para el régimen contributivo.

La revisión de la metodología actual de cálculo de UPC realizada, demuestra, a la luz de la teoría y los métodos tomados como referencia, que su cálculo presenta una insuficiencia estructural en su desarrollo. Esto se debe a que al omitir variables de estado de salud, la prima que se estima, a pesar de ser consistente estadísticamente, con un modelo econométrico estimado que cumple todos los supuestos de una regresión lineal ponderada, no es coherente con la situación financiera y epidemiológica de los aseguradores y genera en estos incentivos indirectos a negar o hacer más complejo el trámite de acceso al seguro (con el objetivo de minimizar los gastos con cargo a la UPC) debido a la asignación incorrecta de recursos destinados a la prestación de servicios de salud, lo que propicia las condiciones para un sistema de salud con indicadores de acceso y equidad más bajos, además de financieramente insolvente.

Este documento compara el ajuste por riesgo actual, realizado por el Minsalud, con una especificación alternativa de ajuste por riesgo, teniendo en cuenta las variables de estado de salud. El ajuste por riesgo, de acuerdo con los resultados, demuestra ser un factor crucial en el diseño de la política pública en salud en Colombia, pues es una variable que puede solucionar problemas de financiamiento, atención en salud y garantía en acceso a los servicios. Los ingresos por la UPC son la principal fuente de ingresos de las aseguradoras en el país, por lo que una subestimación de los verdaderos costos de atención puede incrementar los problemas financieros del sistema de salud.

Sin embargo, la identificación de pacientes en los grupos de riesgo propuestos puede generar incentivos en las aseguradoras para sobre estimar las atenciones a los pacientes con el objetivo de recibir mayores ingresos por UPC. Sin embargo, este riesgo se puede minimizar al realizar una caracterización global por parte del Minsalud, con los reportes de la base de datos de suficiencia de la UPC de cada una de las aseguradoras del país.

Por otro lado, la existencia de ajustes *ex post* como la cuenta de alto costo o los reclamos que han hecho algunas aseguradoras para recibir ingresos acordes a las necesidades de la población cubierta, son un argumento a favor de la inclusión de un modelo de ajuste por riesgo como el propuesto en el presente trabajo. Al demostrar que existe un modelo que puede generar superávit a las aseguradoras después de atender su población cubierta, se está minimizando el riesgo en el que incurren las aseguradoras, pues aunque el gasto en salud es impredecible, lograr contar con una provisión suficiente para garantizar el acceso a las atenciones que necesita una población a través de un modelo de ajuste *ex ante* es suficiente para considerar el cambio en la política pública en salud que estima los valores que se asignan por persona.

Teniendo en cuenta que el modelo que se usa actualmente tiene un poder predictivo bajo, puede asignar recursos, en promedio, de manera adecuada para cada grupo de riesgo siempre que el total de usuarios sea elevado (con una desviación estándar considerable), pero al generar un modelo con un mejor poder predictivo, como el presentado en este trabajo, es factible que el número mínimo requerido de personas afiliadas a una aseguradora que garantice solvencia financiera (además de acceso a los servicios), sea menor al actual. Esto se debe a que la identificación epidemiológica cambia el ingreso por UPC que recibe cada asegurador, beneficiando a los pacientes con enfermedades crónicas instauradas, y generando incentivos para que las aseguradoras compitan entre ellas para cubrir las atenciones de los pacientes.

Una de las limitaciones de la metodología acá presentada está en la calidad de información que reportan las aseguradoras al Minsalud. Dado que la clasificación de los pacientes en los grupos de riesgo propuestos depende de los diagnósticos que estén reportados en la base de datos de Suficiencia de la UPC, errores en el registro de diagnósticos (únicamente el 60 % de los registros de diagnóstico son válidos) pueden afectar el número de afiliados que estén en uno u otro grupo de riesgo, lo que afectaría el ingreso general del asegurador. Este aspecto se puede corregir mejorando los procesos de validación de la información y las prácticas de quienes registran los diagnósticos en los sistemas de información.

ANEXOS

Anexo A. Regresión modelo A: costo sin transformar

Variables	(1) Demog	(2) Demog_Interacc	(3) Grupos_Riesgo
dGrupoEtario2	10,196** (5,048)	774.8 (3,386)	
dGrupoEtario3	5,132** (2,075)	2,314 (3,108)	
dGrupoEtario4	4,590* (2,428)	1,281 (3,332)	
dGrupoEtario5	-803.0 (2,416)	-2,449 (3,918)	
dGrupoEtario6	5,342* (3,074)	2,506 (4,365)	
dGrupoEtario7	1,972 (2,320)	-901.6 (3,461)	
dGrupoEtario8	1,625 (2,354)	-1,973 (3,376)	
dGrupoEtario9	1,797 (2,447)	-1,894 (3,508)	
dGrupoEtario10	-359.3 (2,401)	-5,133 (3,359)	
dGrupoEtario11	35,889*** (7,083)	27,543*** (6,096)	
dGrupoEtario12	3,954* (2,372)	-1,914 (3,189)	
o.dTipoMpio1	-	-	
sex	8,469*** (906.7)	1,740 (4,010)	
1o.sex		-	
1o.dGrupoEtario2		-	
0b.sex#0b.dGrupoEtario2		0 (0)	
0b.sex#1o.dGrupoEtario2		0 (0)	
1o.sex#0b.dGrupoEtario2		0 (0)	
1.sex#1.dGrupoEtario2		23,332* (13,362)	
1o.dGrupoEtario3		-	
0b.sex#0b.dGrupoEtario3		0 (0)	
0b.sex#1o.dGrupoEtario3		0 (0)	
1o.sex#0b.dGrupoEtario3		0 (0)	

1.sex#1.dGrupoEtario3	4,435
1o.dGrupoEtario4	(4,210)
0b.sex#0b.dGrupoEtario4	0
0b.sex#1o.dGrupoEtario4	(0)
1o.sex#0b.dGrupoEtario4	0
1.sex#1.dGrupoEtario4	(0)
1o.dGrupoEtario5	6,110
0b.sex#0b.dGrupoEtario5	(5,117)
0b.sex#1o.dGrupoEtario5	0
1o.sex#0b.dGrupoEtario5	(0)
1.sex#1.dGrupoEtario5	0
1o.dGrupoEtario6	(0)
0b.sex#0b.dGrupoEtario6	2,957
0b.sex#1o.dGrupoEtario6	(4,863)
1o.sex#0b.dGrupoEtario6	0
1.sex#1.dGrupoEtario6	(0)
1o.dGrupoEtario7	4,857
0b.sex#0b.dGrupoEtario7	(6,162)
0b.sex#1o.dGrupoEtario7	0
1o.sex#0b.dGrupoEtario7	(0)
1.sex#1.dGrupoEtario7	0
1o.dGrupoEtario8	(0)
0b.sex#0b.dGrupoEtario8	0
0b.sex#1o.dGrupoEtario8	(0)
1o.sex#0b.dGrupoEtario8	0
1.sex#1.dGrupoEtario8	(0)
1o.dGrupoEtario9	4,991
0b.sex#0b.dGrupoEtario9	(4,639)
0b.sex#1o.dGrupoEtario9	0
1o.sex#0b.dGrupoEtario9	(0)
1.sex#1.dGrupoEtario9	0
1o.dGrupoEtario10	6,874
0b.sex#0b.dGrupoEtario10	(4,824)
0b.sex#1o.dGrupoEtario10	0
1o.sex#0b.dGrupoEtario10	(0)
1.sex#1.dGrupoEtario10	0

0b.sex#1o.dGrupoEtario9	0	
	(0)	
1o.sex#0b.dGrupoEtario9	0	
	(0)	
1.sex#1.dGrupoEtario9	7,083	
	(4,975)	
1o.dGrupoEtario10	-	
0b.sex#0b.dGrupoEtario10	0	
	(0)	
0b.sex#1o.dGrupoEtario10	0	
	(0)	
1o.sex#0b.dGrupoEtario10	0	
	(0)	
1.sex#1.dGrupoEtario10	9,641*	
	(4,930)	
1o.dGrupoEtario11	-	
0b.sex#0b.dGrupoEtario11	0	
	(0)	
0b.sex#1o.dGrupoEtario11	0	
	(0)	
1o.sex#0b.dGrupoEtario11	0	
	(0)	
1.sex#1.dGrupoEtario11	15,779	
	(13,645)	
1o.dGrupoEtario12	-	
0b.sex#0b.dGrupoEtario12	0	
	(0)	
0b.sex#1o.dGrupoEtario12	0	
	(0)	
1o.sex#0b.dGrupoEtario12	0	
	(0)	
1.sex#1.dGrupoEtario12	13,008**	
	(5,076)	
dGRiesgo1		167,215***
		(59,305)
dGRiesgo2		131,583**
		(52,523)
dGRiesgo3		35,703***
		(3,071)
dGRiesgo4		33,660***
		(1,995)
dGRiesgo5		28,041***
		(720.9)
dGRiesgo6		24,848***
		(504.9)
dGRiesgo7		30,641***
		(2,276)
dGRiesgo8		27,050***
		(817.9)
dGRiesgo9		29,744***
		(2,114)
dGRiesgo10		30,234***
		(659.7)

dGRiesgo11	32,753*** (1,296)
dGRiesgo12	31,640*** (711.5)
dGRiesgo13	35,041*** (2,173)
dGRiesgo14	29,793*** (932.4)
dGRiesgo15	38,660*** (3,071)
dGRiesgo16	31,943*** (1,796)
dGRiesgo17	66,373*** (13,136)
dGRiesgo18	49,276*** (5,697)
dGRiesgo19	193,356*** (31,455)
dGRiesgo20	144,362*** (13,479)
dGRiesgo21	65,684*** (6,746)
dGRiesgo22	59,106*** (6,097)
dGRiesgo23	35,696*** (2,655)
dGRiesgo24	35,696*** (2,813)
dGRiesgo25	62,433*** (13,828)
dGRiesgo26	49,910*** (4,660)
dGRiesgo27	516,432*** (137,426)
dGRiesgo28	101,936*** (24,515)
dGRiesgo29	77,539*** (20,713)
dGRiesgo30	61,580*** (11,700)
dGRiesgo31	104,182*** (13,128)
dGRiesgo32	90,823*** (16,887)
dGRiesgo33	40,488*** (3,207)
dGRiesgo34	35,632*** (2,079)
dGRiesgo35	34,976*** (2,112)
dGRiesgo36	46,393*** (4,068)
dGRiesgo37	59,866*** (5,852)
dGRiesgo38	34,111*** (1,801)

dGRiesgo39			140,521*** (29,853)
dGRiesgo40			103,775*** (21,589)
dGRiesgo41			163,395*** (5,401)
dGRiesgo42			157,724*** (18,326)
dGRiesgo43			21,142*** (1,239)
dGRiesgo44			19,361*** (574.9)
dGRiesgo45			58,274*** (8,638)
dGRiesgo46			58,401*** (7,539)
dGRiesgo47			38,729*** (13,140)
dGRiesgo48			34,827*** (4,170)
dGRiesgo49			38,760*** (8,188)
dGRiesgo50			67,469** (32,554)
dGRiesgo51			76,871*** (13,398)
dGRiesgo52			167,381*** (19,117)
dGRiesgo53			100,963** (51,383)
dGRiesgo54			72,941*** (25,801)
dGRiesgo55			576,330*** (25,165)
dGRiesgo56			644,702*** (190,294)
dGRiesgo57			43,664*** (1,505)
dGRiesgo58			49,207*** (2,650)
dGRiesgo59			50,808*** (3,200)
dGRiesgo60			22,594*** (814.2)

Constant	30,729*** (2,057)	34,271*** (3,062)
----------	----------------------	----------------------

Observations	8,361,569	8,361,569	8,361,569
R-squared	0.000	0.000	0.008

Nota: errores estándar robustos en paréntesis.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Fuente: elaboración propia, 2019.

Anexo B. Regresión modelo B: costo logarítmico

Variables	(1) Demog	(2) Demog_Interacc	(3) Grupos_Riesgo
dGrupoEtario2	0.273*** (0.00900)	0.290*** (0.0121)	
dGrupoEtario3	0.128*** (0.00651)	0.196*** (0.00925)	
dGrupoEtario4	-0.0583*** (0.00779)	-0.0424*** (0.0108)	
dGrupoEtario5	0.290*** (0.00728)	0.296*** (0.0105)	
dGrupoEtario6	-0.199*** (0.00761)	-0.186*** (0.0106)	
dGrupoEtario7	-0.363*** (0.00756)	-0.366*** (0.0105)	
dGrupoEtario8	-0.502*** (0.00756)	-0.522*** (0.0105)	
dGrupoEtario9	-0.598*** (0.00765)	-0.610*** (0.0107)	
dGrupoEtario10	-0.678*** (0.00804)	-0.711*** (0.0112)	
dGrupoEtario11	0.599*** (0.0115)	0.606*** (0.0169)	
dGrupoEtario12	-0.676*** (0.00726)	-0.711*** (0.0101)	
o.dTipoMpio1	-	-	
sex	0.0144*** (0.00258)	0.0443*** (0.0122)	
1o.sex		-	
1o.dGrupoEtario2		-	
0b.sex#0b.dGrupoEtario2		0 (0)	
0b.sex#1o.dGrupoEtario2		0 (0)	
1o.sex#0b.dGrupoEtario2		0 (0)	
1.sex#1.dGrupoEtario2		-0.0352* (0.0185)	
1o.dGrupoEtario3		-	
0b.sex#0b.dGrupoEtario3		0 (0)	
0b.sex#1o.dGrupoEtario3		0 (0)	
1o.sex#0b.dGrupoEtario3		0 (0)	
1.sex#1.dGrupoEtario3		-0.196*** (0.0131)	
1o.dGrupoEtario4		-	
0b.sex#0b.dGrupoEtario4		0 (0)	

0b.sex#1o.dGrupoEtario4	0
	(0)
1o.sex#0b.dGrupoEtario4	0
	(0)
1.sex#1.dGrupoEtario4	-0.0302*
	(0.0158)
1o.dGrupoEtario5	-
0b.sex#0b.dGrupoEtario5	0
	(0)
0b.sex#1o.dGrupoEtario5	0
	(0)
1o.sex#0b.dGrupoEtario5	0
	(0)
1.sex#1.dGrupoEtario5	-0.00914
	(0.0146)
1o.dGrupoEtario6	-
0b.sex#0b.dGrupoEtario6	0
	(0)
0b.sex#1o.dGrupoEtario6	0
	(0)
1o.sex#0b.dGrupoEtario6	0
	(0)
1.sex#1.dGrupoEtario6	-0.0226
	(0.0154)
1o.dGrupoEtario7	-
0b.sex#0b.dGrupoEtario7	0
	(0)
0b.sex#1o.dGrupoEtario7	0
	(0)
1o.sex#0b.dGrupoEtario7	0
	(0)
1.sex#1.dGrupoEtario7	0.0198
	(0.0153)
1o.dGrupoEtario8	-
0b.sex#0b.dGrupoEtario8	0
	(0)
0b.sex#1o.dGrupoEtario8	0
	(0)
1o.sex#0b.dGrupoEtario8	0
	(0)
1.sex#1.dGrupoEtario8	0.0637***
	(0.0153)
1o.dGrupoEtario9	-
0b.sex#0b.dGrupoEtario9	0
	(0)
0b.sex#1o.dGrupoEtario9	0
	(0)
1o.sex#0b.dGrupoEtario9	0
	(0)
1.sex#1.dGrupoEtario9	0.0356**
	(0.0154)

1o.dGrupoEtario10	-	
0b.sex#0b.dGrupoEtario10	0	
	(0)	
0b.sex#1o.dGrupoEtario10	0	
	(0)	
1o.sex#0b.dGrupoEtario10	0	
	(0)	
1.sex#1.dGrupoEtario10	0.0864***	
	(0.0162)	
1o.dGrupoEtario11	-	
0b.sex#0b.dGrupoEtario11	0	
	(0)	
0b.sex#1o.dGrupoEtario11	0	
	(0)	
1o.sex#0b.dGrupoEtario11	0	
	(0)	
1.sex#1.dGrupoEtario11	-0.0142	
	(0.0231)	
1o.dGrupoEtario12	-	
0b.sex#0b.dGrupoEtario12	0	
	(0)	
0b.sex#1o.dGrupoEtario12	0	
	(0)	
1o.sex#0b.dGrupoEtario12	0	
	(0)	
1.sex#1.dGrupoEtario12	0.105***	
	(0.0147)	
dGRiesgo1		10.28***
		(0.0840)
dGRiesgo2		10.28***
		(0.0850)
dGRiesgo3		9.235***
		(0.00883)
dGRiesgo4		9.195***
		(0.00927)
dGRiesgo5		9.390***
		(0.00903)
dGRiesgo6		9.348***
		(0.00913)
dGRiesgo7		9.349***
		(0.00944)
dGRiesgo8		9.352***
		(0.00842)
dGRiesgo9		9.036***
		(0.0110)
dGRiesgo10		9.312***
		(0.00615)
dGRiesgo11		9.115***
		(0.00624)
dGRiesgo12		9.316***
		(0.00388)
dGRiesgo13		9.284***
		(0.00721)

dGRiesgo14	9.218*** (0.00560)
dGRiesgo15	9.374*** (0.0122)
dGRiesgo16	9.102*** (0.0107)
dGRiesgo17	9.444*** (0.0304)
dGRiesgo18	9.154*** (0.0274)
dGRiesgo19	10.11*** (0.0243)
dGRiesgo20	9.923*** (0.0168)
dGRiesgo21	8.786*** (0.0146)
dGRiesgo22	8.899*** (0.0151)
dGRiesgo23	9.298*** (0.0152)
dGRiesgo24	9.188*** (0.0146)
dGRiesgo25	9.523*** (0.0336)
dGRiesgo26	9.467*** (0.0233)
dGRiesgo27	10.10*** (0.0483)
dGRiesgo28	9.807*** (0.0289)
dGRiesgo29	9.844*** (0.0641)
dGRiesgo30	9.796*** (0.0459)
dGRiesgo31	9.809*** (0.0279)
dGRiesgo32	9.694*** (0.0287)
dGRiesgo33	9.289*** (0.0173)
dGRiesgo34	9.226*** (0.0122)
dGRiesgo35	9.372*** (0.0225)
dGRiesgo36	9.353*** (0.0177)
dGRiesgo37	9.388*** (0.0167)
dGRiesgo38	9.000*** (0.0118)
dGRiesgo39	10.21*** (0.182)
dGRiesgo40	9.696*** (0.187)
dGRiesgo41	10.39*** (0.0421)

dGRiesgo42			10.16*** (0.0846)
dGRiesgo43			8.342*** (0.00545)
dGRiesgo44			8.352*** (0.00433)
dGRiesgo45			9.811*** (0.145)
dGRiesgo46			9.615*** (0.0993)
dGRiesgo47			9.179*** (0.0414)
dGRiesgo48			9.337*** (0.0376)
dGRiesgo49			10.15*** (0.179)
dGRiesgo50			10.45*** (0.162)
dGRiesgo51			9.852*** (0.0725)
dGRiesgo52			9.804*** (0.0598)
dGRiesgo53			9.654*** (0.916)
dGRiesgo54			10.03*** (0.374)
dGRiesgo55			10.67*** (0.0642)
dGRiesgo56			10.57*** (0.0624)
dGRiesgo57			8.683*** (0.00475)
dGRiesgo58			8.651*** (0.00557)
dGRiesgo59			8.808*** (0.0114)
dGRiesgo60			8.435*** (0.00430)
Constant	9.135*** (0.00624)	9.120*** (0.00887)	
Observations	8,361,569	8,361,569	8,361,569
R-squared	0.040	0.041	0.964

Nota: errores estándar robustos en paréntesis

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$*

Fuente: elaboración propia, 2019.

Anexo C. Regresión modelo C: costo logarítmico ponderado

Variables	(1) Demog	(2) Demog_Interacc	(3) Grupos_Riesgo
dGrupoEtario2	0.273*** (0.0235)	0.272*** (0.0300)	
dGrupoEtario3	0.0513*** (0.0146)	0.118*** (0.0210)	
dGrupoEtario4	-0.293*** (0.0163)	-0.261*** (0.0229)	
dGrupoEtario5	0.292*** (0.0188)	0.298*** (0.0272)	
dGrupoEtario6	-0.447*** (0.0156)	-0.412*** (0.0222)	
dGrupoEtario7	-0.595*** (0.0153)	-0.582*** (0.0219)	
dGrupoEtario8	-0.714*** (0.0151)	-0.714*** (0.0217)	
dGrupoEtario9	-0.780*** (0.0151)	-0.778*** (0.0217)	
dGrupoEtario10	-0.850*** (0.0154)	-0.865*** (0.0221)	
dGrupoEtario11	0.470*** (0.0300)	0.481*** (0.0441)	
dGrupoEtario12	-0.852*** (0.0147)	-0.865*** (0.0212)	
o.dTipoMpio1	-	-	
sex	0.0282*** (0.00438)	0.0615** (0.0277)	
1o.sex		-	
1o.dGrupoEtario2		-	
0b.sex#0b.dGrupoEtario2		0 (0)	
0b.sex#1o.dGrupoEtario2		0 (0)	
1o.sex#0b.dGrupoEtario2		0 (0)	
1.sex#1.dGrupoEtario2		0.0366 (0.0509)	
1o.dGrupoEtario3		-	
0b.sex#0b.dGrupoEtario3		0 (0)	
0b.sex#1o.dGrupoEtario3		0 (0)	
1o.sex#0b.dGrupoEtario3		0 (0)	
1.sex#1.dGrupoEtario3		-0.228*** (0.0298)	
1o.dGrupoEtario4		-	
0b.sex#0b.dGrupoEtario4		0 (0)	

0b.sex#1o.dGrupoEtario4	0
	(0)
1o.sex#0b.dGrupoEtario4	0
	(0)
1.sex#1.dGrupoEtario4	-0.0730**
	(0.0331)
1o.dGrupoEtario5	-
0b.sex#0b.dGrupoEtario5	0
	(0)
0b.sex#1o.dGrupoEtario5	0
	(0)
1o.sex#0b.dGrupoEtario5	0
	(0)
1.sex#1.dGrupoEtario5	-0.00936
	(0.0377)
1o.dGrupoEtario6	-
0b.sex#0b.dGrupoEtario6	0
	(0)
0b.sex#1o.dGrupoEtario6	0
	(0)
1o.sex#0b.dGrupoEtario6	0
	(0)
1.sex#1.dGrupoEtario6	-0.0827***
	(0.0316)
1o.dGrupoEtario7	-
0b.sex#0b.dGrupoEtario7	0
	(0)
0b.sex#1o.dGrupoEtario7	0
	(0)
1o.sex#0b.dGrupoEtario7	0
	(0)
1.sex#1.dGrupoEtario7	-0.0199
	(0.0308)
1o.dGrupoEtario8	-
0b.sex#0b.dGrupoEtario8	0
	(0)
0b.sex#1o.dGrupoEtario8	0
	(0)
1o.sex#0b.dGrupoEtario8	0
	(0)
1.sex#1.dGrupoEtario8	0.0126
	(0.0304)
1o.dGrupoEtario9	-
0b.sex#0b.dGrupoEtario9	0
	(0)
0b.sex#1o.dGrupoEtario9	0
	(0)
1o.sex#0b.dGrupoEtario9	0
	(0)
1.sex#1.dGrupoEtario9	0.00500
	(0.0303)

1o.dGrupoEtario10	-	
0b.sex#0b.dGrupoEtario10	0	
	(0)	
0b.sex#1o.dGrupoEtario10	0	
	(0)	
1o.sex#0b.dGrupoEtario10	0	
	(0)	
1.sex#1.dGrupoEtario10	0.0437	
	(0.0309)	
1o.dGrupoEtario11	-	
0b.sex#0b.dGrupoEtario11	0	
	(0)	
0b.sex#1o.dGrupoEtario11	0	
	(0)	
1o.sex#0b.dGrupoEtario11	0	
	(0)	
1.sex#1.dGrupoEtario11	-0.0210	
	(0.0601)	
1o.dGrupoEtario12	-	
0b.sex#0b.dGrupoEtario12	0	
	(0)	
0b.sex#1o.dGrupoEtario12	0	
	(0)	
1o.sex#0b.dGrupoEtario12	0	
	(0)	
1.sex#1.dGrupoEtario12	0.0521*	
	(0.0295)	
dGRiesgo1		10.14***
		(0.819)
dGRiesgo2		10.23***
		(0.835)
dGRiesgo3		9.264***
		(0.0205)
dGRiesgo4		9.216***
		(0.0215)
dGRiesgo5		9.403***
		(0.0291)
dGRiesgo6		9.353***
		(0.0297)
dGRiesgo7		9.382***
		(0.0329)
dGRiesgo8		9.401***
		(0.0261)
dGRiesgo9		8.943***
		(0.0328)
dGRiesgo10		9.382***
		(0.0140)
dGRiesgo11		9.094***
		(0.0176)
dGRiesgo12		9.352***
		(0.00869)
dGRiesgo13		9.241***
		(0.0202)

dGRiesgo14	9.147*** (0.0129)
dGRiesgo15	9.278*** (0.0313)
dGRiesgo16	8.968*** (0.0222)
dGRiesgo17	9.294*** (0.0688)
dGRiesgo18	8.965*** (0.0542)
dGRiesgo19	10.07*** (0.0414)
dGRiesgo20	9.930*** (0.0257)
dGRiesgo21	8.657*** (0.0186)
dGRiesgo22	8.765*** (0.0196)
dGRiesgo23	9.273*** (0.0269)
dGRiesgo24	9.163*** (0.0238)
dGRiesgo25	9.458*** (0.0681)
dGRiesgo26	9.435*** (0.0424)
dGRiesgo27	10.04*** (0.0768)
dGRiesgo28	9.800*** (0.0515)
dGRiesgo29	9.877*** (0.140)
dGRiesgo30	9.818*** (0.0945)
dGRiesgo31	9.888*** (0.0466)
dGRiesgo32	9.764*** (0.0492)
dGRiesgo33	9.259*** (0.0327)
dGRiesgo34	9.178*** (0.0211)
dGRiesgo35	9.346*** (0.0406)
dGRiesgo36	9.294*** (0.0281)
dGRiesgo37	9.339*** (0.0312)
dGRiesgo38	8.943*** (0.0196)
dGRiesgo39	10.17*** (0.238)
dGRiesgo40	9.531*** (0.211)
dGRiesgo41	10.24*** (0.0637)

dGRiesgo42			10.07*** (0.128)
dGRiesgo43			8.231*** (0.00772)
dGRiesgo44			8.261*** (0.00601)
dGRiesgo45			9.755*** (0.242)
dGRiesgo46			9.721*** (0.136)
dGRiesgo47			9.158*** (0.0690)
dGRiesgo48			9.324*** (0.0603)
dGRiesgo49			10.11*** (0.495)
dGRiesgo50			10.66*** (0.914)
dGRiesgo51			9.693*** (0.105)
dGRiesgo52			9.682*** (0.0718)
dGRiesgo53			9.460*** (1.503)
dGRiesgo54			9.833*** (0.642)
dGRiesgo55			10.48*** (0.0653)
dGRiesgo56			10.45*** (0.0690)
dGRiesgo57			8.632*** (0.00586)
dGRiesgo58			8.630*** (0.00643)
dGRiesgo59			8.805*** (0.0126)
dGRiesgo60			8.407*** (0.00549)
Constant	9.233*** (0.0140)	9.215*** (0.0202)	
Observations	8,361,569	8,361,569	8,361,569
R-squared	0.043	0.043	0.960

Nota: Errores estándar robustos en paréntesis

**** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$*

Fuente: elaboración propia, 2019.

Bibliografía

- A Riascos, R Guerrero, D Chernichovsky, J. D. (2013). Cálculo de la UPC y su ajuste por riesgo. *Políticas En Breve*. Retrieved from https://www.icesi.edu.co/proesa/images/docs/Boletin_5.pdf
- Angel, M., & Fernandez, L. (2015). Cross-validation Cross-validation Cross-validation methods Examples : Model selection References.
- Arlot, S., & Celisse, A. (2009). A survey of cross-validation procedures for model selection, *4*, 40–79. <https://doi.org/10.1214/09-SS054>
- Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *The American Economic Review*, Vol LIII(No. 5), 941–973. Retrieved from https://web.stanford.edu/~jay/health_class/Readings/Lecture01/arrow.pdf
- Barrientos, A., & Firinguetti, L. (1997). Predicting wages from transformed wage equations. *Applied Economics Letters*, *4*(11), 709–710. <https://doi.org/10.1080/758530654>
- Castano, R., & Zambrano, A. (2006). Biased selection within the social health insurance market in Colombia. *Health Policy*, *79*(2–3), 313–324. <https://doi.org/10.1016/j.healthpol.2006.01.010>
- Cepal. (2008). Transformaciones demográficas y su influencia en el desarrollo en América Latina y el Caribe., *2378*, 121.
- Chicaiza Becerra, L. A., & Cabedo Semper, J. D. (2007). Las opciones financieras como mecanismo para estimar las primas de seguro y reaseguro en el sistema de salud colombiano. *Cuadernos de Administración*, *19*(34), 221–236.
- CRES. ACUERDO 009 DE 2009 (2010). Colombia.
- Duan, N. (1983). A Nonparametric Smearing Estimate : Method Retransformation. *Journal of the American Statistical Association*, *78*(383), 605–610.
- Ellis, R. P. (2008). Risk Adjustment in Health Care Markets: Concepts and Applications. *Financing Health Care: New Ideas for a Changing Society*, 177–222. <https://doi.org/10.1002/9783527611294.ch8>
- Flórez Pinilla, K. (2010). La Competencia Regulada en el Sistema de Salud Colombiano. *Revista LEBRET*, *0*(2), 73–99. <https://doi.org/10.15332/rl.v0i2.666>

- Gomez-Suarez, R. (2007). Cream-skimming and risk adjustment in Colombian health insurance system: The public insurer case. *Archivos de Economía*, (336), 1–68.
- Jairo, R., & Kristian, R. (2016). La génesis de la Economía de la Salud en Kenneth Arrow (1963) / La genèse de l'économie de la santé chez Kenneth Arrow (1963) / The genesis of Health Economics in Kenneth Arrow (1963). *Lecturas de Economía VO - 0*, (84), 209. <https://doi.org/10.17533/udea.le.n84a07>
- Ministerio de Salud y Protección Social. (2006). *Estudio de Suficiencia y de los Mecanismos de Ajuste de Riesgo para el Cálculo de la Unidad de Pago por Capitación para garantizar el Plan Obligatorio de Salud 2014* (Vol. 1). Ministerio de Salud y Protección Social.
- Ministerio de Salud y Protección Social. (2008). *Estudio de Suficiencia y de los Mecanismos de Ajuste de Riesgo para el Cálculo de la Unidad de Pago por Capitación para garantizar el Plan Obligatorio de Salud 2018* (Vol. 1).
- Ministerio de Salud y Protección Social. (2012). *Decreto 2562 de 2012*.
- Ministerio de Salud y Protección Social. (2013). *Estudio de Suficiencia y de los Mecanismos de Ajuste de Riesgo para el Cálculo de la Unidad de Pago por Capitación para garantizar el Plan Obligatorio de Salud 2013*.
- Ministerio de Salud y Protección Social. (2014). *Estudio de Suficiencia y de los Mecanismos de Ajuste de Riesgo para el Cálculo de la Unidad de Pago por Capitación para garantizar el Plan Obligatorio de Salud 2014* (Vol. 1). Ministerio de Salud y Protección Social.
- Ministerio de Salud y Protección Social. (2015). *Estudio de Suficiencia y de los Mecanismos de Ajuste de Riesgo para el Cálculo de la Unidad de Pago por Capitación para garantizar el Plan Obligatorio de Salud 2015*.
- Ministerio de Salud y Protección Social. (2016a). *ESTUDIO DE SUFICIENCIA Y DE LOS MECANISMOS DE AJUSTE DE RIESGO PARA EL CÁLCULO DE LA UNIDAD DE PAGO DE CAPITACIÓN PARA GARANTIZAR EL PLAN DE BENEFICIOS EN SALUD*. Retrieved from <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VP/RBC/estudio-suficiencia-upc-2016.pdf>

- Ministerio de Salud y Protección Social. (2016b). Resolución 6411 de 2016, 2016. Retrieved from http://cuentadealtocosto.org/site/images/Publicaciones/normatividad2016/Resolucion_6411_de_2016.pdf
- Ministerio de Salud y Protección Social. (2016c). *SOLICITUD INFORMACIÓN A ENTIDADES PROMOTORAS DE SALUD PARA EL “ESTUDIO DE SUFICIENCIA Y DE LOS MECANISMOS DE AJUSTE DE RIESGO DE LA UNIDAD DE PAGO POR CAPITACIÓN PARA GARANTIZAR EL PLAN DE BENEFICIOS CON CARGO A LA UPC AÑO 2018* (Vol. 2). Retrieved from <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VP/RBC/Solicitud-informacion-estudio-suficiencia-2016.pdf>
- Ministerio de Salud y Protección Social. (2018). *Unidad de Pago por Capitación UPC*. Retrieved from <https://www.minsalud.gov.co/sites/rid/Lists/BibliotecaDigital/RIDE/VP/RBC/soc-unidad-de-pago-por-capitacion-2018.pdf>
- Moreno, J. R. (2010). *Derechos de propiedad del Seguro Obligatorio de Salud en Colombia* (I). Pontificia Universidad Javeriana. Retrieved from <http://www.jstor.org/stable/j.ctt169zszw>
- OMS. (2010). La Financiación De Los Sistemas De Salud.
- Perez, J. L. (2016). Conociendo el Seguro: Teoría General Del Seguro. *Research Gate*, (March), 448. Retrieved from <https://www.researchgate.net/publication/40942409> Conociendo
- Restrepo, D., & López, L. (2012). La unidad de pago por capitación: 17 años en el Sistema General de Seguridad Social en Salud. *Revista Facultad Nacional de Salud Pública*, 30(3), 9. Retrieved from <http://www.scielo.org.co/pdf/rfnsp/v30n3/v30n3a05.pdf>
- Riascos, A., Alfonso, E., & Romero, M. (2012). The Performance of Risk Adjustment Models in Colombian Competitive Health Insurance Market. *Documentos CEDE*, (1657–7191), 20. <https://doi.org/10.2202/1538-0645.1460>
- Riascos, A., Romero, M., & Serna, N. (2017). Risk Adjustment Revisited using Machine Learning Techniques. *Documentos CEDE*, 27, 23.

- Riascos, A., & Vargas, J. F. (2014). The Performance of Risk Adjustment Models in Colombian Competitive Health Insurance Market. *Serie Documentos Cede*, (1657–7191), 76. <https://doi.org/10.2202/1538-0645.1460>
- Romero, M. (2016). Developing A Classification System of Patients Based on Risk Related to Epidemiological and Demographic Variables, *19*(7). Retrieved from https://www.valueinhealthjournal.com/pb/assets/raw/HealthAdvance/journals/jval/JVAL_19_7.pdf
- Schokkaert, E., & Van de Voorde, C. (2004). Risk selection and the specification of the conventional risk adjustment formula. *Journal of Health Economics*, *23*(6), 1237–1259. [https://doi.org/10.1016/S0167-6296\(03\)00040-7](https://doi.org/10.1016/S0167-6296(03)00040-7)
- World Health Organization. (2016). ICD-10 Version:2016. Retrieved May 25, 2019, from <https://icd.who.int/browse10/2016/en>