



**TESIS DE GRADO  
EN INGENIERIA INDUSTRIAL**

**SCORING DE RIESGO PARA ENFERMOS  
CRONICOS EN UNA COMPAÑÍA  
DE MEDICINA PREPAGA**

**Autor: Alejandro Pablo Abatti**

**Directores de la tesis:**

**M.Ing. Alejandra Ochoa**

**Dr. Ezequiel García Elorrio**

**Dr. Ramón García Martínez**

**2007**



## Resumen Ejecutivo

La disponibilidad de fuentes de información ha crecido principalmente gracias al desarrollo tecnológico para el manejo y tratamiento de la misma, contribuyendo a que ésta se produzca en mayor cantidad, con mejor calidad y a bajo costo. Sin embargo, aún hoy muchas empresas desconocen tanto el valor de la misma, como técnicas efectivas para la explotación y búsqueda eficiente del conocimiento deseado.

Uno de los sectores económicos que hoy comienzan a requerir un uso adecuado de los sistemas de información es el sector de la medicina prepaga, donde a partir de la implementación de programas de Disease Management, el conocimiento sobre el consumo y el estado de salud del afiliado es necesario y debe ser analizado a fin de focalizar los recursos eficazmente sobre aquellos grupos de mayor riesgo potencial, en particular sobre aquellos afiliados que padecen enfermedades crónicas.

El presente trabajo propone una alternativa de explotación de información para facilitar el gerenciamiento de enfermos crónicos en los programas de Disease Management de una compañía de medicina prepaga en la Argentina. La alternativa propuesta encara la explotación de información desde un enfoque de minería de datos que combina distintas variables clínicas, demográficas, económicas, biométricas y del estado de salud, detectando el patrón de consumo de distintos subgrupos poblacionales, y obteniendo un scoring de riesgo basado en índices complementarios que cada individuo presente en un momento dado.



## **Executive Brief**

The access to information sources has been growing mainly due to the technological developments for information treatment, thus helping to generate greater quantities of data, with better quality and lower costs. Nevertheless not many companies nowadays are acknowledging either its value, its effective exploitation techniques or its efficient search of the desirable knowledge.

One of the economic industries that is requiring the appropriate use of the information systems is the Health Care Industry, which, by implementing Disease Management Programs, is requiring to know and understand the health status of their affiliates and their consuming tendencies, in order to apply resources efficiently in high risk affiliates, specially in the ones who suffer from chronic diseases.

This paper presents an information exploitation alternative to improve the Management of chronic affiliates in the Disease Management programs of an Argentinean Health Care Company. The alternative presented focalizes the information exploitation from a data mining perspective, combining different variables of clinical, demographic, economic, biometric and health status data, detecting the consuming pattern of the different population segments, and obtaining a scoring based on the risk that each affiliate presents at a given time.





## **Agradecimientos**

A mis padres por todo su apoyo, confianza incondicional y esfuerzo para brindarme la mejor educación.

A la M.Ing. Alejandra Ochoa y el Dr.Ramón García Martínez por su paciencia y valiosos comentarios que me ayudaron a plasmar este trabajo.

Al Dr.García Elorrio por su apoyo y guía constante que me llevaron a buen puerto.

A Sara, Patsy, Ana y María José del Equipo de Coordinación de Cuidados Médicos de OMINT S.A., por toda su voluntad y aliento a lo largo del desarrollo del trabajo.

Al Dr. Santiago González Calvo por sus valiosas observaciones y aportes.

Buenos Aires, Mayo 2007.







---

## INDICE

<b>1. INTRODUCCIÓN</b>	<b>1</b>
<b>2. ESTADO DE LA CUESTIÓN</b>	<b>3</b>
2.1 Paradigmas del mercado actual	3
2.1.1 Tendencia demográfica	3
2.1.2 Tecnología	5
2.1.3 Expectativa de los afiliados	5
2.2 Amenazas y Oportunidades del mercado	6
2.2.1 Amenazas a corto plazo	7
2.2.1.1 Imposibilidad de trasladar totalmente los aumentos de costos a las cuotas.	7
2.2.1.2 Incremento de los riesgos potenciales en la prestación de servicios	7
2.2.1.3 Implementación de un nuevo marco legal para el sector	7
2.2.1.4 Mayores controles de precios	8
2.2.2 Oportunidades a corto plazo	8
2.2.2.1 Incremento de los acuerdos con las obras sociales	8
2.2.2.2 Obtención de una reducción de la carga tributaria	8
2.2.2.3 Expansión al interior del país	8
2.2.2.4 Menor presión de los sustitutos	9
2.2.3 Amenazas a mediano y largo plazo	9
2.2.3.1 Saturación del sector	9
2.2.3.2 Incremento de la desigualdad social y consecuente intensificación de la competencia por un mercado limitado	9
2.2.3.3 Obsolescencia de la Tecnología	10
2.2.3.4 Arribo de nuevos competidores	10
2.2.3.5 Envejecimiento de la población	10
2.2.4 Oportunidades a mediano y largo plazo	10
2.2.4.1 Deficiencias de la salud pública	10
2.2.4.2 Inequidad en la distribución geográfica de la prestación de servicios	11
2.2.4.3 Aumento del nivel de ingresos de la población	11
2.2.4.4 Incremento de la integración vertical	11
2.2.4.5 Incremento de la sustitución de las importaciones de los insumos médicos	12
2.2.4.6 Utilización de las fuentes de información sobre comportamientos de los afiliados	12
2.3 Conclusiones sobre las oportunidades y amenazas del mercado	12
2.4 Estado de los sistemas de flujo de Información	13
2.5 Flujo de Información en la Argentina	14
2.5.1 Flujo de información de consultas médicas en la Argentina	14
2.5.2 Flujo de información Sanatorial en la Argentina	15
2.5.3 Flujo de información de Farmacias en la Argentina	15
2.5.4 Flujo de información de Laboratorios en la Argentina	16
2.5.5 Flujo de información de los Centros de Imágenes en la Argentina	16
2.5.6 Flujo de información de Vacunatorios en la Argentina	17
2.5.7 Flujo de información de servicios de Home Care en la Argentina	17
2.5.8 Flujo de información de Servicios de Emergencia en la Argentina	17
2.5.9 Flujo de información de Salud Mental en la Argentina	18
2.5.10 Conclusiones sobre el flujo de Información en la Argentina	18
2.6 Flujo de Información en los Estados Unidos	19
2.6.1 Flujo de Información de consultas médicas en EEUU	19
2.6.2 Flujo de Información Sanatorial en EEUU	20
2.6.3 Flujo de Información de Farmacia en EEUU	20
2.6.4 Flujo de Información de Laboratorio en EEUU	21
2.6.5 Flujo de Información de Vacunatorios en EEUU	21
2.6.6 Flujo de Información de Centros de Imágenes en EEUU	22
2.6.7 Flujo de Información de Centros de Salud Mental en EEUU	22

2.6.8 Flujo de Información de Servicios de Home Care en EEUU	23
2.6.9 Flujo de Información de Servicios de Emergencia en EEUU	23
2.6.10 Conclusiones sobre el flujo de Información en EEUU	24
<b>2.7. Explotación de la Información para estrategias de “Disease Management”</b>	<b>24</b>
2.7.1 ¿Qué es el Disease Management?	25
2.7.2 El Disease Management en la Argentina	25
2.7.3 El Disease Management del Hospital Italiano – Herramientas utilizadas	26
2.7.4 Programas de Disease Management en OMINT S.A. - Caso de aplicación de la tesis	26
<b>2.8 Herramientas de reconocimiento de patrones</b>	<b>27</b>
2.8.1 Aprendizaje de Máquinas (Machine Learning)	27
2.8.2 Support Vector Machines (SVM)	28
2.8.3 Aprendizaje Bayesiano (Bayesian Learning)	28
2.8.4 Análisis de Componentes Independientes (Independent Component Análisis)	30
2.8.5 Redes Neuronales (Neural Networks)	31
2.8.6 Aprendizaje No Supervisado (Unsupervised Learning)	32
2.8.7 Mapas Autoorganizativos (Self Organizing Maps)	33
<b>3. DETERMINACIÓN DEL PROBLEMA</b>	<b>35</b>
3.1 Objetivo de la tesis	35
3.2 Definición del problema	35
<b>4. SOLUCIÓN PROPUESTA</b>	<b>39</b>
4.1 Solución Propuesta	39
4.2 Contexto general de la solución propuesta	39
4.2.1 ¿Qué riesgo se desea medir?	39
4.2.2 ¿Por qué medir tal riesgo?	40
4.2.3 Dinámica evolutiva del riesgo	41
<b>4.3 – 4.8 Metodología del proyecto</b>	<b>41</b>
<b>4.3 Entendimiento del negocio</b>	<b>42</b>
4.3.1.1 Contexto general	42
4.3.1.2 Contexto puntual del problema	42
4.3.1.3 Área de correspondencia del problema	43
4.3.1.4 Necesidades y expectativas de los usuarios	43
4.3.1.5 Ventajas y desventajas de la solución	43
4.3.1.6 Unidades de negocio	43
4.3.1.7 Recursos organizacionales claves para el proyecto	44
4.3.2 Objetivos del negocio y criterios de éxito	44
4.3.2.1 Objetivos del negocio	44
4.3.2.2 Criterios de Éxito	44
4.3.3 Recursos	44
4.3.4 Consideraciones	45
4.3.5 Riesgos y planes de contingencia	45
4.3.6 Glosario	46
4.3.7 Costos y beneficios	46
4.3.7.1 Beneficios del proyecto	46
4.3.7.2 Costos del proyecto	47
4.3.8 Objetivos y criterios de éxito de la minería de datos	47
4.3.8.1 Objetivos de la aplicación de minería de datos	47
4.3.8.2 Criterios de éxito de la aplicación de minería de datos	48
4.3.9 Etapas del proyecto	48
4.3.10 Herramientas a utilizar	49
<b>4.4 Entendimiento de los datos</b>	<b>49</b>
4.4.1 Extracción de los datos	49

4.4.2 Descripción de los datos	52
4.4.3 Exploración de los datos	55
4.4.4 Calidad y precisión de los datos	56
<b>4.5 Preparación de los datos</b>	<b>56</b>
4.5.1 Selección de los datos	56
4.5.1.1 Variables incluidas en la estratificación	57
4.5.1.2 Variables excluidas de la estratificación	59
4.5.2 Ajuste de formato de los datos	60
4.5.3 Construcción de los datos	60
4.5.4 Integración de los datos	60
4.5.5 Formato de los datos	61
<b>4.6 Modelización</b>	<b>62</b>
4.6.1 Selección de la técnica de modelizado	62
4.6.2 Diseño de experimentos	63
4.6.3 Descripción del modelo	63
4.6.3.1 Seteo de parámetros	63
4.6.3.2 Descripción del modelo	63
4.6.4 Interpretación del modelo	66
4.6.4.1 Interpretación y determinación del Riesgo Cluster	66
4.6.4.2 Interpretación del riesgo por Internación	70
4.6.4.3 Interpretación del riesgo de Sobrevida	75
4.6.4.4 Interpretación del riesgo por combinación de Enfermedad Crónica	75
4.6.4.5 Interpretación del riesgo por exceso y defecto de consultas	78
4.6.4.5.1 Criterio previo a la ponderación del riesgo por consultas	79
4.6.4.6 Interpretación del riesgo por cuidados preventivos	82
4.6.4.7 Interpretación del riesgo del Costo relativo para OMINT.	82
4.6.4.8 Conclusión sobre los riesgos ponderados	83
<b>4.7 Evaluación</b>	<b>84</b>
4.7.1 Evaluación de los resultados respecto de los criterios de éxito del negocio	84
4.7.2 Revisión del proceso de modelización	85
4.7.3 Pasos a seguir	86
4.7.3.1 Listado de posibles acciones	86
4.7.3.2 Decisión	86
<b>4.8 Entrega Final</b>	<b>87</b>
4.8.1 Plan de Entrega	87
4.8.1.1 Información a ser entregada	87
4.8.1.2 Informatización del Proyecto	88
4.8.1.3 Propagación de la información a los usuarios	88
4.8.1.4 Medición de los beneficios del proyecto	88
4.8.1.5 Problemas y contingencias en la entrega de resultados	89
4.8.2 Plan de monitoreo y mantenimiento	89
4.8.2.1 Aspectos dinámicos del entorno	89
4.8.2.2 Medición de la precisión de los resultados	91
4.8.2.3 Caducidad del proyecto	91
4.8.2.4 Plan de mantenimiento y chequeo	91
<b>5 CONCLUSIONES</b>	<b>92</b>
<b>6 FUTURAS LÍNEAS DE INVESTIGACIÓN</b>	<b>94</b>
<b>7 BIBLIOGRAFÍA</b>	<b>95</b>
<b>8 ANEXOS</b>	<b>98</b>
8.1 - Anexo 1.1: Query de información demográfica.	98
8.2 - Anexo 1.2: Query de información de consumo.	98
8.3 - Anexo 1.3: Query de información de gasto.	100

---

8.4 – Anexo 2: Identificación de cada variable categórica. _____	108
8.5 – Anexo 3.1: Características de cada Cluster _____	110
8.6 - Anexo 3.2: Ponderación de gravedad de cada Cluster _____	111
8.7 – Anexo 4: Puntuación de combinaciones de Enfermedades crónicas.	115
8.8 – Anexo 5: Scoring resultante. _____	117

# 1. INTRODUCCIÓN

El mundo se encuentra cada vez más globalizado. Las fronteras son cada vez más difusas y se vislumbra el nacimiento de una cultura global. Con esto, la cantidad de información se multiplicado a valores antes impensados. La disponibilidad de fuentes de conocimiento ha crecido principalmente gracias al desarrollo tecnológico para el manejo y tratamiento de la información, contribuyendo a que ésta se produzca en mayor cantidad, con mejor calidad y a bajo costo.

Sin embargo, todo este volumen de información generado crea mucho ruido, dificultando la búsqueda eficiente del conocimiento deseado. Se hace entonces necesaria una manera de clasificar, ordenar y catalogar la información eficientemente, de forma tal de poder encontrar el conocimiento deseado y poder extraer conclusiones útiles a partir de los datos.

Uno de los sectores económicos que hoy comienzan a requerir un uso adecuado de los sistemas de información es el sector de la medicina prepaga. Ciertamente, estas empresas son uno de los sectores del mercado de la salud que hoy en día encuentran un futuro incierto. En los últimos años, unas pocas compañías líderes del sector han entendido las nuevas reglas del mercado y han reposicionado su estrategia, apuntando sus recursos hacia programas de “Disease Management”, dejando de lado la vieja práctica de considerar al servicio prepago como un simple abono, para pasar a analizarlo como un seguro de riesgo propiamente dicho [DataBank Group, 2005]. Dentro de este cambio de esquema surge inevitablemente la necesidad de profundizar el conocimiento de la población, a fin de concentrar recursos en aquellos grupos de alto costo, conteniendo la situación actual, y previniendo el futuro desarrollo de población no apta en términos económicos [Moroney, 2003].

La nueva estrategia de “Disease Management” supone un cambio radical en la perspectiva sobre el cliente. El viejo paradigma de limitar la utilización o lograr mejores acuerdos de precio con los prestadores médicos dejó ser prioridad. El nuevo paradigma establece que el verdadero beneficio radica en la salud del afiliado. Estos programas funcionan a partir de un seguimiento exhaustivo de los afiliados enfermos, donde se evalúa periódicamente su estado de salud, y ante la presencia de desvíos o anomalías sobre los patrones deseables, se trabaja junto con el afiliado para que este mejore su estado de salud y calidad de vida. Si goza de buena salud, o bien si a este logra detectársele potenciales enfermedades a tiempo, el impacto económico para la empresa será menor.

En este sentido, el conocimiento sobre el consumo y el estado de salud del afiliado es necesario, y debe ser analizado a fin de focalizar recursos eficazmente sobre aquellos grupos de mayor riesgo potencial. Sin embargo, si bien este

cambio de estrategia ha generado que la recolección y utilización de la información sea un tema prioritario, los esfuerzos iniciales se vieron limitados principalmente por la ineficiente explotación de las fuentes de información, la poca calidad de los datos, y el desconocimiento sobre como abordar su utilización.



## 2. ESTADO DE LA CUESTIÓN

*En este capítulo se describen los paradigmas del mercado de la salud actual (Inciso 2.1); se describen las amenazas y oportunidades del mercado (inciso 2.2); se describe el Estado de los Sistemas de flujo de información en la Argentina y en los EEUU (Inciso 2.5 a 2.6); explotación de la información para estrategias de Disease Management (inciso 2.7); y herramientas de reconocimiento de patrones aplicadas en proyectos similares (inciso 2.8).*

### 2.1 Paradigmas del mercado actual

En los últimos años la industria de la medicina prepaga en la Argentina ha estado tratando de responder una pregunta: ¿Es el sistema estable a largo plazo?

La respuesta aún parece incierta para muchas empresas sector, aunque sin lugar a dudas todas coinciden en un contexto marcado por un vertiginoso aumento del costo médico, objeto del desarrollo de nuevas tecnologías cada vez más complejas, así como la creciente innovación en tratamientos tendientes a mejorar la expectativa de vida de sujetos enfermos [Wong *et al*, 2005].

En este sentido, las perspectivas a futuro del mercado pueden delimitarse en tres aspectos indicativos a saber: la tendencia demográfica, la tecnología y las expectativas de los afiliados.

A continuación se detallará el impacto de cada uno de ellos en el mercado, para luego justificar la razón de ser del cambio de esquema en el negocio de la medicina prepaga.

#### 2.1.1 Tendencia demográfica

Las tendencias demográficas – edad, sexo y crecimiento – impactan directamente sobre el consumo y utilización de servicios relacionados a la salud.

Es sabido que el avance de la medicina ha sido uno de los factores que más ha influido en que la expectativa de vida de los individuos sea cada vez mayor. Nuevas drogas que curan enfermedades antes desconocidas; nuevas terapias para el tratamiento del cáncer; nuevas cirugías para el transplante de órganos; y tantas otras, han generado individuo tenga una alta expectativa de vida, y una alta

probabilidad de que su enfermedad sea detectada, y que algún tratamiento sea aplicable para sanarla o bien vivir con mayor calidad con ella. Sin embargo, este inobjetable avance atenta cada vez más contra la financiación de la salud. En este sentido, en el pasado la gente moría por causas “naturales”, cuando en realidad eran algún tipo de enfermedad, para la cual hoy tendremos un nombre y un tratamiento disponible. Lógicamente el impacto actual del conocimiento de causa, genera que las empresas de medicina gasten más en sus afiliados, y estos a su vez, demanden mejores tratamientos por el servicio que abonan.

Así mismo, el progresivo envejecimiento de la población junto con la mayor expectativa de vida es otro de los grandes problemas a enfrentar. El gráfico contiguo (Figura 2.1.1) [Interdepartmental Committee on Aging and Seniors Issues, 2001] muestra la evolución futura de los grupos mayores a 65 años sobre el total de la población para los años venideros.

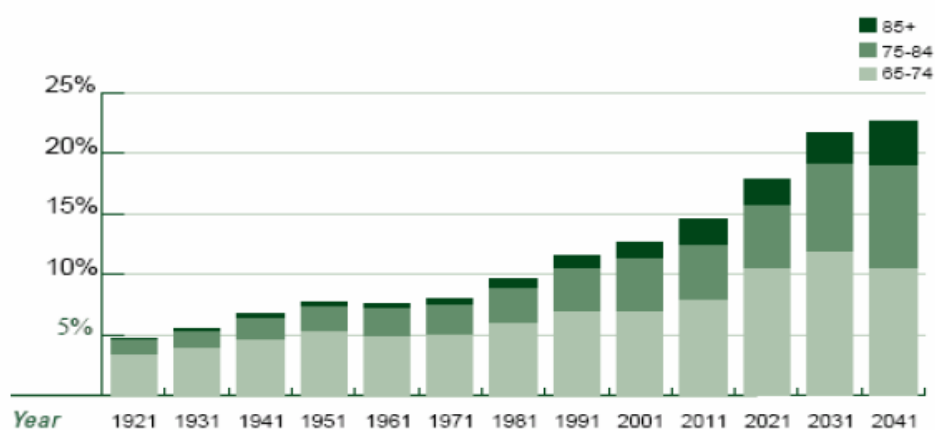


Figura 2.1.1 – Evolución futura de los grupos mayores a 65 años.

En términos económicos, el gasto en salud para una persona mayor de 65 años es 5 veces el de una persona de 20 años, y sigue aumentando año a año. Queda claro que si la población de alta edad aumenta, el gasto será cada vez mayor.

Otro de los aspectos no menores son las enfermedades crónicas. El 20% de la población posee al menos una enfermedad crónica. Si bien los números hablan por sí solos, es necesario remarcar que sólo este grupo gasta el 41% de los recursos del sistema, y lo hará mientras permanezca afiliado. El gasto de un individuo crónico tiene un piso fijo de por vida, y el variable a medida que progresa su enfermedad tiene una pendiente más pronunciada que el resto de los enfermos.

Por último otro factor demográfico de relativo impacto es el sexo. Por un lado el sexo femenino tiene una expectativa de vida más larga que el hombre, pero más importante aún es el hecho que es más susceptible a desarrollar enfermedades crónicas, y por consiguiente a utilizar más recursos.

### 2.1.2 Tecnología

El desarrollo de nuevas tecnologías para el tratamiento y detección de enfermedades impacta en los costos de la salud. Las nuevas tecnologías, por el tipo de desarrollo que tienen, presentan un costo superior a sus antecesoras. Según estudios extranjeros, el impacto de una nueva tecnología es del 25% del total del aumento del gasto de un año a otro [Whellan *et al*, 2005].

Respecto a los tipos de tecnologías innovadoras, no sólo se presentan en el campo de la maquinaria, sino también en lo que respecta la industria farmacéutica. El desarrollo de nuevas drogas han sido en los últimos 25 años el principal componente de gasto para las empresas de medicina prepaga [Dickson and Jacobzone, 2003]. Entre 1990 y 2000 el gasto per cápita en medicación aumentó un 93%, mucho más que el 40% de aumento promedio del resto de los gastos [CIHI, 2003]. Este aumento se debe principalmente al desarrollo de nuevas drogas en reemplazo de las preexistentes. Su razón de ser es objeto del costo de investigación y desarrollo de las mismas.

Así mismo, en los últimos años, el desarrollo de drogas se ha expandido a tratar enfermedades que hace 10 años no era posible, tales como HIV, osteoporosis, problemas de ansiedad, y otras tantas.

### 2.1.3 Expectativa de los afiliados

La demanda por parte de los afiliados ha cambiado notablemente. Sus demandas, variables según su contexto – edad, sexo, poder adquisitivo, educación, estilo de vida, expectativas, y plan contratado – ya no son tan influenciadas por los profesionales de la salud, sino por la información que ellos mismos obtienen de otros medios [IBM Business Consulting Services, 2003]. Así mismo se evidencia un incremento en el sentido de responsabilidad en cuanto a la prevención de enfermedades [PricewaterhouseCoopers, 2002].

Las demandas de los afiliados no sólo vienen de la mano de factores demográficos tales como la edad y el sexo, sino también objeto de la mayor demanda de cuidado que exigen al momento del tratamiento, especialmente aquellos que padecen enfermedades crónicas. Tales demandas se traducen en exigencias para las compañías de medicina prepaga, las cuales deben juzgar desde el punto de vista económico y el beneficio para el afiliado, sus demandas potenciales. Debe indicarse que este tipo de compañías es la 5ta industria con más reclamos en la Argentina.

## 2.2 Amenazas y Oportunidades del mercado

En la siguiente matriz (matriz 2.2) se listan las amenazas y oportunidades del contexto macroeconómico de la medicina prepaga en la Argentina.

Impacto	Probabilidad	Corto Plazo		Mediano Y Largo Plazo	
		Amenazas	Oportunidades	Amenazas	Oportunidades
ALTO	ALTA	Imposibilidad de trasladar incrementos de costos a los precios	Utilización estratégica de las fuentes de información sobre comportamientos de los afiliados	Saturación de la oferta del sector	Utilización estratégica de las fuentes de información sobre comportamientos de los afiliados
	MEDIA	Mayores controles de precios	Reducción de cargas tributarias	Obsolescencia Tecnológica	Aumento del nivel de ingresos de la población
	BAJA			Arribo de nuevos competidores	Incremento de la integración vertical
MEDIO	ALTA		Incremento de los acuerdos con las obras sociales Expansión al interior del país	Envejecimiento de la población	
	MEDIA	Nuevo marco normativo		Incremento desigualdad social y consecuente incremento en la rivalidad competitiva	Mejora de la condición sanitaria de la población
	BAJA				Deficiencias en salud pública
BAJO	ALTA	Incremento de riesgos potenciales	Menor presión de los sustitutos		Inequidad en distribución geográfica de servicios
	MEDIA				Exportación de servicios
	BAJA				

Matriz 2.2 – Amenazas y Oportunidades del Mercado de la Salud

## **2.2.1 Amenazas a corto plazo**

### **2.2.1.1 Imposibilidad de trasladar totalmente los aumentos de costos a las cuotas.**

Como fue noticia durante el 2006, los incrementos de costos no han podido trasladarse a los precios. La fuerte resistencia gubernamental para frenar los aumentos de las cuotas ha hecho que las empresas se enfrenten ante un escenario en el cual los costos continúan incrementándose pero no pueden ser trasladados a los precios al afiliado. De más está decir que el impacto que ello tendría frente a una desregulación estatal haría resentir la cantidad de afiliados.

### **2.2.1.2 Incremento de los riesgos potenciales en la prestación de servicios**

Uno de cada cinco médicos enfrenta demandas por mala praxis ante la Justicia. Se teme que las denuncias contra quienes ejercen la medicina se multipliquen por cinco a lo largo de esta década, en muchos casos impulsadas por una "industria del juicio" que aprovecha la posibilidad de litigar sin gastos y puede así hacer colapsar el sistema sanitario. El incremento de demandas provocó que el 65 por ciento de los médicos, en forma particular, debieran contratar seguros cada vez más onerosos por el aumento de la siniestralidad. De persistir esta tendencia, la cadena de valor del negocio de las prepagas puede resentirse al incrementarse los costos directamente (primas) e indirectamente (atención médica defensiva) [DataBank Group, 2005].

### **2.2.1.3 Implementación de un nuevo marco legal para el sector**

Una eventual normativa que imponga un régimen de capitales mínimos, el registro obligatorio en la SSS, nivelación del régimen impositivo, etc, cambiaría la dinámica sectorial en forma notable. En líneas generales, las reglas tienden a favorecer a las medianas y grandes prepagas y buscan proteger al usuario. Como contrapartida, se espera que desaparezcan las empresas fantasma y la competencia desleal. Pero también existe el riesgo concreto que el nuevo marco legal beneficie a algunos actores en detrimento de otros a título discrecional.

### **2.2.1.4 Mayores controles de precios**

En un contexto de un año electoral en el cual la inflación ha pasado a ser uno de los temas relevantes, es muy probable que el sector deba experimentar mayores controles y presiones en su operatoria y fundamentalmente, en lo que se refiere a los precios en los planes. Sin embargo, es muy previsible que tales presiones sean muy intensas para los planes más económicos, pero no tanto para los planes que tengan como clientes primordiales sectores de altos ingresos.

## **2.2.2 Oportunidades a corto plazo**

### **2.2.2.1 Incremento de los acuerdos con las obras sociales**

La cantidad de convenios desarrollados entre las prepagas y las obras sociales es, aún, relativamente escaso. Asimismo, las obras sociales que más beneficiarios convocan no han actuado en esta línea lo que abre un escenario positivo para el crecimiento de las cápitas en la medicina prepaga.

### **2.2.2.2 Obtención de una reducción de la carga tributaria**

Si se llega a un contexto en el cual para el Estado sea una preocupación el mantenimiento de los precios de los planes de prepagas, es factible que el gobierno nacional y provinciales se encuentren dispuestos a reducir la presión tributaria sobre el sector como contraprestación en una eventual negociación. En este sentido, el 2007 por ser un año electoral, es particularmente interesante para un escenario de tales características [DataBank Group, 2005].

### **2.2.2.3 Expansión al interior del país**

Los líderes están decididos a avanzar hacia el interior del país debido a que una cobertura nacional les permite acceder a los grandes clientes corporativos. Para ello, en un principio, acuerdan con diferentes redes hasta obtener economías de escala suficientes como para establecerse en forma directa.

Sin embargo, el arribo a cada plaza requiere de convenios con diferentes colegios médicos que, en ocasiones, operan con pequeñas prepagas del interior del país. En consecuencia, necesitarán absorber o anular a estas pequeñas prepagas para poder operar en las ciudades del interior del país.

#### **2.2.2.4 Menor presión de los sustitutos**

Un escenario macroeconómico más estable predispone favorablemente a los individuos para un mayor consumo, entre ellos la salud, aún cuando las condiciones reales no hayan cambiado sustancialmente.

Esta situación podría llevar a algunos beneficiarios a cambiar de plan por uno de mayores prestaciones, o bien podría trasladar la demanda de los seguros de salud a los planes más bajos, o incluso, podría recuperar parte de las cápitas que salieron del sistema privado para pasar al público durante la crisis.

En consecuencia, la presión de estos sustitutos podría disminuir, aunque siempre relacionada con el segmento de particulares ya que el segmento corporativo no utiliza estos sustitutos.

### **2.2.3 Amenazas a mediano y largo plazo**

#### **2.2.3.1 Saturación del sector**

Si se mantienen las condiciones y la evolución sectorial de los últimos años, el sector puede llegar a saturarse puesto que sólo sobrevivirán algunos actores estratégicos que puedan hacer frente a una competencia muy intensa en un contexto de rentabilidades bajas.

Lo más previsible es que sean los líderes quienes logren hacerlo, con lo cual un número importante de seguidores y prácticamente todos los oportunistas se vean obligados a realizar acuerdos o ventas a las empresas mejor posicionadas.

#### **2.2.3.2 Incremento de la desigualdad social y consecuente intensificación de la competencia por un mercado limitado**

El sistema de salud Argentino ha tenido un gasto de 32.000 millones de pesos en 2005. Sin embargo, a pesar del incremento en valores absolutos de las erogaciones, la desigualdad de acceso a los servicios de salud es muy notoria. La desocupación y el empleo en negro, agravaron las dificultades de los argentinos para acceder a una mínima cobertura sanitaria. La implicancia de corto plazo para el sector de las prepagas es que la intensidad de la competencia por los segmentos más interesantes del mercado puede incrementarse, dado que aún cuando las condiciones macroeconómicas sean buenas, una importante proporción de la población puede permanecer fuera del espectro de acción de las prepagas.

### **2.2.3.3 Obsolescencia de la Tecnología**

El sector de la salud tiene asociada una inflación que le es propia y que supera la media de cualquier estado. La razón fundamental de esta inflación es la incorporación permanente y sistemática de tecnología médica a través de medicamentos, tratamientos, equipos médicos, etc. que incrementan los costos médicos y vuelven obsoletas muy rápidamente las versiones anteriores.

En efecto, la búsqueda del control de las enfermedades y tratamientos menos cruentos y más eficaces a través de la investigación se vuelca cada vez más rápidamente a nivel local debido a la facilidad de las comunicaciones modernas y la colaboración entre los diferentes equipos médicos entre países.

### **2.2.3.4 Arribo de nuevos competidores**

Debido a la madurez del sector, la sobreoferta en algunos servicios y la moderada rentabilidad, es poco probable que nuevos competidores decidan ingresar en el sector de forma tan notoria que modifiquen las reglas de la competencia. Esto dependerá de la importancia del grupo inversor si ésta competirá con los líderes o seguidores o con una estrategia oportunista, o bien puede hacerlo a través de la adquisición o de una alianza con los actuales operadores [DataBank Group, 2005]..

### **2.2.3.5 Envejecimiento de la población**

Tal como fue abordado oportunamente, la población mundial envejece sin remedio ya que la esperanza de vida crece y los índices de nacimiento caen. Por otra parte, la población mayor requiere de cuidados más caros lo que implica un costo médico superior para su atención.

Esta situación genera un sostenido aumento en los costos que, de no poder trasladarse a los precios o compensarse con una mayor eficiencia en la operación, atentaría contra la rentabilidad del negocio.

## **2.2.4 Oportunidades a mediano y largo plazo**

### **2.2.4.1 Deficiencias de la salud pública**

Dado que gran parte de los hospitales atraviesa severos problemas de personal y provisión de insumos básicos, y además los horarios de atención por parte de los



médicos no superan en promedio las cuatro o cinco horas diarias, insuficientes para hacer frente a las demandas de una superpoblación de pacientes, si las empresas de medicina prepaga pueden ofrecer servicios que suplan a las funciones que el Estado debe proveer, el mercado potencial es inmenso.

#### ***2.2.4.2 Inequidad en la distribución geográfica de la prestación de servicios***

La difusión de los servicios de salud es notoriamente desigual en las diferentes regiones del país. Si bien a nivel nacional se destinan cerca de 200 pesos anuales per cápita promedio, en la Capital la asignación anual por persona alcanza los 380 pesos, mientras que en distritos provinciales como Buenos Aires o Tucumán se dispone sólo de 160 pesos para una población más empobrecida y con mayores necesidades. Esta coyuntura puede ser una oportunidad para las propuestas que logren posicionarse como un sustituto adecuado y accesible a sectores de limitados recursos en el interior, actualmente muy marginados del acceso a la medicina prepaga.

#### ***2.2.4.3 Aumento del nivel de ingresos de la población***

Este punto está emparentado con el anterior en el sentido que, si el nivel de ingresos de la población aumenta, los desempleados podrán acceder a una prepaga en forma individual y los beneficiarios de obras sociales podrían, como ya ha ocurrido con anterioridad a la crisis, contar con una doble cobertura para lograr beneficios adicionales como el acceso a ciertos prestadores o mayor comodidad en el servicio.

#### ***2.2.4.4 Incremento de la integración vertical***

El aumento del grado de concentración brinda a los líderes la posibilidad de continuar integrándose verticalmente como una herramienta para obtener una mejor posición en costos.

Cabe destacar que esta integración vertical no necesariamente tiene que realizarse a través de la compra de un prestador, sino a través de alianzas con los mismos, especialmente con los prestigiosos que agregan valor a la cartilla.

#### **2.2.4.5 Incremento de la sustitución de las importaciones de los insumos médicos**

Si bien algunos insumos no pueden ser sustituidos localmente porque el país carece de la tecnología y la economía de escala necesarias, todavía existen numerosos insumos que pueden, inversión mediante, ser producidos en Argentina [DataBank Group, 2005].

La incorporación de insumos locales reduce los costos totales y evita la dependencia de la evolución del tipo de cambio.

#### **2.2.4.6 Utilización de las fuentes de información sobre comportamientos de los afiliados**

Uno de los aspectos menos claros para las compañías de medicina prepaga, pero con mayor impacto en el negocio es la utilización de la información sobre los patrones de consumo de sus afiliados, y así como el estado de salud que estos presentan.

Tal información aún no presenta grandes avances en cuanto a su facilidad de obtención, pero el impacto económico al implementar estrategias de negocio, tales como programas de prevención de enfermedades o programas de disease management, ha sido efectivamente constatado en otras partes del mundo, en particular en los EEUU.

### **2.3 Conclusiones sobre las oportunidades y amenazas del mercado**

El mercado de la salud encuentra un escenario sumamente difícil. Por un lado la regulación estatal impide el traslado de los aumentos a los afiliados; los afiliados por su cuenta demandan cada día un mejor servicio; y los servicios cada día, si bien son más eficientes en términos de beneficios para la salud, son cada vez más caros. En este contexto, la rentabilidad de las empresas se ve cada vez más acotada, y por ello es necesario que se busquen nuevas estrategias de negocio.

Una de las posibles estrategias de negocio será abordada por el siguiente proyecto de tesis, donde se evaluará una novedosa estrategia basada en los sistemas de información de las compañías de medicina prepaga. Hacia fines del 2006 en la Argentina sólo se encuentran dos compañías que han incurrido por este camino, y sobre una de ellas, OMINT S.A. se realizará el presente proyecto basado en la aplicación de redes neuronales para la detección de patrones de

consumo e inferencia del estado clínico en los afiliados, a fin de encausar objetivamente los recursos en los programas de “Disease Management” de la compañía.

Será preciso por consiguiente conceptuar en primer lugar cómo está concebido el flujo de información entre la empresa y el afiliado, para luego entender cuales son las ventajas y potenciales limitaciones del proyecto, y que el lector pueda comprender el fundamento práctico de este tipo de estrategia.

## **2.4 Estado de los sistemas de flujo de Información**

La recolección de información para el mercado de salud en la Argentina es un tema muy discutido en cuanto a su legalidad. Si bien las compañías no infringen ninguna ley al recolectar “datos”, no pueden utilizar los mismos con fines que determinen la antiselección o indeseabilidad de un cliente. Sin embargo, sí es válido que la utilicen con fines de análisis de su negocio, y en aquellos casos que impliquen un beneficio para el afiliado.

El estado de la tecnología para recolectar este tipo de información en la Argentina se encuentra aún en una etapa de desarrollo, y es muy probable que se mantenga así por muchos años. Si se realiza una comparación con los sistemas Norteamericanos, la brecha es abismal. No sólo por el tipo de información que logran recolectar, sino por la calidad de la misma.

La recolección de la información clínica es una tarea compleja. Las fuentes de información para detectar el consumo médico de un afiliado agrupan a ocho tipos distintos de prestadores, a saber:

- Médicos particulares
- Sanatorios y Centros
- Laboratorios
- Farmacias
- Centros de Odontología
- Vacunatorios
- Centros de Imágenes
- Servicios de HomeCare
- Servicios de Emergencia
- Servicios de Salud Mental

Lógicamente el flujo de información de cada uno de estos prestadores tiene sus particularidades y es interesante comparar como es el flujo de la misma en la Argentina contra el mejor sistema actual, aplicado en EEUU.

## 2.5 Flujo de Información en la Argentina

Los siguientes cuadros representan el flujo de información desde el afiliado hasta la compañía de medicina prepaga.

### 2.5.1 Flujo de información de consultas médicas en la Argentina

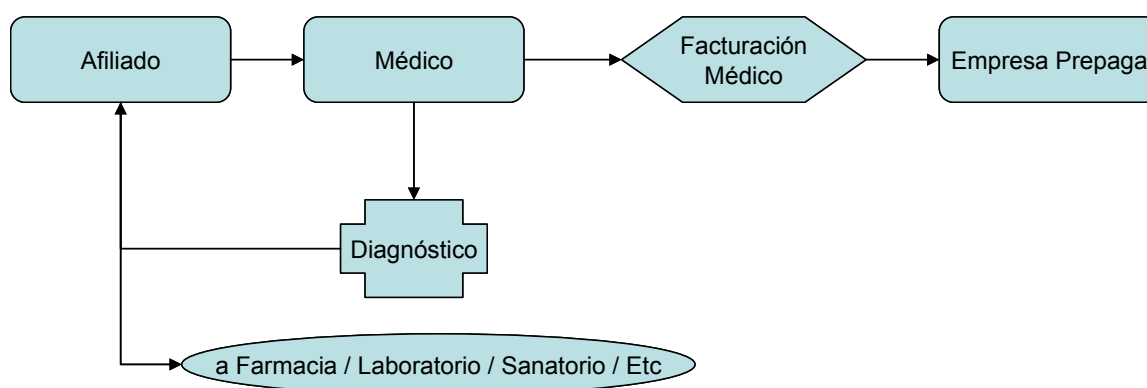


Figura 2.5.1 – Flujo de Información de consultas médicas en la Argentina

Como puede observarse (figura 2.5.1), cuando el afiliado realiza una visita en consultorio a algún médico, la empresa de medicina prepaga recibe únicamente el dato de la visita médica al momento de pagar la facturación. Tal dato cuenta solamente con el motivo de la visita (“Consulta en consultorio”) y fecha de la misma, por lo cual únicamente se puede saber a que tipo de especialista visitó el afiliado y cuando lo hizo. Pero nada puede saber el estado clínico del socio o la enfermedad que pueda llegar a tener, ya que el diagnóstico es información confidencia entre el médico y su paciente. Así mismo, el delay de la información es de aproximadamente 3 meses.

## 2.5.2 Flujo de información Sanatorial en la Argentina

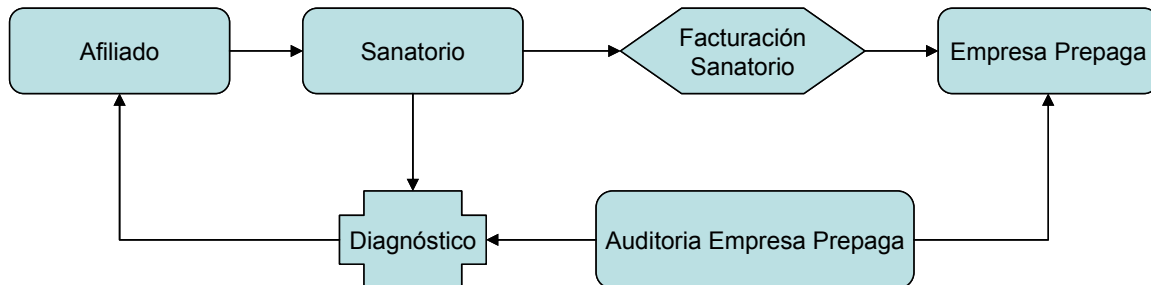


Figura 2.5.2: Flujo de Información Sanatorial en la Argentina

Con los sanatorios la información recibida es algo más clara (ver figura 2.5.2). La empresa puede enterarse de la internación tanto por la facturación enviada por el sanatorio, como por el sector de auditoría de internaciones que diariamente censa a los sanatorios recibiendo información sobre la causa primera de la internación. Sin embargo, no recibe información sobre que enfermedades o sucesos se le encontraron al paciente, medicamentos consumidos y aplicados durante la internación, y estado de salud posterior a la alta. El delay de la información es de 3 meses para la facturación y diaria para el censo de internación.

## 2.5.3 Flujo de información de Farmacias en la Argentina

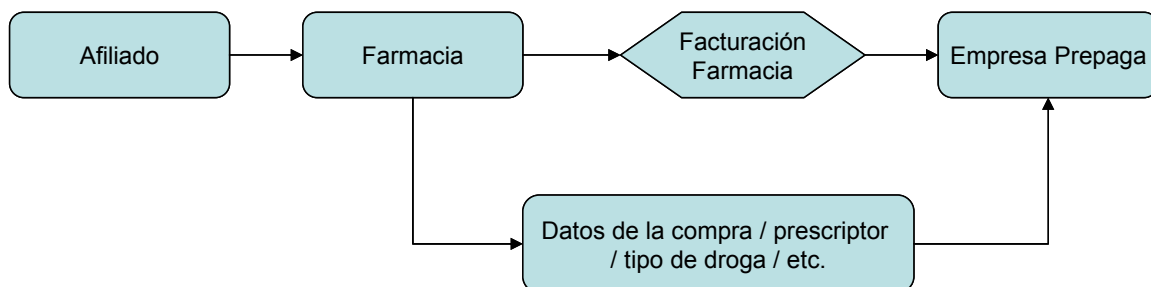


Figura 2.5.3: Flujo de Información de Farmacias en la Argentina

Respecto a la utilización de farmacias (ver figura 2.5.3), la información recibida es perfecta. El servicio se encuentra instalado bajo un sistema de POS nacional, linkado con una base común de medicamentos provista por la red de laboratorios, donde se complementan los datos del producto adquirido. La empresa recibe datos sobre el tipo de medicamento adquirido, la acción

farmacológica del mismo, el prescriptor de la misma, la cantidad adquirida y hasta el lote del producto. En cuando al delay, este es de un mes.

## 2.5.4 Flujo de información de Laboratorios en la Argentina

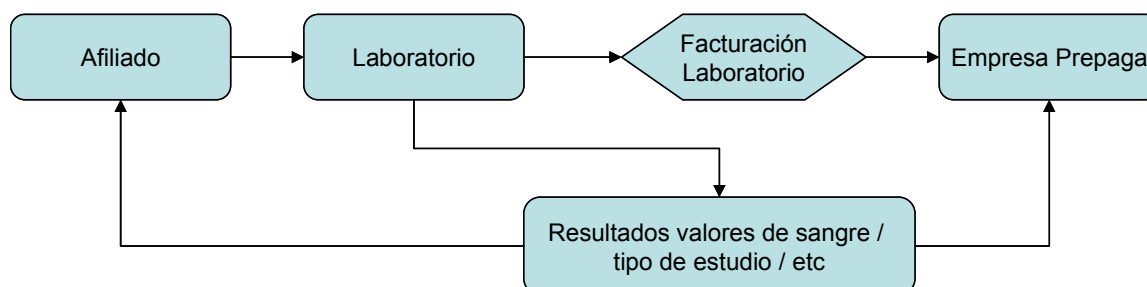


Figura 2.5.4 – Flujo de Información de Laboratorios en la Argentina

La información de laboratorios es bastante completa dependiendo del proveedor (ver figura 2.5.4). El 70% del mercado de análisis clínicos se encuentra agrupado bajo una misma red que está en condiciones recolectar información sobre los resultados de los análisis de los afiliados (tipo de estudio realizado, resultados de cada estudio y prescriptor), por lo cual tal información puede ser adquirida por las empresas de medicina prepaga. Sin embargo, la calidad de los datos es aún bastante baja en cuando a inconsistencias de la data. El delay es de 3 a 6 meses.

## 2.5.5 Flujo de información de los Centros de Imágenes en la Argentina

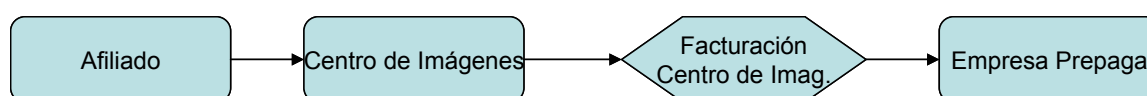


Figura 2.5.5 – Flujo de Información de Centros de Imágenes en Argentina

Respecto a los centros de imágenes (figura 2.5.5), estos proveen únicamente información relevante al tipo de estudio realizado y el médico prescriptor de la orden, El delay de la información es de 3 meses.

## 2.5.6 Flujo de información de Vacunatorios en la Argentina

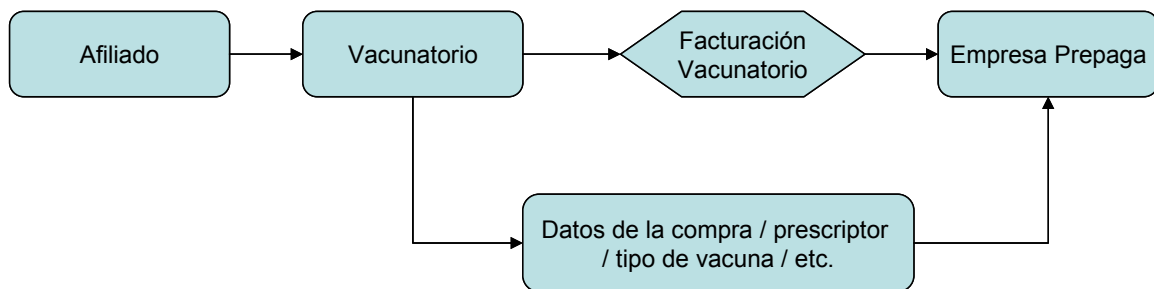


Figura 2.5.6 – Flujo de Información de Vacunatorios en Argentina

Los centros vacunatorios se encuentran en una fase piloto de recolección de datos. Solamente el 60% de los mismos ha logrado recolectar información precisa y de utilidad para las empresas de medicina prepaga. El delay actual de la información es de 2 meses (ver figura 2.5.6).

## 2.5.7 Flujo de información de servicios de Home Care en la Argentina

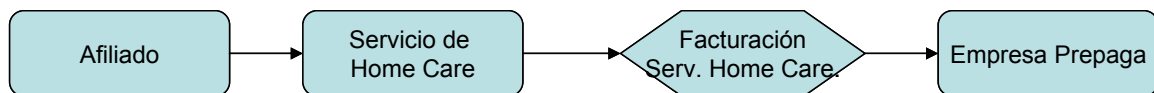


Figura 2.5.7 – Flujo de Información de servicios de Home Care en Argentina

El afiliado que solicita servicios de Home Care es aquel que debido a su condición física debe estar bajo supervisión, pero su gravedad no es tal que implique continuar internado en un centro sanatorial.

La empresa de medicina prepaga recibe únicamente información sobre el servicio prestado con la facturación, cuyo delay es de 3 meses (ver figura 2.5.7).

## 2.5.8 Flujo de información de Servicios de Emergencia en la Argentina

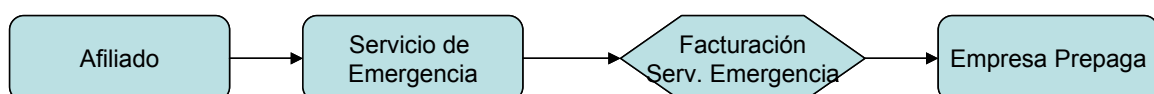


Figura 2.5.8 – Flujo de información de Servicios de Emergencia en Argentina

El servicio de emergencias supone al servicio de ambulancias o servicios asistenciales domiciliarios.

La empresa prepaga sólo recibe información del servicio prestado con la facturación. El delay de la misma es de 3 meses (ver figura 2.5.8).

### 2.5.9 Flujo de información de Salud Mental en la Argentina

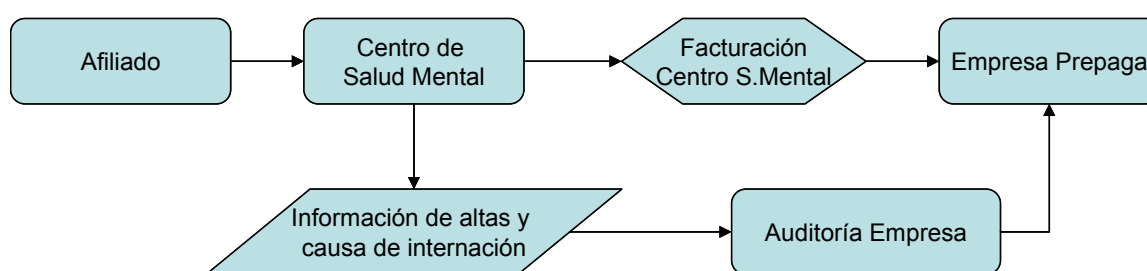


Figura 2.5.9 – Flujo de Información de Salud Mental en Argentina

Las internaciones en los centros de salud mental son censados diariamente por la empresa de medicina prepaga, y mensualmente desde el centro de internación se le envía información referente a la causa de ingreso de internación y las altas y bajas del mes. Sin embargo, la información no tiene reglas de codificación, y no presenta información sobre la evolución, diagnósticos intermedios del paciente y medicación administrada. El delay de la información es mensual (ver figura 2.5.9).

### 2.5.10 Conclusiones sobre el flujo de Información en la Argentina

Si se analiza el flujo en su conjunto, existe un dato clave al cual las empresas de medicina prepaga en la actualidad no pueden acceder, y ello es el diagnóstico clínico. Es claro que sin tal información, el estado de salud del socio es desconocido. Tal falencia hace muy dificultosa una tarea de análisis para establecer cuál es la situación puntual de cada afiliado, no pudiendo por consiguiente establecer acciones de contención tanto para el beneficio del afiliado como de la empresa. Por ello mismo las estrategias de implementación de programas de disease management se ven limitadas en sus resultados. Si bien queda evidenciado que relacionando las distintas fuentes de información se puede llegar a inferir el estado de salud de un afiliado, debido al delay en la obtención de la información, tal estado será tardío. Sin embargo el problema de ello radica básicamente en cuál será la información relevante a utilizar, ya que como se ha comentado oportunamente, la calidad del dato de algunas fuentes es baja. De todos modos, y tal como se abordará más adelante, la posibilidad de



utilizar dicha data es posible, y las conclusiones que pueden obtenerse son valederas, siendo una mejor aproximación frente a una ausencia completa de información.

## 2.6 Flujo de Información en los Estados Unidos

La diferencia principal del flujo de información se encuentra en que el Estado Nacional exige la codificación de los diagnósticos y procedimientos clínicos y sanatoriales. Esta codificación se viene utilizando de la década del 80, y según el tipo de práctica existen códigos denominados ICD-9, DRG y HCPAC donde no solo se encuentra registrada el tipo de enfermedad, sino sus variantes, diagnósticos sobre cada estudio, procedimientos realizados, procedimientos en internación, administraciones dentro de la internación, y otras tantos efectuados por cada actor del círculo de atención médica. La informatización del sistema fue la clave para su éxito.

Los siguientes cuadros representan el flujo de información desde el afiliado hasta la compañía de medicina prepaga.

### 2.6.1 Flujo de Información de consultas médicas en EEUU

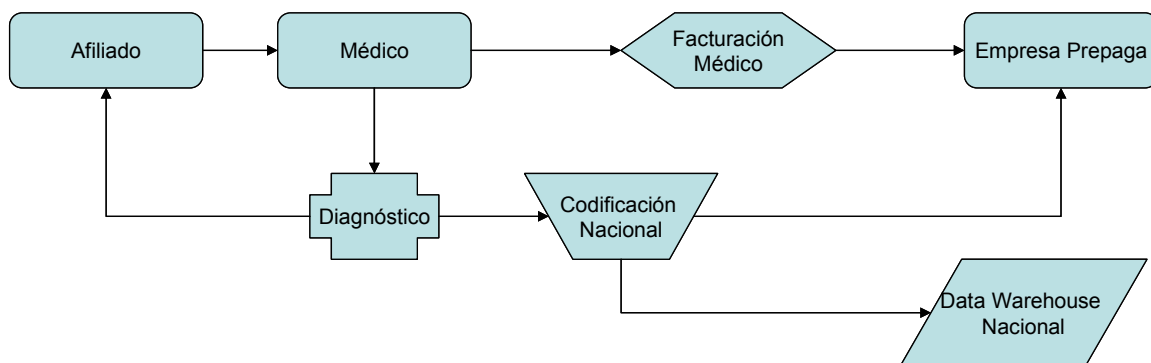


Figura 2.6.1 – Flujo de Información de consultas médicas en EEUU.

Al realizar el afiliado una visita a un consultorio externo, tanto la empresa de medicina prepaga como el mismo Estado reciben la información online de la visita médica, y el diagnóstico establecido según la codificación (ver figura 2.6.1). Es decir, en el instante se conoce el estado de salud del afiliado y la causa de la visita. Así mismo, el tema de la codificación no es menor. La categorización del diagnóstico es común a nivel nacional, por lo cual no error de interpretación para una misma causa.

## 2.6.2 Flujo de Información Sanatorial en EEUU

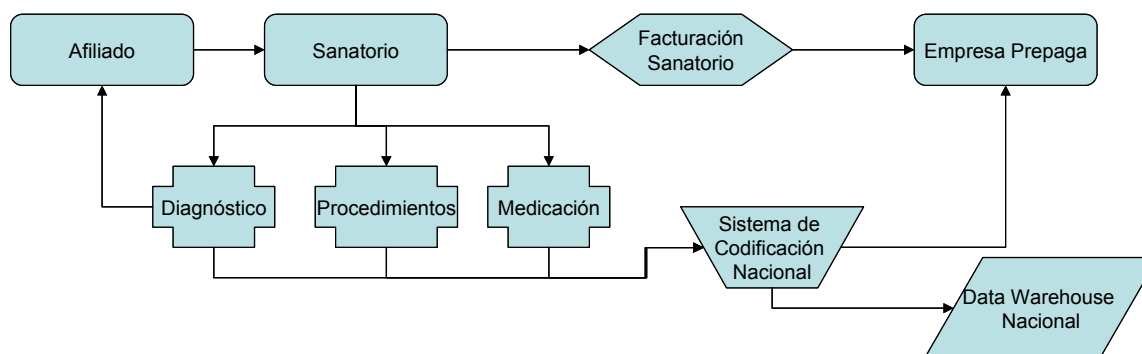


Figura 2.6.2 – Flujo de Información sanatorial en EEUU.

En el caso de las internaciones o visitas sanatoriales, se registran los diagnósticos, procedimientos y medicaciones utilizadas de acuerdo a la codificación correspondiente. Tal información es codificada y enviada al sistema (ver figura 2.6.2). Debemos hacer mención en este punto la fuerte inversión tecnológica que se requirió para llevar a cabo esta recolección. En todos los centros sanatoriales se instalaron cientos de terminales y se contrataron auxiliares administrativos que ingresen los códigos correspondientes ante cada procedimiento / diagnóstico / medicación aplicada al paciente. El costo de mantener tal sistema es sumamente elevado.

## 2.6.3 Flujo de Información de Farmacia en EEUU

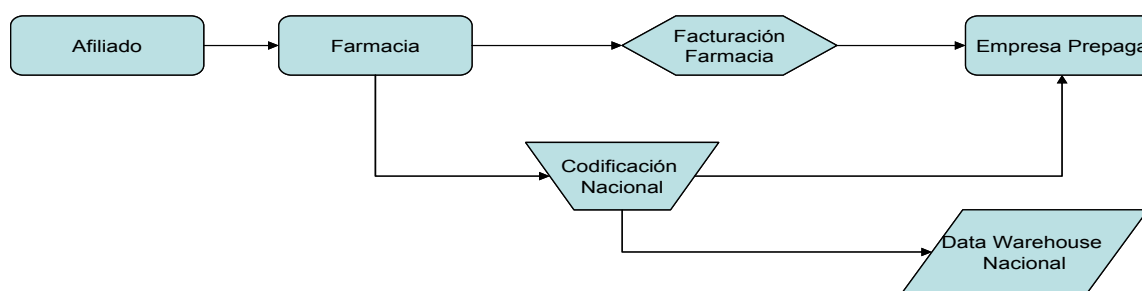


Figura 2.6.3 – Flujo de Información de farmacia en EEUU.

El sistema de información aplicado es similar al utilizado en nuestro país. Sin embargo presenta la diferencia de estar codificada la utilización de la medicación para cada tipo de enfermedad (ver figura 2.6.3).

## 2.6.4 Flujo de Información de Laboratorio en EEUU

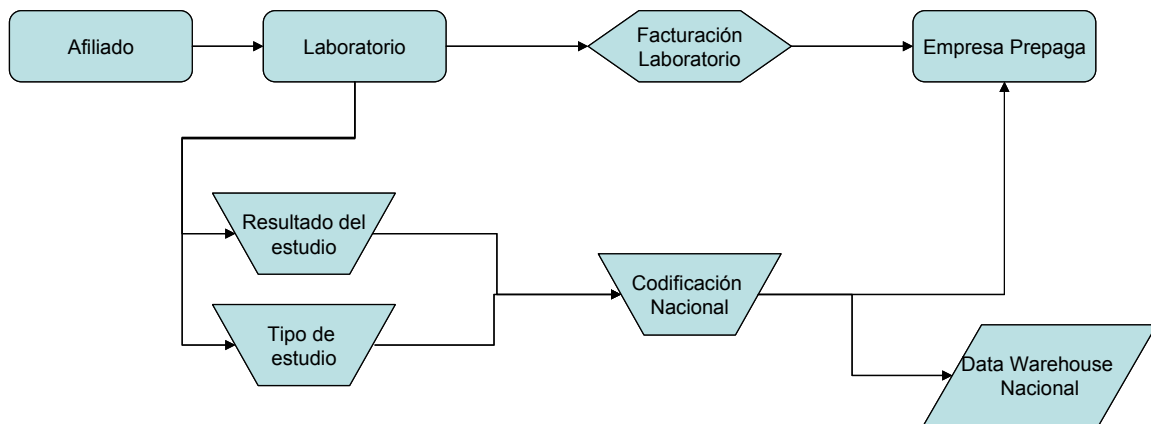
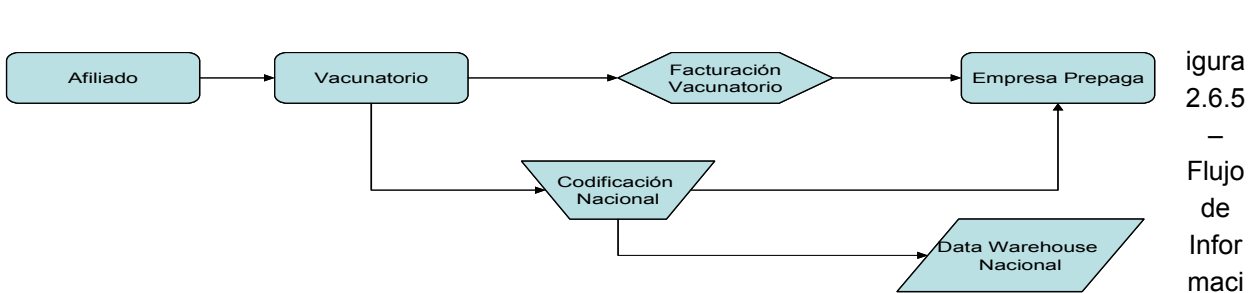


Figura 2.6.4 – Flujo de Información de Laboratorio en EEUU.

Del mismo modo que el sistema de información aplicado en la Argentina, el laboratorio provee la información sobre el resultado y el tipo de estudio. Sin embargo, la diferencia radica nuevamente en la precisa codificación que se le aplica a la misma, así como la codificación del diagnóstico del técnico de laboratorio (ver figura 2.6.4).

## 2.6.5 Flujo de Información de Vacunatorios en EEUU



Flujo de Vacunatorios en EEUU.

F

Figura 2.6.5 – Flujo de Información

En cuanto a los vacunatorios, el flujo de información es similar al nuestro, exceptuando nuevamente el tema de la codificación, aunque se está aplicando una metodología de codificación similar (ver figura 2.6.5).

Sí es preciso destacar el intensivo seguimiento que se realiza en cuanto al conocimiento de las vacunas que cada afiliado se ha administrado a lo largo de su vida.

## 2.6.6 Flujo de Información de Centros de Imágenes en EEUU

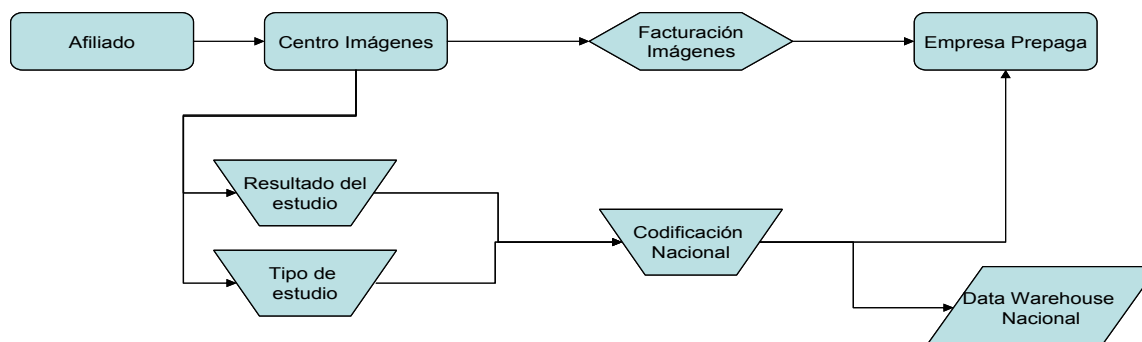


Figura 2.6.6 – Flujo de Información Centros de Imágenes en EEUU.

El flujo de información se completa con el diagnóstico sobre la imagen y la posterior codificación de las mismas. Sin embargo aún no se ha implementado un registro sobre cada imagen recolectada (ver figura 2.6.6).

## 2.6.7 Flujo de Información de Centros de Salud Mental en EEUU

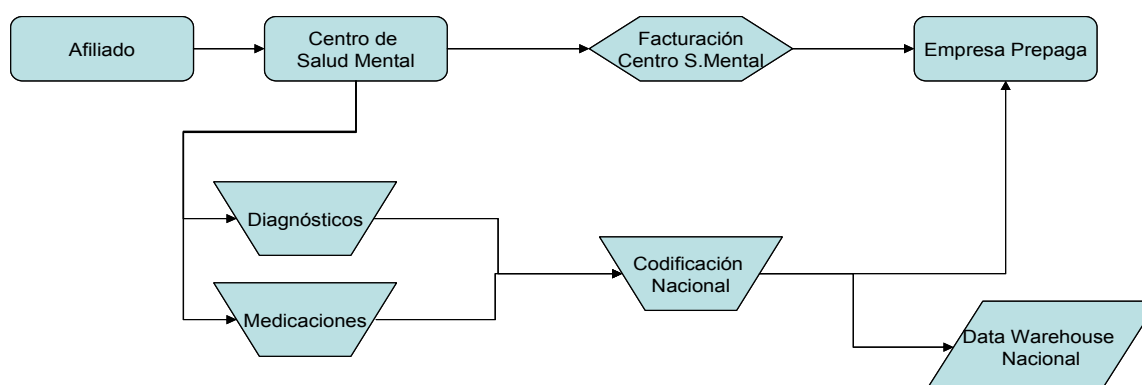


Figura 2.6.7 – Flujo de Información Salud Mental en EEUU.

Las internaciones en instituciones mentales son administradas muy eficazmente. Se codifica la información durante el tratamiento, obteniendo datos sobre diagnósticos intermedios, medicación administrada, tiempos de tratamiento, médicos intervinientes, causas de la internación, y sobre todo tratamientos prescritos que el paciente debe cumplir luego el alta (ver figura 2.6.7).

## 2.6.8 Flujo de Información de Servicios de Home Care en EEUU

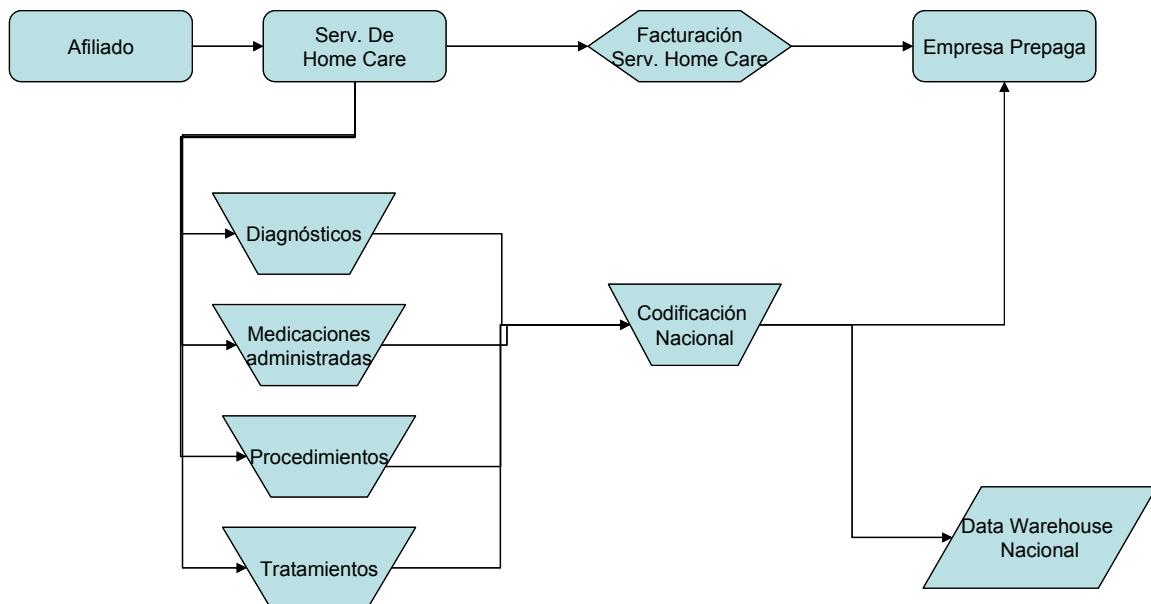


Figura 2.6.8 – Flujo de Información de Servicios de Home Care en EEUU.

El servicio de Home Care por su parte provee la información codificada referente a los tratamientos y procedimientos aplicados, medicación administrada, y diagnósticos durante la intervención (ver figura 2.6.8).

## 2.6.9 Flujo de Información de Servicios de Emergencia en EEUU

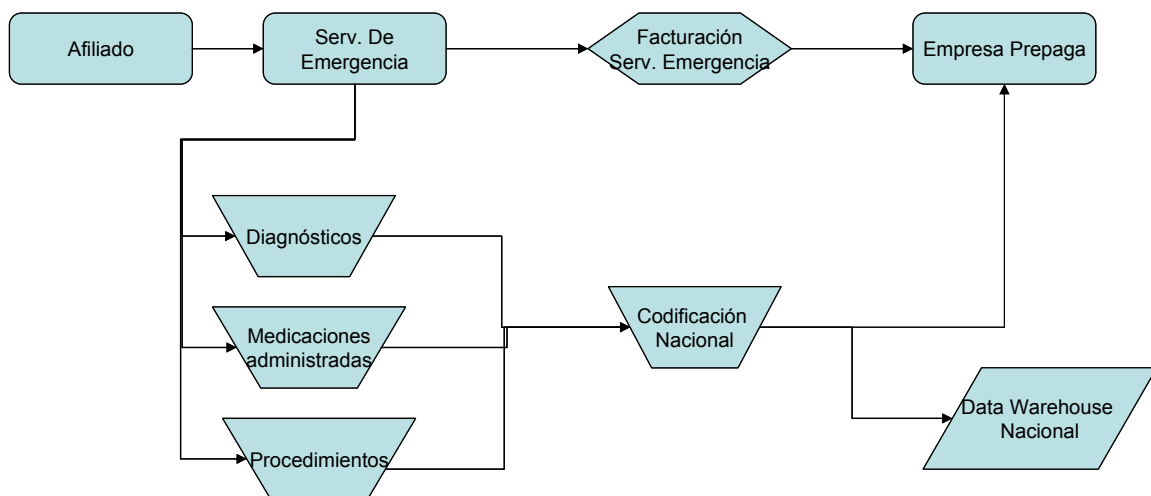


Figura 2.6.9 – Flujo de Información de Servicios de Emergencia en EEUU.

La información es codificada de acuerdo a los procedimientos realizados, el diagnóstico inicial de la emergencia, las medicaciones administradas durante la misma (ver figura 2.6.9).

### **2.6.10 Conclusiones sobre el flujo de Información en EEUU**

Sin lugar a dudas la información recolectada es sumamente precisa y cierra todo el círculo de la salud. La recolección de esta información no es caprichosa, sino que está orientada a resolver una problemática nacional tanto sobre la utilización de los servicios de salud, como sobre la progresión de las condiciones de salud de la población. En este sentido, al poseer tal claridad sobre los diagnósticos de los afiliados, pueden focalizarse los recursos más eficazmente, así como realizar estudios científicos sobre enfermedades que condicionan a la población.

Así mismo queda evidenciado como las compañías de medicina prepaga de ese país puede conocer el estado de salud de sus afiliados. Por consiguiente, la estrategia de implementación de programas de disease management es efectiva, permitiendo focalizar recursos en aquellos individuos que más riesgo representan.

### **2.7. Explotación de la Información para estrategias de “Disease Management”**

Si se extraen conclusiones de los tres aspectos indicados en el punto 2, es indudable que el costo para las empresas de medicina prepaga es cada vez mayor, y no hay buena política que pueda y deba frenarlo. Es evidente que limitar la utilización de servicios no es una opción, por lo cual la que parece ser la mejor manera de apalear esta situación a largo plazo es justamente lo contrario, incentivar al uso de los servicios. Este pensamiento viene de la mano de la medicina preventiva. Es indudable que la detección de enfermedades en estadios tempranos ahorra dinero en cuanto hecho drástico que derive en cirugías, transplantes e internaciones será evitado.

El Disease Management por consiguiente supone establecer programas específicos que se focalicen en afiliados enfermos, o con altas probabilidades de contraer ciertas enfermedades. Estos programas funcionan a partir de un seguimiento exhaustivo de los afiliados enfermos, donde se evalúa periódicamente su estado de salud, y ante la presencia de desvíos o anomalías sobre los patrones deseables, se trabaja junto con el afiliado para que este mejore su estado de salud y calidad de vida.

En consecuencia, este cambio de esquema ha llevado a que sea prioritaria la información sobre el estado de salud de los afiliados, así como los patrones de consumo que estos tienen.

Para abordar el problema antes descrito, existe una rama de las ciencias que conjuga varias disciplinas conocidas como Reconocimiento de Patrones (RP).

El reconocimiento de patrones es una útil herramienta no sólo para el mundo de los negocios, sino que sirve para guiar la acción en muchos otros campos. Existen aplicaciones que usan software especiales de reconocimiento de patrones para buscar tendencias y movimientos anómalos y por ende, detectar outliers, acciones fraudulentas o desvíos de las reglas esperadas. Las aplicaciones para programas de Disease Management pueden ser diversas, desde controles de consumo, clasificación automática de severidad y reconocimiento de patrones anómalos entre otros.

### **2.7.1 ¿Qué es el Disease Management?**

El Disease Management o Gerenciamiento de Enfermedades es una estrategia sistémica que busca el mejoramiento de la salud en pacientes que padecen enfermedades específicas. Esta estrategia está basada en la interacción de grupos clínicos interdisciplinarios, explotación eficiente de la información, educación y comunicación activa entre el paciente y el coordinador. En este sentido, esta estrategia considera que el cuidado de la salud puede ser entregado más eficientemente si aquellos pacientes enfermos toman un rol activo en su cuidado y su evolución es guiada conjuntamente por expertos en salud a fin de prevenir o demorar la progresión de su enfermedad.

### **2.7.2 El Disease Management en la Argentina**

En la Argentina sólo existen dos casos de considerable interés. El primero de ellos es llevado a cabo por el Hospital Italiano desde hace 5 años; el otro es llevado a cabo por la prepaga OMINT desde hace 3 años. Los resultados obtenidos en ambos lugares son bastante distintos, principalmente debido a las fuentes de información a las que pueden acceder.

### **2.7.3 El Disease Management del Hospital Italiano – Herramientas utilizadas**

Los programas establecidos en este hospital comunitario pueden ser comparables directamente con los realizados en los EEUU. Ello se debe principalmente a que los flujos de información que utilizan provienen de un sistema cerrado. Por ser este un hospital comunitario, los únicos centros de atención son en su hospital o consultorios externos de su misma empresa. Así mismo, los laboratorios de análisis clínicos y centros de imágenes se encuentran en su hospital, así como el servicio de farmacia y emergencias. Por consiguiente, los afiliados únicamente pueden utilizar servicios dentro del propio sistema hospitalario, por lo cual, la información no se pierde.

Toda la información es codificada, tanto consumos como diagnósticos, por lo cual el seguimiento que puede establecerse sobre el afiliado es total. Las herramientas utilizadas corresponden a software que validan reglas preestablecidas y detectando probabilidades de sucesos críticos.

### **2.7.4 Programas de Disease Management en OMINT S.A. - Caso de aplicación de la tesis**

En esta compañía privada de medicina prepaga se comenzó en el 2004 con este tipo de programas. Los resultados en sí no son tan consistentes como el caso anterior, ya que como toda empresa del rubro, presenta dificultades para acceder a información de diagnósticos.

En sus comienzos, la manera de detectar aquellos afiliados que requerían seguimiento se basaba principalmente en los censos de internación y un ranking de los afiliados que más recursos consumían. Sobre ellos comenzaron los programas, y logró demostrarse que en alguno de ellos, como lo es el PAID (Programa de Atención Integral al Diabético) los afiliados mejoraron su estado de salud. Sin embargo, la necesidad de utilizar correctamente los escasos recursos del programa, evidenciaba que era necesaria una explotación más adecuada de la información.

La explotación de la información no sólo requiere de la selección de un software adecuado, sino también de la evaluación de los flujos de información recolectados por la compañía. Como se mencionó oportunamente, estos programas requieren actuar rápidamente sobre los individuos, ya que de nada sirve detectar un evento tardío.



Sin duda, los avances tecnológicos en el campo de la computación han contribuido al auge de la disciplina de reconocimiento automático de patrones. No se puede estudiar el fenómeno del Reconocimiento de Patrones sin considerar las herramientas informáticas que lo soportan. Es preciso por ende detectar que tipo de herramientas de explotación de la información son utilizables y fueron probadas con éxito para encausar programas de Disease Management.

## **2.8 Herramientas de reconocimiento de patrones**

A continuación se listará un conjunto de herramientas de reconocimiento de patrones que fueron utilizadas para proyectos de similares características. Lamentablemente la escasa literatura académica sobre programas de Disease Management en los EEUU no indica puntualmente el tipo de herramientas utilizadas [Crawford *et al*, 2005]. Sí en la gran mayoría indica la utilización de herramientas de data mining, pero sin hacer alusión al tipo de herramienta. Por consiguiente se determinará el potencial uso de cada uno desde el punto de vista del caso a resolver.

### **2.8.1 Aprendizaje de Máquinas (Machine Learning)**

El aprendizaje de máquinas es un área de la inteligencia artificial relacionada con el desarrollo de técnicas que permiten a las computadoras “aprender”. Más específicamente, el aprendizaje de máquinas es un método para crear programas de computadoras mediante el análisis de “sets” de datos.

El aprendizaje de máquinas está íntimamente ligado a la estadística, ya que ambos campos estudian el análisis de datos, pero a diferencia de la estadística, el aprendizaje de máquinas está relacionado con la complejidad algorítmica de implementaciones computacionales. [Internet 1].

El potencial uso de estas herramientas será la detección de reglas dentro del set de datos y su validación por medios estadísticos.

#### ***General learning: Pattern Recognition, Neural Networks, Machine Learning***

- DTREG Decision Trees for regression and classification
- HTK, Hidden Markov modelling Toolkit
- Hierarchical Clustering Explorer for Windows

- LEARNSC Learning and soft computing: SVM, NN and Fuzzy Logic Models in Matlab
- Partek Pattern Recognition and Recognition Technologies
- PRTOOLS Pattern Recognition toolbox in Matlab, by Bob Duin
- SPR/ANLIB, C-libraries for Statistical Pattern Recognition and Artificial Neural Networks
- Statistical Pattern Recognition Toolbox
- Tilburg Memory Based Learner

### 2.8.2 Support Vector Machines (SVM)

Los SVM son un grupo de métodos de aprendizaje supervisado relacionado usados para clasificación y regresión. Básicamente es un algoritmo para crear funciones a partir de un grupo de datos de entrenamiento.

Cuando se utiliza para clasificación, el algoritmo SVM crea un “hiperplano” que separa a los datos en dos clases con un margen máximo. Dados algunos datos de ejemplo marcados “sí” o “no”, un hiperplano de máximo margen que divide los “sí” de los “no” es identificado, de tal manera de que la distancia entre el hiperplano y los ejemplos más cercanos (el margen) se maximice. [Internet 2].

Este tipo de herramientas puede ser utilizada para la clasificación de riesgo, y la determinación de la inclusión o no dentro de un grupo.

#### **Support Vector Machines (SVM)**

- BSVM A decomposition method for bound-constrained SVM formulations
- Least Squares - Support Vector Machines Matlab/C Toolbox
- mySVM SVM implementation for pattern recognition and regression
- Nearest Point Algorithm by Sathiya Keerthi (in FORTRAN)
- SMO training of SVM Sequential Minimal Optimization in C++
- SVMLight

### 2.8.3 Aprendizaje Bayesiano (Bayesian Learning)

El aprendizaje Bayesiano, o “Bayesian learning” es un proceso simple de:

1. Especificar modelos de “distribución de probabilidad previa” (o “previo”)
2. Integrar usando la ley de Bayes con respecto a toda la información observada para computar una “distribución de probabilidad posterior” (o “posterior”)
3. Predecir de acuerdo a la posterior.

A su vez tiene varias ventajas sobre otros programas de aprendizaje:

1. Interpolación Los métodos de aprendizaje Bayesianos interpolan totalmente de forma ingenieril, Cuando se enfrentan a cualquier problema de aprendizaje, existe la elección de cuánto tiempo y esfuerzo realiza un humano contra lo que realiza una computadora. Cuando se crea un sistema de ingeniería, se construye un modelo del mundo y luego se encuentra un buen controlador para ese modelo. Los métodos Bayesianos interpolan hasta ese extremo porque el “previo” Bayesiano puede ser una función delta de un modelo del mundo real. Lo que esto significa es que una receta de “pensar más” (acerca de especificar un “previo”) y “computar más” (para calcular un “posterior”) eventualmente tiene éxito. Otros enfoques de aprendizaje de máquinas no poseen esta garantía.
2. Métodos cuasi-Bayesianos (near-Bayesian) y de lenguaje Bayesiano tienen un lenguaje asociado para especificar “previos” y “posteriores”. Esto es significativamente útil cuando se trabaja en la parte de “pensar más” de la solución.
3. El Aprendizaje Bayesiano Intuitivo involucra especificar un “previo” e integrar, dos actividades que son útiles universalmente (se pueden aplicar para otras partes del problema)

Con todas estas ventajas, el aprendizaje Bayesiano es un programa fuerte. Sin embargo, existen también algunas desventajas significativas:

1. Información teóricamente inviable. Especificar un “previo” es extremadamente difícil. Prácticamente, se debe especificar un número real para cada ajuste de los parámetros del modelo.
2. Inviabilidad Computacional. Incluso si se logra especificar correctamente un “previo” de un sistema complejo, el posterior cómputo de un “posterior” puede ser extremadamente difícil. Esto significa que se requiere de aproximación computacional.

3. No automatización. La parte de “pensar más” requiere de personas. Mientras haya problemas nuevos de aprendizaje de máquinas, seguirá existiendo la necesidad de ingenieros Bayesianos para solucionarlos. [Internet 3]

Estas herramientas pueden ser utilizadas para determinar probabilidades a posteriori de un set de datos dados ciertos cambios en el juego de variables interrelacionadas.

### ***Bayesian Learning and Gaussian Processes***

- BayesiaLab
- Flexible Bayesian Modelling Bayesian regression and classification models based on Neural Networks and Gaussian processes
- MAP-1 Gaussian processes for regression by Carl Rasmussen
- Tpros and Cpros Gaussian processes package by Mark Gibbs
- Software Packages for Graphical Models / Bayesian Networks

### **2.8.4 Análisis de Componentes Independientes (Independent Component Analysis)**

El Análisis de Componentes Independientes (ICA) es una técnica estadística y computacional para revelar factores ocultos que subyacen bajo grupos (sets) de variables, mediciones o señales aleatorias

El ICA define un modelo generativo para los datos multivariados observados, los cuales son generalmente una gran base de datos de muestras. En el modelo, se asume que las variables de datos son combinaciones lineales de algunas variables latentes desconocidas, y el sistema de mezclado tampoco se conoce. También se asume que las variables latentes son no gaussianas y mutuamente independientes, y se las denomina los componentes independientes de los datos observados. Estos componentes independientes también, denominados “fuentes” o “factores”, también pueden ser hallados mediante el ICA.

El ICA está superficialmente relacionado con el análisis de componentes principales y el análisis de factores. El ICA, sin embargo, es una técnica mucho más poderosa, capaz de encontrar los factores o fuentes subyacentes donde los métodos clásicos fallan completamente.

Los datos analizados por ICA pueden ser originados desde diversas clases de campos de aplicación, incluyendo imágenes digitales, bases de datos de documentos, indicadores económicos y mediciones psicométricas. En muchos casos, las mediciones se dan como un grupo de señales paralelas o series de tiempo; el término “fuente de separación ciega” se utiliza en estos casos. Ejemplos típicos de esto son mezclas de señales de discurso simultáneas que han sido recogidas por varios micrófonos, ondas cerebrales grabadas por múltiples sensores, señales de radio interferentes arribando a un teléfono celular, o series de tiempo paralelas obtenidas de algún proceso industrial. [Hyvärinen, 2001]

### ***Independent Component Analysis (ICA)***

- Blind Source Separation Using Temporal Predictability in Matlab
- ICA toolbox in Matlab
- ICALAB for Signal and Image Processing in Matlab

### **2.8.5 Redes Neuronales (Neural Networks)**

Una Red Neuronal Artificial (ANN, del inglés “Artificial Neural Network”) es un paradigma de procesamiento de información que está inspirado en la forma en que los sistemas nerviosos biológicos, tales como el cerebro, procesan la información. El elemento clave de este paradigma es la estructura novedosa del sistema de procesamiento de la información. Se compone de un gran número de elementos de procesamiento altamente interconectados (neuronas) funcionando al unísono para resolver problemas específicos.

Las ANN, al igual que las personas, aprenden mediante el ejemplo. Una ANN está configurada para una aplicación específica, tal como el reconocimiento de patrones y clasificación de datos, a través de un proceso de aprendizaje. El aprendizaje en sistemas biológicos involucra ajustes a las conexiones sinápticas que existen entre las neuronas. Esto es cierto también para las ANN.

Las redes neuronales, con su habilidad para extraer significado de datos complicados o imprecisos, pueden usarse para extraer patrones y detectar tendencias que son muy complejas para poder ser detectadas tanto por humanos como por otras técnicas informáticas. Puede pensarse a una red neuronal entrenada como un “experto” en la categoría de información a la que ha sido asignada para analizar. Este experto puede ser utilizado entonces para proveer

proyecciones dadas nuevas situaciones de interés y contestar preguntas del tipo “qué pasaría si”.

Otras ventajas incluyen:

1. Aprendizaje adaptativo: La habilidad de aprender cómo hacer tareas basado en los datos dados para entrenamiento o como experiencia inicial.

2. Auto-Organización: Una ANN puede crear su propia organización o representación de la información que recibe durante el tiempo de aprendizaje

3. Operación en Tiempo Real: Las computaciones de una ANN pueden ser realizadas en paralelo. Actualmente, se está diseñando y fabricando hardware especial para aprovechar esta capacidad.

4. Tolerancia a Fallas vía Codificación de la Información Redundante: La destrucción parcial de una red lleva a la correspondiente degradación de la performance. Sin embargo, algunas capacidades de la red pueden ser retenidas a pesar de daños importantes a la red [Stergiou y Siganos].

### ***Self-Organizing Maps (SOM) and Neural Networks (NN)***

- Artificial Neural Network toolbox in Matlab
- NNUGA Neural Network Using Genetic Algorithms
- PINN Pseudo-Inverse associative Neural Networks, a library in C++
- SNNS (Stuttgart Neural Network Simulator)
- SOM\_PAK and LVQ\_PAK
- SOM

### **2.8.6 Aprendizaje No Supervisado (Unsupervised Learning)**

El Aprendizaje No Supervisado es un método de aprendizaje de máquinas donde el modelo se ajusta a las observaciones. Se distingue del Aprendizaje Supervisado por el hecho de que no existe una salida (output) “a priori”. En el aprendizaje no supervisado, se recoge un set de datos de objetos de entrada (input) que luego es tratado como un grupo de objetos de variables aleatorias.

El aprendizaje no supervisado se puede utilizar en conjunto con la inferencia Bayesiana para producir probabilidades condicionales (aprendizaje supervisado) para cada una de las variables aleatorias.

El aprendizaje no supervisado puede ser útil también para la compresión de datos. Fundamentalmente, todos los algoritmos de compresión de datos dependen

explícita o implícitamente en una distribución probabilística sobre una serie de inputs. Otra forma de aprendizaje no supervisado es el “clustering”, como por ejemplo los mapas autoorganizativos, a continuación [Internet 4].

Este tipo de herramientas puede utilizarse para detectar automáticamente casos anómalos o tendencias negativas para en el cuidado de los pacientes.

### ***Unsupervised Learning***

- CLUTO Clustering high-dimensional datasets
- Fast Fuzzy Cluster
- Snob Mixture modelling

### **2.8.7 Mapas Autoorganizativos (Self Organizing Maps)**

Los Mapas Autoorganizativos (SOM, del inglés “Self-Organizing Map”) son un tipo especial de algoritmo de redes neuronales, introducidas por Teuvo Kohonen en 1982. El algoritmo SOM, también conocido como Mapa de Kohonen, es uno de los algoritmos de redes neuronales más conocidos. En contraste con muchas otras redes neuronales que usan aprendizaje supervisado, el SOM se basa en aprendizaje no supervisado.

El SOM es un tipo único de red neuronal en el sentido de que construye un mapa que preserva la topología, desde el espacio multidimensional a unidades de mapa de tal forma de que las distancias relativas entre puntos (datos) se conservan. Las unidades de mapa, o neuronas, generalmente forman un entramado regular bidimensional, donde la ubicación de una neurona contiene información semántica, El SOM puede entonces servir como una herramienta de agrupamiento, o “clustering”, de datos multidimensionales. Debido a su forma típicamente bidimensional, es además fácil de visualizar.

Otra ventaja importante del SOM es su capacidad de generalizar. En otras palabras, puede interpolar entre entradas previamente encontradas. [Internet 5].

### ***Self-Organizing Maps (SOM) and Neural Networks (NN)***

- Artificial Neural Network toolbox in Matlab
- NNUGA Neural Network Using Genetic Algorithms
- PINN Pseudo-Inverse associative Neural Networks, a library in C++

- SNNS (Stuttgart Neural Network Simulator)
- SOM\_PAK and LVQ\_PAK
- SOM



### **3. DETERMINACIÓN DEL PROBLEMA**

*En este capítulo se describen el Objetivo de la tesis (inciso 3.1) y los macro-aspectos y micro-aspectos que concluyen con la Definición del problema (inciso 3.2).*

#### **3.1 Objetivo de la tesis**

Aplicar un proceso de exploración de la información con el propósito de efficientizar la focalización de recursos en los programas de Disease Management de OMINT S.A.

#### **3.2 Definición del problema**

Dentro del contexto general de la medicina prepaga en la Argentina, tal como fue analizado en el Inciso 2, uno de los problemas que se plantea surge de la creciente utilización de recursos médicos y económicos por parte de la población con padecimientos crónicos. En la terminología médica existe un concepto denominado “piramidización del gasto”. Esta caracterización indica que el gasto y la composición poblacional del sistema de salud, tanto pública como privada, responden a una estructura netamente piramidal, donde la base de la pirámide poblacional representa una pequeña porción de la pirámide de gasto. En lo que ello respecta, si se segmenta la población en tres estratos dados por individuos sanos, individuos ambulatorios o con episodios frecuentes, e individuos crónicos, se denota que la amplia base ocupada por el 83% de los individuos está conformada por los estratos sanos y ambulatorios, mientras que la pequeña cúspide restante se encuentra conformada por padecientes crónicos (ver figura 3.2.1-1).

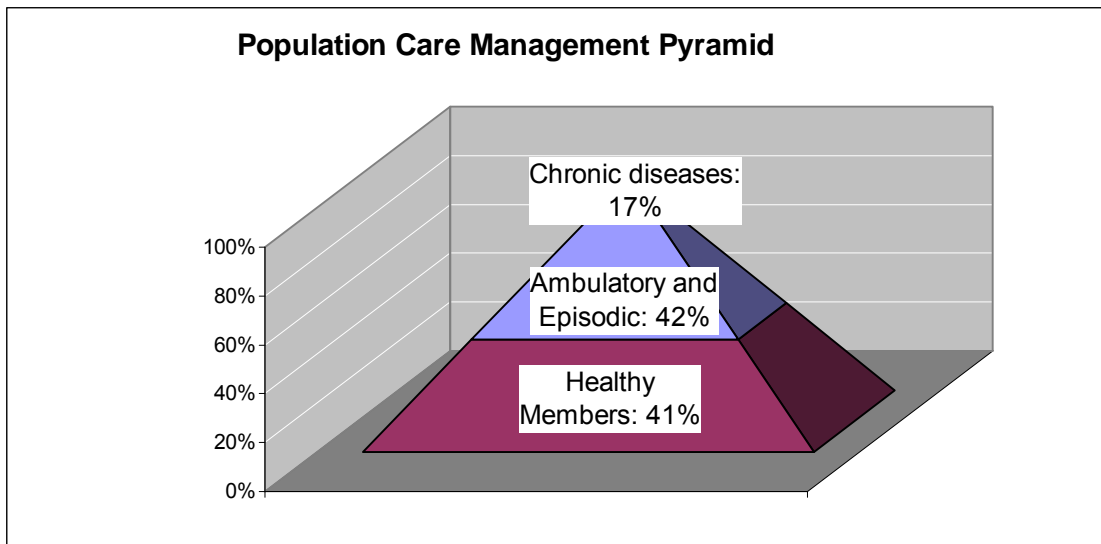


Figura 3.2.1-1 – Estratos de los cuidados de salud.

Evaluando el gasto médico de cada uno de estos estratos, se evidencia una desproporción sumamente marcada en la utilización de recursos, tal como se muestra la figura 3.2.1-2 contigua.

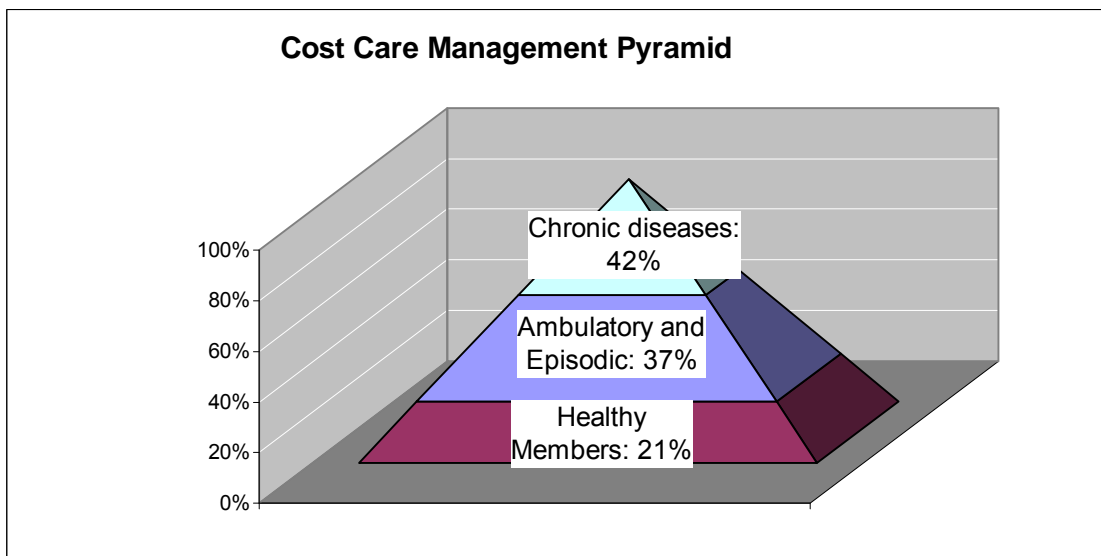


Figura 3.2.1-2 – Estratos del costo en salud.

Esta desproporción es llamativa en cuanto a que los individuos crónicos representados por el 17% del total gastan el 42% de los recursos, mientras que el 41% de la población restante consume solamente el 21% de los recursos.

Ciertamente queda reflejada la verdadera importancia que requiere el gerenciamiento de los padecientes crónicos<sup>1</sup>.

De acuerdo a ello queda evidenciada la necesidad de focalizarse en este grupo poblacional, a fin de controlar el gasto actual, prevenir un mayor gasto futuro y eficientizar los recursos disponibles en pos de una mejor calidad médica y rentabilidad.

Dentro de este contexto más específico es que surgen los programas de Disease Management o Chronic Disease Management. Estos programas presentan dos razones que impulsan su permanencia: una puramente económica, y otra en lo que respecta a la calidad de vida del paciente.

La primera de ellas implica que este tipo de programas tendrán sentido únicamente si evidencian un impacto económico considerable, especialmente al largo plazo, dado que para su funcionamiento utilizan una gran cantidad de recursos, desde humanos hasta tecnológicos. La evidencia de esto es escasa, principalmente debido a la dificultad que presenta su medición, y sobre todo, debido a que son programas que demuestran su aptitud al largo plazo. Sin embargo, evaluando estos para el caso de OMINT, se ha demostrado que los afiliados enrolados en los programas de Coordinación de Cuidados Médicos han gastado en promedio un 18% menos en los últimos dos años. La metodología de cálculo de este valor se detalla en el “Informe sobre el impacto económico de los programas de Case Management” [OMINT S.A., 2007], y en términos de dinero, el ahorro acumulado durante el período fue de alrededor de \$450.000,00. Así mismo, si se analiza concretamente el gasto durante el mes de mayor gasto, aquellos socios enrolados en los programas presentan en promedio un gasto del 34.2% menos que aquellos que no lo están [OMINT S.A., 2007]. Ello indica que el enrolamiento del socio juega un papel fundamental en el ahorro impartido, y en especial si este es previo al mes de mayor consumo. Lógicamente tal período no puede ser reconocido anticipadamente, aunque sí pueden inferirse previamente ciertos patrones que impliquen una presunción de riesgo potencial.

Respecto a las razones de calidad de vida, numerosos estudios comprueban que los programas de Disease Management tienen un marcado impacto en el mejoramiento de la salud de sus pacientes a través del gerenciamiento efectivo de los enfermos por medio de equipos interdisciplinarios de profesionales en salud [Lorig K, 2001]. Así mismo no debe olvidarse la implicancia económica que este hecho presenta.

---

<sup>1</sup> La estructura de segmentación poblacional y de costo representada corresponde al caso particular de OMINT S.A. Estos porcentajes pueden variar entre las distintas compañías, aunque la conclusión es similar en cuanto a que la población crónica utiliza la mayor cantidad de recursos.

A partir de lo expuesto se manifiesta que los programas de Disease Management son responsables de un beneficio económico considerable y deben ser eficientemente administrados. Sin embargo, la evidencia presentada por OMINT S.A., indica que en pos de maximizar los recursos utilizados deben perfeccionarse los sistemas de información, especialmente en lo que se refiere a detección y gerenciamiento de pacientes. Dado que el ahorro de estos programas es determinado por la efectividad de las coordinadoras y equipos médicos en la administración e inclusión de los pacientes, debe estudiarse un enfoque que contribuya a la detección temprana de enfermos, y priorización de su gerenciamiento.

En particular, el presente proyecto buscará entonces dar respuesta a un problema de explotación ineficiente de los sistemas de información dentro de los programas de Disease Management de OMINT S.A. Estos presentan una seria dificultad en cuanto a una utilización ineficaz de sus recursos, objeto de la baja calidad de la información empleada para la inclusión y gerenciamiento de sus pacientes. En este sentido, el procedimiento utilizado hacia fines del 2006 supone recabar información proveniente de los censos diarios de internación, y en particular de un ranking semestral de socios con alto gasto. Si bien tras analizar los flujos de información en la Argentina tal proceder es coherente, ello altamente ineficaz, en cuanto que evalúa información tardía y peor aún, información basada únicamente en gastos. La realidad que plantea la tendencia del mercado presentada en el Inciso 2, requiere que los recursos invertidos en estos programas sean eficazmente aprovechados, y por ello es necesario un nuevo enfoque basado en una explotación de la información de una manera más eficiente.

## 4. SOLUCIÓN PROPUESTA

*En este capítulo se describen la solución propuesta (inciso 4.1); se describe el contexto general de la solución propuesta: ¿qué riesgo se desea medir? (inciso 4.2.1); se describen el ¿por qué medir tal riesgo? (inciso 4.2.2); y la dinámica evolutiva del riesgo (inciso 4.2.3). Se describen los pasos de la metodología CRISP-DM utilizada): Entendimiento del negocio (inciso 4.3); Entendimiento de los datos (inciso 4.4); Preparación de los datos (inciso 4.5); Modelización (inciso 4.6); Evaluación (inciso 4.7) y Entrega Final (inciso 4.8).*

### 4.1 Solución Propuesta

La explotación ineficiente de los sistemas de información para el gerenciamiento de los programas de Disease Management de OMINT S.A. será resuelto a partir del desarrollo de un scoring de riesgo para la población de afiliados con padecimientos crónicos. La alternativa propuesta encara la explotación de información desde un enfoque de minería de datos que combina distintas variables clínicas, demográficas, económicas, biométricas y del estado de salud, detectando el patrón de consumo de distintos subgrupos poblacionales, y aplicando un scoring de riesgo basado en índices complementarios que cada individuo presente en un momento dado. Esto último implica que la solución propuesta tiene un factor dinámico en el riesgo a ponderar, a partir que permite priorizar los pacientes, y a su vez automatizar la actualización diaria en los sistemas de información, eliminando por completo el delay que el antiguo procedimiento presenta.

### 4.2 Contexto general de la solución propuesta

#### 4.2.1 ¿Qué riesgo se desea medir?

En el contexto de la solución propuesta es necesario definir a que se denomina riesgo. La literatura en la materia indica que existen una amplia variedad de riesgos potenciales que pueden ser medidos. Estos riesgos pueden incluir eventos terminales como lo es la muerte, eventos biológicos como lo es el deterioro en el estado de salud o fallas en el funcionamiento de órganos vitales, eventos clínicos como lo son las internaciones hospitalarias, u eventos económicos como lo es el costo de salud entre tantos otros.

La mayoría de los estudios publicados en círculos científicos ha definido el riesgo en términos del riesgo de hospitalización o internación. Algunos estudios son muy específicos y se focalizan en el riesgo de la reinternación [Parker JP, 2003] o la temprana reincidencia en emergencias [McCusker J, 2000], mientras otros evalúan el riesgo de “utilización” desde un aspecto más amplio, como la utilización de servicios ambulatorios y de emergencia [Wahls TL, 2004]. Así mismo existen estudios que consideran que existen otro tipo de riesgos que son más dificultosos de medir bajo el marco del costo e internación. Estos incluyen riesgos como la muerte, el decaimiento funcional [Dendukuri N, 2004], complicaciones clínicas [Selby JV, 2001] y enfermedad [Barber M, 2001]. A pesar que la muerte no es necesariamente un evento costoso por sí solo, la evidencia muestra que el último año de vida se caracteriza por un alto costo en servicios médicos y por ello es importante para el sector prepago [Petersen LS, 2005].

Sea cual fuere el enfoque abordado, todos concluyen con mucha lógica que el esquema central gira en torno a que un pequeño número de pacientes debe ser clasificado con alto riesgo y ser candidato a utilizar gran cantidad de recursos.

#### **4.2.2 ¿Por qué medir tal riesgo?**

Existen dos razones principales para tratar de identificar quienes son aquellos sujetos de alto riesgo y en consecuencia presumibles de ser costosos. La primera de ellas es con objeto de determinar una prima de seguro más justa y real en relación al gasto actual y futuro en salud. Esto es conocido como “ajuste de riesgo” y es utilizado por empresas aseguradoras de salud, o comúnmente conocidas como empresas de medicina prepaga. La segunda razón es para identificar pacientes en los cuales una apropiada intervención podría mejorar sustancialmente su cuidado y prevenir futuras complicaciones en lo que a su salud respecta. Aunque estas dos razones son muy diferentes, ambas concuerdan en reconocer que, en todo sistema, existen recursos escasos que deben ser utilizados de la manera más apropiada. La primera razón es puramente financiera y busca compensar a las aseguradoras por el costo del servicio médico. La segunda se fundamenta en identificar el target apropiado al que deben focalizarse intensamente los escasos recursos, a fin de lograr un mejor cuidado de la salud, y en pos de ahorrar dinero en el largo plazo.

Dada la naturaleza aseguradora de las empresas de medicina prepaga, estas tienen un gran incentivo para desarrollar técnicas de predicción de riesgo a partir de la necesidad de poder cubrir sus costos de prima en base al mix de severidad de los afiliados que componen su cartera. Pacientes enfermos incurren en altos gastos, por lo cual existe un incentivo económico por parte de las aseguradoras a evadir su admisión. El ajuste de riesgo busca corregir esto, o al menos, reducir las

razones de ello, asegurando que las aseguradoras resguarden sus costos a partir de una prima más justa. En un sistema sin ajuste de riesgo, las prepagas admitirían únicamente individuos sanos, ya que son los más baratos de cuidar. Basándonos en términos económicos informados en ciertos estudios, el 76.3% de los miembros sanos consumen el 14% del total de gasto, mientras que el 15.7% representado por lo enfermos, consumen el 75.7% del total [Huges AS, 2004].

Respecto a las técnicas de identificación o estratificación de pacientes acorde a su riesgo, estas son útiles para establecer que pacientes requieren una intervención adecuada y eficiente de recursos en pos de mejorar su salud, reducir costos futuros y facilitar un mejor planeamiento de su cuidado. La prevención temprana por medio de una intervención adecuada es un beneficio global tanto para la empresa como para el paciente, logrando una mejor sustentabilidad económica a largo plazo. Si el sistema es capaz de detectar aquellos sujetos en riesgo, el uso eficiente de recursos permitirá focalizarse en su cuidado, y de una forma más efectiva.

#### **4.2.3 Dinámica evolutiva del riesgo**

Otro tema que se remarca en la literatura es la dinámica evolutiva del riesgo. La mayoría de los programas de case management asumen que el nivel de riesgo individual y su estado de salud es invariable. Ello es erróneo y los resultados de algunos estudios lo confirman; en un estudio dentro de una compañía de medicina prepaga, los miembros de alto costo representaban el 1% y gastaban el 21% del total en 1998. Al año siguiente, ese mismo 1% representaba solamente el 7% del gasto total [Dove HG, 2003]. Ello es debido a que el consumo de recursos de los miembros de mayor gasto generalmente decrece, aún en casos de no haber sido intervenidos por estos programas. Esto adhiere una complejidad extra al proceso de identificación de pacientes aptos para ser intervenidos.

#### **4.3 – 4.8 Metodología del proyecto**

Para el desarrollo del proyecto se utilizó como guía la metodología CRISP-DM en sus diferentes etapas. Cada una de ellas se expondrá a continuación.

## **4.3 Entendimiento del negocio**

### **4.3.1.1 Contexto general**

La causa raíz de la ineficiencia de los programas de Disease Management por parte de las empresas de medicina prepaga en la Argentina radica en la ineficiente utilización de las fuentes de información. Hasta el momento no se ha desarrollado un enfoque global que combine distintos tipos de variables que puntualicen aspectos diversos del estado de salud, consumo e incidencia económica del paciente. La perspectiva actual es netamente lineal, donde el paciente de alto gasto es candidato a ser incluido dentro de los programas y es objeto de un seguimiento más cuidadoso. Si se considera que el gasto es consecuencia de numerosas variables que, combinadas, dan como resultado un gasto elevado, se evidencia que la perspectiva actual es ineficiente. Así mismo el alto gasto es consecuencia de una utilización desmedida de recursos, razón que no necesariamente implica un deterioro en la salud del paciente. En este sentido la inclusión directa en los programas es objetable y probablemente sea sesgada.

Una identificación objetiva de los potenciales pacientes y de priorización para el gerenciamiento es necesaria, y requiere un enfoque global que puede ser provisto por la minería de datos. La minería de datos es una solución eficaz frente a problemas de alta complejidad que requieren la combinación de diversas fuentes de información, para la extracción de conocimiento útil, buscando patrones, comportamientos, agrupaciones, secuencias o tendencias que puedan generar algún modelo que permita comprender mejor el contexto para la toma efectiva de decisiones en estos programas.

### **4.3.1.2 Contexto puntual del problema**

Los programas de Disease Management de OMINT S.A. presentan una seria dificultad funcional objeto de la baja calidad de la información empleada para la inclusión y gerenciamiento de sus pacientes. El procedimiento utilizado hacia fines del 2006 supone recabar información proveniente de los censos diarios de internación, y en particular de un ranking semestral de socios con alto gasto. Si bien tras analizar los flujos de información en la Argentina tal proceder es coherente, ello altamente ineficaz, en cuanto a que evalúa información tardía y parcial, basada únicamente en gastos. Así mismo tal proceder no permite una correcta priorización de los pacientes para su gerenciamiento diario.



#### **4.3.1.3 Área de correspondencia del problema**

El problema detallado corresponde a los Programas de Coordinación de Cuidados Médicos, los cuales engloban los Programas de Disease Management de OMINT S.A. a través de un equipo interdisciplinario de profesionales de la salud denominado Equipo de Coordinación de Cuidados Médicos.

#### **4.3.1.4 Necesidades y expectativas de los usuarios**

El Equipo de Coordinación de Cuidados Médicos manifiesta una necesidad puntual de eficientizar sus recursos para el gerenciamiento de sus pacientes, a partir de un sistema que provea información más temprana, con mayor poder resolutorio y que implique una facilidad para la utilización en el gerenciamiento diario.

#### **4.3.1.5 Ventajas y desventajas de la solución**

La solución detallada en el inciso 4.1 presenta ventajas en cuanto a que prioriza sobre quienes se deben focalizar los recursos más fuertemente, a partir de relevar variables clínicas, de salud, económicas, de consumo demográficas y biométricas. Las variables relevadas se combinan a partir de diferentes índices que complementariamente brindan un enfoque global sobre el paciente. Así mismo presenta una componente dinámica al permitir actualizar el score diariamente y reordenar los pacientes de riesgo prioritario.

Por otro lado, la solución propuesta presenta una desventaja objeto de la parametrización matemática de las variables que concluyen en un puntaje de riesgo individual. Esto elimina la componente subjetiva y de conocimiento fino que puede tener el coordinador sobre el paciente, pudiendo en algún caso puntuar un riesgo menor o mayor al que subjetivamente pueda ser estipulado por el experto médico. Así mismo este scoring fue concebido únicamente para pacientes crónicos.

#### **4.3.1.6 Unidades de negocio**

El proyecto concierne a la gerencia de Área Médica, a la cual competen la distribución y utilización de los recursos económicos y humanos hacia los afiliados.

#### **4.3.1.7 Recursos organizacionales claves para el proyecto**

El proyecto depende de la gerencia de Área Médica, bajo el desarrollo del sector de Programas Especiales. Este sector es clave para el proyecto al ser el encargado del desarrollo e implementación de nuevos proyectos, así como del gerenciamiento de los Programas de Coordinación de Cuidados Médicos.

#### **4.3.2 Objetivos del negocio y criterios de éxito**

##### **4.3.2.1 Objetivos del negocio**

Implementar un enfoque integral que permita a los Programas de Disease Management de OMINT S.A. puntualizar sobre que pacientes crónicos deben focalizarse y priorizar los recursos, con objeto de lograr la mejora en la salud del paciente a través de la educación en el autocontrol y control preventivo, gastando eficientemente los recursos que estos utilizan, controlando la situación actual y previniendo el mayor gasto futuro.

##### **4.3.2.2 Criterios de Éxito**

Obtener un “enfoque integral” a partir de un scoring resultante que contemple aspectos económicos, demográficos, clínicos, biométricos y del estado de salud, excediendo las limitaciones actuales respecto al uso de información que los programas de Disease Management presentan.

Así mismo debe contemplar una reducción notable en los tiempos de detección y actualización de la información, incorporando una componente dinámica que permita priorizar los recursos a asignar.

#### **4.3.3 Recursos**

Los recursos a utilizar para la realización de este proyecto serán los siguientes:

- Recursos Humanos:

- Exploración de la información: Analista de Programas Especiales de OMINT S.A.
- Conocimiento médico-clínico: Coordinadores de los Programas de Coordinación de Cuidados Médicos de OMINT S.A.

- Recursos Técnicos: Sistemas de información de OMINT S.A.

#### 4.3.4 Consideraciones

El desarrollo del proyecto responde a una necesidad de negocio actual para los programas de Disease Management de OMINT S.A., donde el scoring dinámico a ser establecido fue considerado una alternativa viable frente a la mecánica de los sistemas de información de la compañía al momento del comienzo del proyecto.

Así mismo, el relevamiento de los datos será de una ventana temporal correspondiente al año 2005, por lo cual al momento de la implementación en los sistemas deberá estipularse el período a considerar.

A lo largo del desarrollo del proyecto se preservará la confidencialidad de los pacientes a partir de la mutación de su número identificadorio.

#### 4.3.5 Riesgos y planes de contingencia

Los principales riesgos durante el desarrollo del proyecto son los siguientes:

- Inaccesibilidad de la información.
  - o Plan de contingencia: Acceder a los permisos necesarios de las bases depuradas o buscar fuentes alternativas de alimentación desde las bases de datos crudas a las que sí se tiene acceso.
  
- Incongruencia entre los resultados arrojados y el criterio médico
  - o Plan de contingencia: La razonabilidad de los resultados debe ser en términos médicos, por lo cual ante un resultado incoherente se optará por utilizar el criterio médico.
  
- Validez de los índices a determinar.
  - o Plan de contingencia: Los índices de riesgo que alimenten el score deben ser estadísticamente válidos. Sin embargo en ocasiones ello es imposible dado que se desconocen las causas de ciertas relaciones. Por consiguiente, se optará por determinar el índice de una manera ad-hoc de acuerdo al criterio médico.

### 4.3.6 Glosario

- *OMINT S.A.*: Compañía de Medicina Prepaga radicada en la Argentina.
- *Programas Especiales*: Sector de I+D en OMINT S.A. a cargo de los Programas de Coordinación de Cuidados Médicos.
- *Programas de Coordinación de Cuidados Médicos*: Refiere a los Programas de Disease Management de OMINT S.A.
- *Case Management*: Programas de gerenciamiento por casos.
- *Coordinadores de Cuidados Médicos*: Refiere al equipo de especialistas en salud que atienden y coordinan los programas de Disease Management de OMINT S.A.
- *Gerenciadores*: Refiere a los coordinadores de los programas de Coordinación de Cuidados Médicos de OMINT S.A.
- *SSS*: Superintendencia de Salud
- *CRISP-DM*: Metodología de desarrollo de proyectos de minería de datos.
- *SOM*: Self Organizing Maps

### 4.3.7 Costos y beneficios

#### 4.3.7.1 Beneficios del proyecto

El proyecto en cuestión tiene un alto impacto económico y de gerenciamiento para OMINT S.A., aunque la medición de los mismos resulta compleja. Oportunamente se ha comentado que los programas de Disease Management de OMINT han demostrado que los afiliados enrolados en los programas de Coordinación de Cuidados Médicos han gastado en promedio un 18% menos en los últimos dos años. La metodología de cálculo de este valor se detalla en el “Informe sobre el impacto económico de los programas de Case Management” [OMINT S.A., 2007], y en términos de dinero, el ahorro acumulado durante el período fue de alrededor de \$450.000,00. Por otra parte, analizando el gasto durante el mes de mayor gasto, aquellos socios ya enrolados en los programas presentan en promedio un gasto del 34.2% menos que aquellos que no lo están [OMINT S.A., 2007]. Ello indica que el enrolamiento del socio juega un papel fundamental en el ahorro impartido, y en especial si este es previo al mes de mayor consumo. El scoring propuesto debería mejorar la detección temprana de los socios a incluir, por lo cual el ahorro a percibir debe ser mayor.

Así mismo, dado que el scoring determinará descendientemente el riesgo de cada socio, se podrá focalizar puntualmente sobre aquellos socios con cierto rango mínimo, gerenciando por consiguiente los casos más graves y con mayor probabilidad de impacto económico.

La complejidad en la medición de estos aspectos radica en que aún no es posible determinar la cantidad de pacientes que un coordinador puede gerenciar, ni así tampoco la respuesta real en el consumo de los pacientes frente a un seguimiento más puntual y focalizado en aquellos de mayor riesgo. Es decir, hasta el momento se desconoce si los pacientes gerenciados se encuentran en un estado de alta gravedad o no respecto a otros, ya que su evolución no es estudiada con datos actualizados, ni relevantes a la salud. Ello implica que el estudio económico realizado no necesariamente incluye los casos más riesgos en términos de salud y perspectivas de consumo. Es por ello que es imposible en este punto determinar un número que magnifique el beneficio, aunque es claro que el enfoque propuesto en este proyecto mejorará ampliamente los problemas actuales al incluir tanto las variables económicas que actualmente se utilizan, así como otras tantas que infieren el estado de salud del paciente.

#### **4.3.7.2 Costos del proyecto**

En lo que respecta a los costos del proyecto, estos son netamente humanos, siendo utilizados el analista de Programas Especiales y el Equipo de Coordinación de Cuidados Médicos. Los recursos utilizados en lo referente a software y hardware son nulos debido a que las herramientas de minería de datos a utilizar son gratuitas, y los sistemas informáticos ya se encuentran instalados.

#### **4.3.8 Objetivos y criterios de éxito de la minería de datos**

##### **4.3.8.1 Objetivos de la aplicación de minería de datos**

La aplicación de minería de datos para la resolución del problema estipulado tiene como objeto la estratificación de la población respecto a diversas variables médicas, biométricas, de salud, de consumo y económicas, a fin de detectar grupos con ciertos patrones propios que infieran el riesgo del paciente.

#### **4.3.8.2 Criterios de éxito de la aplicación de minería de datos**

Identificación de los grupos poblacionales o clusters dentro de la población de enfermos crónicos de OMINT S.A. Así mismo se deben establecer las particularidades de cada grupo identificado de un modo tal que pueda computársele un riesgo determinado.

#### **4.3.9 Etapas del proyecto**

El proyecto de minería de datos contará de las siguientes etapas:

##### **Etapas I – Relevamiento de los datos**

La etapa de relevamiento consiste en buscar dentro de las distintas fuentes de información de OMINT las variables determinadas a relevar.

Tiempo estimado de la tarea: 1.5 meses.

##### **Etapas II – Preparación de los datos**

Esta etapa consta de la adecuación de la información relevada previa a ser procesada.

Tiempo estimado de la tarea: 1 mes.

##### **Etapas III – Elección de la Herramienta**

La siguiente etapa trata la elección de la herramienta a utilizar para el proyecto, en lo que a minería de datos respecta.

Tiempo estimado de la tarea: 2 semanas.

##### **Etapas IV – Procesamiento de los datos**

Esta etapa supone el procesamiento de datos con la herramienta seleccionada.

Tiempo estimado de la tarea: 1.5 meses.

##### **Etapas V – Evaluación del Output**

La salida del procesamiento de los datos debe ser evaluada a fin de establecer si cumple con los objetivos y criterios de éxito estipulados previamente.

Tiempo estimado de la tarea: 2 meses.

### **Etapa VI – Determinación de los índices de riesgo**

El objetivo del negocio es la ponderación de un índice de riesgo. En este sentido deben desarrollarse los índices estipulados previamente.

Tiempo estimado de la tarea: 1 mes.

### **Etapa VII – Evaluación final del Scoring**

Tras la realización del scoring debe evaluarse su resultado por parte de los expertos, a saber, los coordinadores del ECCM.

Tiempo estimado de la tarea: 2 semanas.

## **4.3.10 Herramientas a utilizar**

Las herramientas que pueden ser utilizadas se listaron en el Inciso 2.8. En vista del objetivo de minería de datos para este problema particular, se utilizarán las técnicas SOM, en particular el algoritmo de Kohonen dada su sencillez y eficiencia de procesamiento. El mismo es un algoritmo iterativo que detecta similitudes entre los datos, asignando cada individuo a un cluster determinado. El proceso itera sucesivamente hasta que encuentre la convergencia o estabilidad en la asignación de los clusters a cada individuo. La delimitación de cada cluster se realiza a partir de un procedimiento de ponderación relativa, donde la asignación de cada individuo al cluster correspondiente se basa en el método de mínimos cuadrados al centroide, es decir, la menor distancia al centro de cada cluster determina el cluster al que pertenece. Una vez definido el cluster de pertenencia a cada individuo, se puede determinar las particularidades del cluster y asignársele un riesgo determinado.

## **4.4 Entendimiento de los datos**

### **4.4.1 Extracción de los datos**

Los datos recopilados de los sistemas de información de OMINT S.A. para cada paciente son los siguientes:

#### **I. Información demográfica**

##### **a. Edad**

- b. Sexo
- c. Lugar de residencia
- d. Plan adquirido

La información se encuentra dentro del Data Warehouse, en la tabla de “indicadores..benef\_a\_fecha\_completo\_hist”, y corresponde al período enero a diciembre 2005. Los queries utilizados para su obtención se detallan en el anexo 1.1.

## **II. Información de consumo**

- a. Cantidad de Cirugías
- b. Cantidad de Rehabilitaciones
- c. Cantidad de Prótesis y Ortésis
- d. Cantidad de Imágenes
- e. Cantidad de Medicación Consumida
- f. Cantidad de Prestaciones Consumidas
- g. Cantidad de Internaciones
- h. Cantidad de Consultas a especialistas

La información se encuentra dentro del Data Warehouse, en la tabla de “dw..dwv\_gastos”, y corresponde al período enero a diciembre 2005. Los queries utilizados para su obtención se detallan en el anexo 1.2.

## **III. Información de gasto**

- a. Gasto total en el período

La información se encuentra dentro del Data Warehouse, en la tabla de “dw..dwv\_gastos”, y corresponde al período enero a diciembre 2005. Los queries utilizados para su obtención se detallan en el anexo 1.3.

## **IV. Información Clínica**

- a. Tipo de Internación
- b. Causa de Internación



La información se encuentra dentro de los censos de internación del área de Auditoría Médica. Estos censos contienen datos referentes a las órdenes de internación, por lo cual debe cruzarse dicho dato con la información contenida en el data warehouse, a fin de obtener el afiliado correspondiente. Así mismo debe cruzarse la información referente a la causa de internación dada por el ICD-9, y obtener así la categoría genérica de causa de internación.

## V. Información de Salud

### a. Enfermedades Crónicas padecidas

Respecto al set de enfermedades crónicas seleccionado este fue determinado a partir de un intensivo estudio de la bibliografía médica, relevando aquellas enfermedades crónicas de mayor importancia en términos de impacto económico, severidad clínica y viabilidad de detección en los sistemas de información. A partir del estudio de “Health Care: Chronic condition list” [OpenPlan, 2005], se establecieron las categorías y enfermedades dentro de estas a ser relevadas, a saber (tabla 4.4.1):

Categorías Crónicas seleccionadas			
Alergias	Cáncer	Cardiovascular	Diabetes
Dislipidemias	Enfermed. Obstructivas vía aérea	Endocrina	Enfermedad ocular
Gastrointestinal	Glandular	HIV	Infecciosa
Músculo-Esquelética	Neuronal	Óseo / Articular	psiquiátricas
Discapacidad	Sangre	Inmunosupresores	Tratamiento dolor
Estados prohemorrágicos	Terapias de reemplazo hormonal	Renal	

Tabla 4.4.1 – Categorías Crónicas seleccionadas.

Si bien algunas no responden estados crónicos como lo es el cáncer y la gran mayoría de las enfermedades psiquiátricas, se consideró que dada la alta incidencia económica de sus tratamientos, no debían ser excluidas del estudio.

Así mismo, debe aclararse que si bien a priori ciertas categorías de las listadas parecieran que deben estar agrupadas, tal es el caso de las enfermedades endocrinas y la diabetes, se las ha considerado separadas dada la importancia de estas respecto al resto de las enfermedades que restan dentro de la categoría.

Determinadas las enfermedades crónicas, se deben detectar aquellos socios que las padecen. Si se relevan las fuentes de información con las cuenta el sistema prepago en la Argentina, la detección de los mismos no se evidencia a simple vista. Esta información se halla encubierta principalmente bajo los

consumos del rubro farmacéutico. Por consiguiente, se estipuló un sistema que permita inferir el consumo de medicación crónica.

Tal sistema se basa en la identificación de pacientes a partir del relevamiento de aquellas monodrogas y/o acciones farmacológicas incluidas en el manual farmacéutico<sup>2</sup> y empleadas para el tratamiento del set de enfermedades crónicas determinado. A partir de ellas se determinó que medicamentos aplicaban a los tratamientos seleccionados. El mayor problema que se presenta surge del hecho que algunos de los medicamentos tienen utilidades para tratamientos ambulatorios (tratamiento de enfermedades no crónicas). A fin de salvar esta problemática, se estudió paralelamente el patrón de consumo de la medicación por parte de pacientes, identificando que en promedio la mínima cantidad de prescripciones o compras anuales era de seis. Ello implica que cada un período de dos meses se realiza una compra de medicación. De esta manera, se evita la variación entre las unidades (cajas) que requiere cada medicamento particular para el tratamiento de cada una de las enfermedades especificadas.

Por otro lado, algunas enfermedades crónicas como el cáncer, discapacidad, HIV, Renal y diabetes, se detectaron por otro camino, dado que estas requieren medicación y/o tratamientos fuera del circuito de adquisición común al resto de los afiliados. Estos por ser medicaciones o tratamientos de alto costo, requieren una autorización especial, por lo cual la información sobre sus padecientes se encuentra disponible en los sistemas de información de la empresa, en particular dentro de los programas de Discapacidad, Diabetes, Renal, Cáncer y de Medicación del Alto Costo.

#### 4.4.2 Descripción de los datos

El relevamiento de los datos arrojó 20.693 socios de OMINT S.A. con una o más enfermedades crónicas. Los campos resultantes para cada socio fueron los siguientes (tabla 4.4.2):

Variable	Tipo de dato / unidad
Edad	años
Sexo	F o M
Provincia	categoría
Línea	categoría
Q_Cirugias	unidades
Q_Protesis	unidades
Q_Imag	unidades
Q_Rehab	unidades

<sup>2</sup> Este se denomina ManualDat dentro de los círculos de la salud.

Q_ Consultas	unidades
Q_Medic	unidades
Q_Internaciones	unidades
ICD-9 Categ	categorica
Tipo Internación	categorica
Importe	pesos
Q_Total_Consumida	unidades
Cantidad de Enfermedades Crónicas	unidades
Cardiovascular	binaria
Inmunosupresores	binaria
Terapias de reemplazo hormonal	binaria
Alergias	binaria
Dislipidemias	binaria
Efermed Obstructivas vía aérea	binaria
Enfermedad ocular	binaria
Diabetes	binaria
Gastrointestinal	binaria
Psiquiatritas	binaria
Tratamiento dolor crónico	binaria
Endocrina – Glandular	binaria
Músculo - Óseo – Esquelética	binaria
Sangre - Estados Prehemorrágicos	binaria
Discapacidad	binaria
Neuronal	binaria
HIV	binaria
Cáncer	binaria
Renal	binaria
Infecciosa	binaria

Tabla 4.4.2 – Variables relevadas

En cuanto a las variables categóricas, la identificación de cada una de ellas se lista en el anexo 2.

Respecto a las razones del relevamiento específico de cada variable son las siguientes:

- **Información demográfica:** El sexo y edad son variables que influyen en el riesgo del paciente, siendo a su vez características necesarias para la segmentación de grupos poblacionales. La edad aumenta el riesgo de sufrir complicaciones en la salud debido al deterioro natural del organismo, y así mismo aumenta el impacto económico para la prepaga. Numerosos estudios destacan la importancia de este factor, donde, a modo de ejemplo, la población enrolada mayor a 65 años, la cual representa el 12.5% del total, gasta el 43% de los recursos [Wong *et al*, 2005]. A su vez, el aumento de la edad favorece la aparición de patologías crónicas, tales como diabetes, osteoporosis, problemas cardiovasculares, entre otros [Wong *et al*, 2005]. Con respecto al sexo, el sexo femenino es más propenso a

contraer enfermedades crónicas, mientras que el masculino a sufrir fatalidades [Wong *et al*, 2005]. Así mismo el sexo femenino tiende a consumir más medicación que el masculino, siendo importante el gasto en farmacia [Wong *et al*, 2005].

Por otro lado la información referente a la línea y provincia son variables socioeconómicas que pueden representar su utilidad en la segmentación.

- **Información de consumo:** La cantidad de prestaciones y tipo de prestaciones utilizadas por cada afiliado es necesaria para evaluar el patrón de consumo de cada grupo poblacional. Ciertos consumos sin indicativos del cuidado preventivo del sujeto, mientras que los excesos denotan una complicación en su estado de salud. Así mismo la utilización de servicios sanatoriales es indicador de un suceso adverso en la salud del afiliado e indica un impacto considerable en riesgo [Kings Fund, 2005].
- **Información de gasto:** Los especialistas proyectan que en los EEUU el gasto en salud será del 17.7% del PBI para el 2010, un 41% más que en 2002 [Moroney, 2003]. Este número refleja el fuerte aumento tanto en los precios de los servicios, como también en la utilización de los mismos por parte de los afiliados. El incremento en la utilización se encuentra determinado por el envejecimiento poblacional, el descubrimiento de nuevas terapias y una mayor demanda por parte de los usuarios. Todo esto se traduce en un aumento en el costo para las prepagas, que impacta directamente en su rentabilidad. En lo que al afiliado respecta, el gasto del mismo infiere el tipo de recursos que este está consumiendo, y a mayor costo de estos mayor la complejidad del tratamiento y por ende de la enfermedad. Así mismo este valor es un claro indicador de la priorización sobre su gerenciamiento por parte del prepago.
- **Información Clínica:** El impacto económico de la internación ronda el 30% del total [Moroney, 2003] hacia el año 2001. Independientemente de ello, las causas y tipos de internación dan pauta de la gravedad clínica del afiliado. El estado de salud frente a un evento de internación se ve comprometido, y conocer la causa y el tipo de la misma es información valiosa para el gerenciamiento del afiliado dentro de los programas de Disease Management.

Por otro lado, el padecimiento enfermedades crónicas conlleva un riesgo inherente de sufrir una internación. La probabilidad de sufrir este evento puede extraerse a partir de estudios biométricos en la población analizada. Así mismo con esta información puede discriminarse el riesgo inherente de

internación causado por la cantidad y combinación de enfermedades crónicas que un individuo presente. A modo de ejemplo, es mayor el riesgo de internación de un sujeto con enfermedades cardiovasculares y de dislipidemias, frente a otro sujeto con enfermedades cardiovasculares y psiquiátricas.

- **Información de Salud:** Las enfermedades que padece el afiliado son un valioso dato para inferir la prioridad del mismo sobre otro, así como para delimitar coherentemente el riesgo de este frente a otro individuo. En este sentido, no necesariamente un sujeto que tenga cierto patrón de consumo y con cierto tipo de enfermedad puede tener el mismo riesgo que otro con similar patrón de consumo, peor otro tipo de enfermedad.

#### 4.4.3 Exploración de los datos

Estudiando la información recopilada y evaluando su futura utilidad para determinar el riesgo potencial de cada paciente, se determina que el scoring debe estar definido por una serie de índices complementarios entre sí que evalúen el impacto de diferentes aspectos clínicos, demográficos, económicos, de salud y de consumo. El proyecto inicial suponía la inferencia del riesgo a partir de la utilización directa de minería de datos. Sin embargo, evaluando los datos recopilados, se evidencia que la utilización de las técnicas SOM por sí solas arrojarán resultados parciales que pueden ser complementados por otro lado. En este sentido, las técnicas SOM serían empleadas para dar lugar a la segmentación de la población identificando diferentes clusters que por sí solos infieran un riesgo particular. Las variables que se utilicen para esta técnica deben ser variables globales, que puntualicen ciertos aspectos diferenciales entre los grupos o cluster. En lo que a ello respecta, ciertas variables como lo son el tipo de internación o la causa de internación son variables puntuales de cada individuo, por lo cual no deben ser incluidas al aplicar esta técnica. No obstante, este tipo de variables contiene información sumamente útil para el gerenciamiento de los programas de Disease Management, ya que son manifestaciones del riesgo clínico o de salud. Por ende, se determinarán una serie de índices complementarios, dentro de los cuales se encontrará un índice referente al riesgo del grupo poblacional determinado a partir de las técnicas SOM.

#### 4.4.4 Calidad y precisión de los datos

La calidad de los datos relevados es sumamente alta, excepto en aquellos datos extraídos de los censos de internación y aquellos inferidos por el consumo de medicación.

Las bases de datos extraídas del data warehouse son bases crudas, pero con datos pre-filtrados, por lo cual no existen errores de formato o duplicaciones.

En cuanto a los datos de los censos de internación, estos presentan problemas de formato y en algunos casos de error de entrada. Estos últimos no permiten la correcta identificación del socio al suceso de internación, debiendo buscarse manualmente. Respecto al formato, los datos de los ICD-9 se encuentran varias causas para un mismo paciente, por lo cual se opta por seleccionar el primer registro.

Respecto la enfermedad crónica de cada paciente inferida a partir del consumo de medicación, puede haber un sesgo menos en el sistema de detección utilizado.

### 4.5 Preparación de los datos

#### 4.5.1 Selección de los datos

La identificación de los cluster o subgrupos poblaciones se realizó mediante la utilización de una herramienta de SOM denominada NNClust. Esta herramienta al ser una macro del Excel, presenta un cierto impedimento respecto a la eficiencia de procesamiento de masas de datos de gran tamaño. Por consiguiente, dado que la cantidad de enfermos crónicos detectados a partir de los criterios establecidos en la etapa II era de más de 20.000 casos, se tomo una muestra aleatoria de la misma de 8000 casos, validando la correspondencia estadística entre ambas poblaciones<sup>3</sup>. La tabla 4.5.1 contigua detalla el contraste porcentual entre la muestra y la población.

		Población	Muestra
n		20692	8000
Edad	Media	52.3	52.2
	Desvío	20.6	20.5
Sexo	%	59.5%	59.5%
Cantidad de	Media	1.95	1.97

<sup>3</sup> La correspondencia estadística puede ensayarse mediante un test de igualdad de medias para cada una de las variables. Sin embargo debido al alto tamaño muestral ello pierde sentido. De todos modos, a fines de justificar la aseveración, se realizó cada uno de los test obteniendo un p-value promedio de 0.0002.

Enfermedades	Desvío	1.3	1.3
Internaciones		13,4%	13,7%
Enfermedades	Cardiovascular	36,8%	36,7%
	Inmunosupresores	0,2%	0,2%
	Terapias de reemplazo hormonal	4,2%	4,2%
	Alergias	4,2%	4,2%
	Dislipidemias	19,1%	19,2%
	Efermed Obstructivas vía aérea	4,7%	4,9%
	Enfermedad ocular	5,7%	5,4%
	Diabetes	8,1%	8,3%
	Gastrointestinal	11,0%	11,1%
	Psiquiatrias	27,5%	28,4%
	Tratamiento dolor crónico	0,0%	0,0%
	Endocrina – Glandular	16,3%	16,0%
	Músculo - Óseo – Esquelética	8,5%	8,9%
	Sangre - Estados Prehemorrágicos	6,9%	6,9%
	Discapacidad	2,3%	2,7%
	Neuronal	22,0%	21,8%
	HIV	0,5%	0,4%
	Cáncer	3,1%	3,1%
Renal	0,3%	0,2%	
Infecciosa	0,1%	0,1%	

Tabla 4.5.1 – Correspondencia porcentual población-muestra.

#### 4.5.1.1 Variables incluidas en la estratificación

De acuerdo con el objetivo de estratificación dentro de la población de enfermos crónicos, se debió seleccionar un set de variables que a priori presenten un sentido lógico y no causal. En este sentido, las variables seleccionadas debían ser consistentes en el hecho de representar los diferentes aspectos establecidos en base a términos:

1. Económicos
2. de Consumo
3. de Salud
4. Demográficos

Por consiguiente se determinaron las siguientes variables de acuerdo a aspectos de:

##### 1. Económicos:

- i. - Gasto en el período considerado

El impacto económico de la utilización de prestaciones por parte del paciente queda determinado por el gasto del mismo, siendo a su vez una

variable representativa y diferencial de los distintos estratos poblacionales. La estratificación poblacional en base al gasto obedecerá al objetivo de la identificación de aquellos grupos de mayor impacto económico para OMINT S.A.

## 2. de Consumo:

- i. – Cantidad de medicación consumida
- ii. – Cantidad de cirugías
- iii. – Cantidad de prótesis
- iv. – Cantidad de rehabilitaciones
- v. – Cantidad de imágenes
- vi. – Consultas a especialistas
- vii. – Cantidad de Internaciones

La segmentación poblacional se basa “en agrupar a los individuos del mercado en grupos similares por sus hábitos, de tal manera que exijan un tratamiento operacional diferente y diferenciado”, [Grimaldi, 1996], y por consiguiente las variables de consumo son uno de los aspectos más relevantes. Así mismo, la identificación del patrón de consumo de cada cluster cumplimenta los objetivos del proyecto en lo referente a las mejoras de la información brindada a los equipos de coordinación para el gerenciamiento.

## 3. de Salud:

- i. – Tipo de enfermedades

Alergias	Cáncer	Cardiovascular	Diabetes
Dislipidemias	Efermed. Obstructivas vía aérea	Endocrina	Enfermedad ocular
Gastrointestinal	Glandular	HIV	Infeciosa
Músculo-Esquelética	Neuronal	Óseo / Articular	Psiquiátricas
Discapacidad	Sangre	Inmunosupresores	Tratamiento dolor
Estados prohemorrágicos	Terapias de reemplazo hormonal	Renal	

Tabla 4.5.1.1 – Tipo de enfermedades seleccionadas

El objeto de aplicación del proyecto en lo que a minería de datos respecta, supone la estratificación de la población. Ciertamente una de las variables diferenciales de mayor peso entre sujetos es su condición de salud. En este estudio la condición de la misma viene dada por las diferentes enfermedades que el sujeto pueda presentar (ver tabla 4.5.1.1).



Por otro lado, la identificación de los grupos en base a las enfermedades que exhiban brindará una base necesaria para la posterior ponderación del riesgo de acuerdo a los recursos poseídos por los programas de Disease Management de OMINT S.A.

#### 4. Demográficos:

i. Edad

ii. Sexo

Las variables demográficas son variables distintivas entre los grupos y deben incluirse. La distinción de los clusters de acuerdo a la edad y sexo son necesarios al momento de la ponderación del riesgo, justificación que se abordará en el inciso 4.6.3 y 4.6.4.

#### **4.5.1.2 Variables excluidas de la estratificación**

Cierto tipo de variables fueron excluidas durante la aplicación de minería de datos por objeto de las razones detalladas en el inciso 4.3 y 4.4. En lo que a ello respecta, la segmentación de la población requiere la utilización de variables globales, y diferenciales en un amplio aspecto. Las variables listadas a continuación debieron ser prescindidas por esta razón.

#### **1. Demográficos**

##### **i. Lugar de residencia**

Su exclusión se debe a que la población de afiliados de OMINT S.A. se encuentra el 78.8 % en Capital Federal y Buenos Aires, y un 13% en Córdoba, y por tal motivo el carácter diferencial de ello no compete a esta variable.

##### **ii. Plan adquirido**

Si bien el tipo de plan es indicador del nivel socioeconómico del afiliado, en el caso de OMINT, el costo de los planes médicos no es tal por objeto de ser planes de alto costo. Ello implica que los afiliados corresponden a un nivel económico alto, siendo por consiguiente su diferencia poco significativa.

## 2. Clínicos

Las variables clínicas fueron excluidas en su totalidad debido a que no compete segmentar por variables propias de un evento azaroso como lo es una internación. Este evento no compete a todo individuo, y no es propio de una condición particular de todos los entes dentro de un grupo. Este evento es causa de variables particulares de cada sujeto, por lo cual no corresponde utilizarlas como variable de diferenciación.

### 4.5.2 Ajuste de formato de los datos

Los datos seleccionados no presentaron reparos en lo que respecta a su ajuste previo al procesamiento.

### 4.5.3 Construcción de los datos

La única variable que fue adicionada a las listadas en el punto 3.1 fue la “Cantidad de enfermedades crónicas”. Esta variable surge de sumar la cantidad de enfermedades que cada individuo presenta. La razón de incluir este número es su importancia como parámetro de distinción poblacional, más allá del tipo de enfermedad que se presente.

### 4.5.4 Integración de los datos

Este inciso se refiere al cruce de información entre diferentes tablas o bases de datos. Para el proyecto en cuestión los cruces de información se estipulan en los queries anexos, y su objeto es combinar la información de un mismo afiliado que proviene de diferentes fuentes, obteniendo por consiguiente una fila por cada afiliado, con sus respectivos datos encolumnados. A modo de resumen las tablas combinadas fueron las siguientes:

- indicadores..benef\_a\_fecha\_completo U Dw..dwv\_gastos
- indicadores..benef\_a\_fecha\_completo U Dw..Especialidades
- indicadores..benef\_a\_fecha\_completo U Dw..lineas
- indicadores..benef\_a\_fecha\_completo U Dw..prestaciones

### 4.5.5 Formato de los datos

Las variables que debieron ser recodificadas previo al procesamiento en la herramienta NNClust fueron las siguientes:

- Edad: se codificó en los siguientes rangos etáreos.

Rangos Etareos	
Rango	Id
. 0 a 1	1
. 2 a 14	2
. 15 a 35	3
. 36 a 50	4
. 50 a 65	5
. +65	6

Tabla 4.5.5-1: Categorías etareas.

- Sexo: La codificación es la siguiente.

Sexo	Id
M	0
F	1

Tabla 4.5.5-2: Categorías de sexo.

- Cantidad de medicación consumida

Q_Medic	
Rango	Id
.0-3	0
.4-12	1
.13+	2

Tabla 4.5.5-3: Categorías de cantidad de medicación consumida.

- Cantidad de prótesis

Q_Prótesis	
Rango	Id
.0	0
.>0	1

Tabla 4.5.5-4: Categorías de cantidad de prótesis.

- Cantidad de rehabilitaciones

Q_Rehab	
Rango	Id
.0	0
.>0	1

Tabla 4.5.5-5: Categorías de cantidad de Rehabilitación.

– Cantidad de imágenes

Q_Imágenes	
Rango	Id
.0	0
.1	1
.2 a 4	2
>4	3

Tabla 4.5.5-6: Categorías de cantidad de Imágenes.

– Consultas a especialistas

Q_consultas	
Rango	Id
de 0	0
de 1-4	1
de 5-10	2
de 11-19	3
de 20-39	4
de +39	5

Tabla 4.5.5-7: Categorías de cantidad de consultas.

– Cantidad de Internaciones

Q_Internaciones	
Rango	Id
0	0
1 a 2	1
3 a 4	2
5 o más	3

Tabla 4.5.5-8: Categorías de cantidad de internaciones.

## 4.6 Modelización

### 4.6.1 Selección de la técnica de modelizado

En el inciso 4.3.10 del presente proyecto se estableció la conveniencia de la utilización de la técnica SOM (Self Organizing Maps), a través de la herramienta NNClust. El SOM es un tipo único de red neuronal en el sentido de que construye un mapa que preserva la topología, desde el espacio multidimensional a unidades de mapa de tal forma de que las distancias relativas entre puntos (datos) se conservan. Las unidades de mapa, o neuronas, generalmente forman un entramado regular bidimensional, donde la ubicación de una neurona contiene información semántica, El SOM puede entonces servir como una técnica de agrupamiento, o “clustering”, de datos multidimensionales.

## 4.6.2 Diseño de experimentos

Sobre la masa de datos original, se seleccionaron aleatoriamente 8000 casos para su procesamiento con la herramienta estipulada, tal como se estipula en el inciso 4.5.1. El entrenamiento de la herramienta con los datos es un parámetro estipulado antes del procesamiento, y fue fijado en 50 ciclos. Respecto a la evaluación del modelo, se seleccionaron 1069 casos sobre los que se tenía información previa sobre su estado de salud y parámetros de consumo en base al seguimiento frecuente de las coordinadoras de los programas de Disease Management de OMINT S.A., validando la sensibilidad del riesgo estipulado por el modelo.

## 4.6.3 Descripción del modelo

### 4.6.3.1 Seteo de parámetros

Los parámetros seteados en la herramienta fueron los siguientes:

- Cantidad de neuronas: 16

Este número se estableció con objeto de lograr una discriminación moderada de grupos poblacionales.

- Ciclos de entrenamiento: 50

Debido a la sobrecarga de procesamiento, este es un valor adecuado para racionalizar el tiempo empleado en el mismo.

- Parámetro inicial de aprendizaje: 0.9
- Parámetro final de aprendizaje: 0.1
- Decaimiento en el aprendizaje: Exponencial

Todos estos responden a seteos estándar de la herramienta.

### 4.6.3.2 Descripción del modelo

La herramienta seleccionada presenta un algoritmo iterativo que va detectando similitudes entre los datos, asignando cada individuo a un cluster determinado. El proceso itera sucesivamente hasta encontrar la convergencia o estabilidad en la asignación de los clusters a cada individuo. La asignación de cada individuo al cluster correspondiente se basa en el método de mínimos cuadrados ponderados al centroide, es decir, la menor distancia al centroide de cada cluster determina el cluster al que pertenece.

En cuanto a la cantidad de clusters que pueden presentarse en una población, la herramienta permite asignar una cierto límite a los mismos. Si bien ello supone forzar la capacidad de discriminación, debe tenerse en cuenta que si la premisa es la detección de grupos diferenciales dentro de una masa poblacional, es preciso delimitar la cantidad de subgrupos a un tamaño razonable. Desde ya que lo ideal en lo que se refiere a “discriminación” sería asignar un cluster a cada individuo. Ello desde el punto de vista de la clasificación y selección es sumamente válido, pero es irracional desde el punto de vista práctico, ya que tener 20.000 clusters no tiene ninguna utilidad. Por consiguiente, se debe establecer con criterio una cierta cantidad de clusters que sean manejables para la aplicación particular de cada problema. En este caso, dado que la cantidad de variables es importante, y así mismo el tipo de aplicación exige que la discriminación entre individuos sea elevada, se establecieron 16 clusters iniciales (Cluster PP), tal como se definió en los parámetros iniciales del inciso anterior. Es decir, se forzó a la herramienta a iterar y converger hasta que la masa poblacional asigne a cada individuo a uno de los 16 clusters establecidos. La cantidad de individuos asignados a cada cluster fue la siguiente (ver matriz 4.6.3.2-1):

<b>Cluster 1</b>	<b>Cluster 2</b>	<b>Cluster 3</b>	<b>Cluster 4</b>
813	1726	142	734
<b>Cluster 5</b>	<b>Cluster 6</b>	<b>Cluster 7</b>	<b>Cluster 8</b>
11	169	20	16
<b>Cluster 9</b>	<b>Cluster 10</b>	<b>Cluster 11</b>	<b>Cluster 12</b>
146	177	371	574
<b>Cluster 13</b>	<b>Cluster 14</b>	<b>Cluster 15</b>	<b>Cluster 16</b>
254	266	461	2119

Matriz 4.6.3.2-1 – Cantidad de individuos por cluster principal.

, siendo su proyección topográfica (ver figura 4.6.3.2):

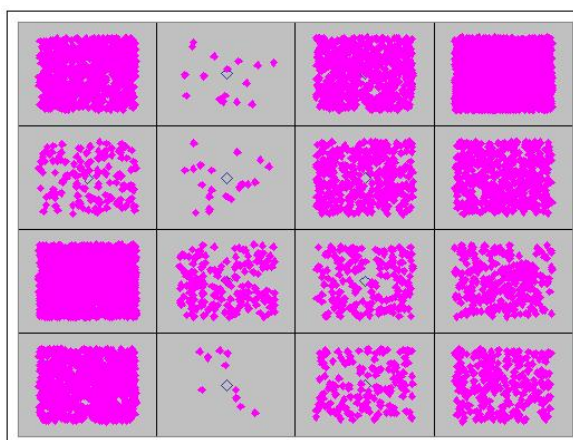


Figura 4.6.3.2 – Representación de los clusters obtenidos

Sin embargo en este punto se observó un problema de concepto devenido de la falta de discriminación que el resultado obtenido presentaba. Se evidenciaba que cada uno de los 16 clusters resultantes eran sumamente genéricos. Esto implica que entre ellos se diferenciaban por macro-aspectos. Este concepto se establece a partir que cada uno de los clusters detectados presenta ciertas características inherentes que lo definen. Ello implica que por ejemplo el Cluster X es una población que “en promedio” presenta cierto rango etáreo, sexo, cierto tipo de enfermedades y con cierto patrón de consumo y gasto. Por consiguiente es claro que puede asignársele un cierto riesgo a cada cluster de acuerdo a una serie de aspectos preestablecidos. Sin embargo, existe una fuerte razón que impone que se establezca un criterio aún más discriminatorio entre los subgrupos, y ello viene dado por el tipo de variables utilizadas. El concepto de individuos para este proyecto son personas, y las variables utilizadas sobre las mismas son en su mayoría demográficas y de salud. Estas últimas son variables sumamente sesgadas e influenciadas por infinidad de causas. En términos médicos, una misma afección de salud en dos personas puede deberse a causas distintas, con síntomas múltiples. En consecuencia, puede haber dos sujetos con el mismo set inicial de enfermedades, sexo, edad, y mismo estado de salud, y sin embargo, el gasto y consumo de ambos puede ser totalmente diferente. Esta particularidad puede ejemplificarse claramente si extrapolamos este concepto al caso de la clasificación del reino animal, donde encontraremos las clases y subclases. En resumen, la primera etapa de clasificación delimita a la población de acuerdo a macro-aspectos. En consecuencia, se establece una segunda etapa de modelizado que corresponde a una apertura adicional dentro de cada cluster, a fin de detectar micro-aspectos.

La cantidad de clusters secundarios (Cluster Secc) en este caso es variable, y se detalla a continuación (ver matriz 4.6.3.2-2):

Cluster PP	Cluster Secc
1	9
2	16
3	9
4	9
5	1
6	9
7	3
8	1
9	3
10	4
11	9
12	7
13	9
14	9
15	16
16	16

Matriz 4.6.3.2-2 – Cantidad de clusters secundarios en cada cluster primario.

Por consiguiente, la población fue estratificada en 130 subgrupos determinados por 16 grupos genéricos. Las características de cada grupo y subgrupo se detallan en el Anexo 3.1.

#### **4.6.4 Interpretación del modelo**

El modelo obtenido en dos etapas discrimina sensiblemente los grupos y subgrupos dentro de la población de enfermos crónicos de OMINT S.A. Si se evalúa detenidamente cada uno de ellos, se observa que la selección de variables ha sido correcta, en cuanto permite una ponderación de riesgo a cada subgrupo. Esta ponderación o Riesgo Cluster, objetivo final del proyecto en su etapa de minería de datos, surge de relevar la incidencia de la combinación de las diferentes variables demográficas, de consumo y de salud. Así mismo, el resultado del modelo manifiesta el patrón de consumo de cada subgrupo, evidenciando sobre cuales subgrupos se deben focalizar las acciones que tiendan a lograr una utilización más racional de los recursos.

Por otro lado, al modelo obtenido deben imputársele otra serie de riesgos complementarios tal como fue expuesto en el inciso 4.3.5. La razón de ello radica en que es posible y necesario un detalle más preciso sobre la condición de cada paciente a fin de efectivizar el gerenciamiento de estos. Estos índices complementarios tienen un fundamento médico-económico sumamente específico, y serán detallados en los incisos siguientes.

##### **4.6.4.1 Interpretación y determinación del Riesgo Cluster**

Una vez estratificada la población por medio de mapas autoorganizados, se le asigna a cada cluster o subgrupo un riesgo determinado en base a las particularidades del mismo. A continuación se detallan los pasos establecidos para la determinación del riesgo cluster.

###### **4.6.4.1.1 Categorías de puntuación**

La puntuación de cada cluster, se basó en el riesgo establecido en base a las siguientes categorías (ver tabla 4.6.4.1.1):



Categorías evaluadas			
Consumo	Gasto	Edad	Salud

Tabla 4.6.4.1.1 – Categorías de riesgo ponderadas

Estas categorías abarcan todo el espectro de variables empleadas durante la estratificación. Luego cada categoría es puntuada porcentualmente en razón de la Alta Gravedad (High risk), Mediana Gravedad (Médium risk) o Baja Gravedad (Low risk) que presentase.

#### 4.6.4.1.2 Criterio de clasificación de categorías

El criterio empleado para la clasificación de cada cluster, es decir, para la imputación de la alta, media o baja gravedad, se basó en el contraste de cada cluster secundario contra el valor promedio del total de clusters del grupo. Es decir, se lo contrastó contra su cluster principal. Así mismo, la calificación de Alta, Media o Baja gravedad para cada categoría, se estableció a partir del siguiente criterio:

- Alta gravedad:  $> 1.5 \sigma$  de  $\mu$
- Media gravedad: entre  $\pm 1.5 \sigma$  de  $\mu$
- Baja gravedad:  $< 1.5 \sigma$  de  $\mu$

donde  $\mu$  representa a la media del cluster principal, y  $\sigma$  al desvío estándar del mismo.

#### 4.6.4.1.3 Criterio de puntuación de categorías

Posteriormente se debió asignar un puntaje al riesgo de cada categoría. La referencia de puntuación de acuerdo al riesgo se establece en la siguiente matriz 4.6.4.1.3:

	Q consumida	Q Importe	Q Edad	Q enf
High	100%	100%	100%	100%
Med	50%	50%	50%	50%
Low	0%	0%	0%	0%

Matriz 4.6.4.1.3 – Porcentajes de riesgo asignados por categoría.

A modo de ejemplo, si la categoría presenta un riesgo alto, se le asignará un valor del 100%, si es medio del 50% o bajo del 0%.

#### 4.6.4.1.4 Criterio de puntuación de categorías

A continuación se asignó el riesgo a las 81 combinaciones posibles dadas por la matriz 4.6.3.2-2. Ciertas combinaciones no competen al no haberse detectado un cluster con tales características dentro de la población. En el anexo 3.2 se detalla la ponderación de la gravedad de cada cluster.

#### 4.6.4.1.5 Importancia relativa de la categoría

Finalmente debe computarse un solo valor para cada cluster que unifique las cuatro categorías. Por consiguiente se estableció una ponderación ad-hoc de pesos relativos a saber (ver tabla 4.6.4.1.5):

Peso relativo			
Q consumida	Q Importe	Q Edad	Q enf
10%	40%	15%	35%

Tabla 4.6.4.1.5 – Ponderación final de cada categoría.

Los pesos relativos utilizados para la ponderación de las categorías de riesgo fueron establecidos de acuerdo a la premisa que el importe o gasto, y así tanto la enfermedad, son los dos parámetros más distintivos entre cada subgrupo. Los criterios lógicos para la elección del valor ponderado son los siguientes.

**4.6.4.1.5.1 Ponderación del consumo:** La evidencia de los datos demuestra que el consumo es una variable de moderada significancia, en cuanto el alto o bajo consumo no es una variable sobre la cual las Coordinadoras de los Programas de Coordinación de Cuidados Médicos de OMINT S.A. deban, ni pueden influir, ya que ese número abarca todo tipo de consumo que el individuo haya hecho. Ello implica que si recibe gran cantidad de prescripciones de medicación debido a que la misma trae pocas unidades por envase; o por ejemplo realiza chequeos de laboratorio completos que implican decenas de estudios, la cantidad consumida será importante. Esta es la razón por la cual se ha considerado que su peso relativo debe ser moderado.

**4.6.4.1.5.2 Ponderación de la edad:** En cuanto a la edad, la misma si bien es distintiva de cada cluster, no debe ser considerada prioritaria para definir el riesgo del grupo. En este sentido, si su prioridad fuese alta, se estaría sobrevalorando a los grupos etéreos más grandes, hecho obvio desde el punto de vista que la edad

indefectiblemente aumenta el riesgo. Ello generaría un sesgo que impediría focalizarse sobre grupos de menor edad y similar riesgo.

**4.6.4.1.5.3 Ponderación de la Cantidad de enfermedades:** Con respecto a la cantidad enfermedades, su ponderación es alta por inferir gravedad en la condición del grupo: a mayor cantidad de enfermedades, mayor riesgo.

Puede objetarse el hecho de no ponderar cada cluster por el tipo de enfermedades que lo definen. Sin embargo, la razón de no hacerlo radica en que si bien es una condición que distingue a un cluster de otro, la relación entre consumo, edad e importe se encuentran correlacionadas con las enfermedades que lo especifican. Así mismo, debe entenderse que de establecer criterios de puntuación en base a la enfermedad o enfermedades que definen cada cluster, puede sesgarse nuevamente el resultado al considerar ciertas enfermedades más importantes que otras.

Por otro lado, considerando el funcionamiento de las redes neuronales, ponderar el tipo de enfermedades específicas de cada cluster, sería en definitiva ponderar un promedio sobre las enfermedades de los individuos que conforman cada subgrupo. Ello generaría un grosero error en la estimación del riesgo, ya que habrá sujetos que padezcan cierta enfermedad, y otros que no; en consecuencia, el resultado del riesgo ponderado sería más bien macro. Es decir, si la utilización fuera simplemente puntuar el riesgo de cada cluster respecto de otro, se le agrega a la ponderación de los mismos un score similar al utilizado en el inciso 4.6.4.6 para la puntuación de la enfermedad y combinación crónica, y arrojará con menos detalle cuales clusters son más riesgosos que otros. Por consiguiente, como el objetivo de un scoring es poder determinar un riesgo de alta sensibilidad a las particularidades individuales, se decidió ponderar las enfermedades por fuera del cluster, logrando así un resultado más preciso y detallado.

**4.6.4.1.5.4 Ponderación del costo:** La alta incidencia del costo se debe principalmente a la importancia que tiene el factor económico implica en la aplicación de este scoring. Ello supone que en vista que el campo de aplicación de este proyecto será para una empresa de medicina prepaga que busca la maximización de sus recursos en pos de una mejor rentabilidad y posición competitiva en términos de cuidado a sus afiliados, el costo debe ser el factor distintivo que indique el riesgo latente entre los distintos clusters.

#### **4.6.4.1.6 Conclusiones sobre el Riesgo Cluster**

Basado en la estratificación de la población por medio de redes neuronales, pondera el riesgo para el paciente combinando variables demográficas (según edad y sexo), variables referentes a enfermedades crónicas y tipo de enfermedades, variables de consumo y de gasto. En este sentido, la combinación de todas estas variables determina distintos grupos o “Clusters” dentro de la población, los cuales presentan un patrón de consumo y enfermedades que tienen un riesgo asociado particular que puede ser ponderado por medio del cálculo descrito.

#### **4.6.4.2 Interpretación del riesgo por Internación**

La internación es un evento de suma importancia clínica que necesariamente debe ser considerado. La implicancia de este suceso conlleva una alta gravedad en términos económicos y de salud para el paciente, y por consiguiente debe ser ponderado. Las fuentes de información relevadas de los censos sanatoriales diarios, proveen datos de alta significancia que pueden parametrizarse.

Este índice pondera el riesgo causado por la cantidad de internaciones, ajustando tal valor por cuatro riesgos aditivos dados por:

##### **a - Cantidad de internaciones:**

La cantidad de reinternaciones en el período infiere de algún modo el estado de salud del paciente. A mayor cantidad de internaciones es presumible un estado de salud más grave, y es por ello que debe ponderarse el riesgo creciente frente a la reinternación.

##### **b - Probabilidad de internación dada la combinación de enfermedad crónica.**

El padecimiento de enfermedades crónicas conlleva un riesgo inherente de sufrir una internación. La probabilidad de sufrir este evento puede extraerse a partir de estudios biométricos en la población analizada. La razón de ajustar el índice por esta variable, se basa en que es necesario discriminar el riesgo inherente de internación causado por la cantidad y combinación de enfermedades crónicas que un individuo presente. A modo de ejemplo, es mayor el riesgo de internación de un

sujeto con enfermedades cardiovasculares y de dislipidemias, frente a otro sujeto con enfermedades cardiovasculares y psiquiátricas.

#### **c - Gravedad del suceso en base al ICD-9<sup>4</sup>.**

Este ajuste implica considerar la causa de la última internación. Esta información es sumamente valiosa en cuanto a que permite conocer la causa clínica que llevó al paciente a internarse.

Ciertamente se debe ponderar un riesgo diferencial para cada tipo de enfermedad. Estas vienen dadas de acuerdo a una codificación internacional denominada International Classification of Diseases (ICD-9), que clasifica las causas de internación en 19 subgrupos genéricos. Por consiguiente, se ponderó un riesgo ad-hoc en base a la capacidad de acción del equipo de Coordinación de Cuidados Médicos (CCM) y a los programas de Disease Management implementados por OMINT.

#### **d - Tipo de intervención (Quirúrgica, parto, quimioterapia, etc.).**

El riesgo de internación debe reflejar el motivo de la misma. En consecuencia se le debe ajustar un riesgo por el tipo de intervención que sufrió el paciente. Tal ponderación se efectuó en base a la capacidad de acción del equipo de Coordinación de Cuidados Médicos (CCM).

A modo de ejemplo, el riesgo de un paciente intervenido por un parto es menor que aquel intervenido por una cirugía cardiovascular.

#### **4.6.4.2.1 Factores del riesgo por Internación**

La determinación del riesgo de cada factor listado se detalla a continuación. Sin embargo antes de proseguir debe aclararse que la metodología implementada para determinar cada uno de ellos es puede parecer subjetiva. En este sentido, debe aclararse que la estipulación de un riesgo no es tarea sencilla, ya que no existe una forma correcta o incorrecta de hacerlo, sino simplemente una manera más racional. El ajuste de riesgos es la es la parte inevitable de cualquier proceso que implique la toma de decisiones, y como tal, esta sujeto a subjetividades en su modelizado. Ciertos riesgos son simples de modelizar debido a la pequeña cantidad de interacciones que afectan su entorno, pero otros, especialmente los riesgos ligados a ciencias médicas, son complejos por objeto de la gran cantidad

---

<sup>4</sup> ICD-9: International Classification of Diseases

de variables que influye directa o indirectamente sobre el entorno. Es por ello que el lector evidenciará la simplicidad de algunos, y en otros deberá aceptar las razones fundamentadas para su razón.

**a - Cantidad de internaciones:** El riesgo asociado al evento de reinternación sobre el estado de salud del paciente debió ser establecido a medida, debido a que no se encontró investigación alguna en la literatura médica que indique una metodología puntual para su cálculo. Sí bien existen numerosos estudios que determinan modelos predictivos para cierto tipo de enfermedad o factores de riesgo asociados bajo ciertas condiciones que derivan en la reinternación, no se adecuan a los objetivos de este proyecto. Por consiguiente, se estableció que el riesgo responde a una ecuación asintótica dada por

$$R = \frac{Q-1}{Q} \quad (4.6.4.2.1)$$

donde R es el riesgo computado y Q la cantidad de reinternaciones del paciente. La tabla de riesgo 4.6.4.2.1-1 arroja la siguiente evolución:

Aditividad por internación	
Q Internaciones	% extra de riesgo
1	0,0%
2	50,0%
3	66,7%
4	75,0%
5	80,0%
6	83,3%
7	85,7%
8	87,5%
9	88,9%
10	90,0%
11	90,9%
12	91,7%
13	92,3%
14	92,9%
15	93,3%
16	93,8%
17	94,1%
18	94,4%
19	94,7%
20	95,0%
21	95,2%
22	95,5%
23	95,7%
24	95,8%

Tabla 4.6.4.2.1-1 – Porcentajes de aditividad por riesgo de internación.

**b - Probabilidad de internación dada la combinación de enfermedad crónica.**

Este factor se computa a partir del ratio de

$$R = \frac{Q_{Internaciones}}{Q_{Internados}} \quad (4.6.4.2.1-1)$$

por tipo de enfermedad. Matemáticamente este riesgo determina la probabilidad de sufrir una reinternación según la enfermedad padecida. Médicamente este cómputo cobra sentido en cuanto a que asigna al paciente una probabilidad computada a partir del común de la población que padece la misma enfermedad.

**c - Gravedad del suceso en base al ICD-9<sup>5</sup>.**

El riesgo atribuido por cada una de las causas de internación es indiscutiblemente subjetivo. No existe juicio válido que permita estipular matemáticamente que una causa de internación es más grave que otra. Por consiguiente, se ponderó cada una de las categorías del ICD-9 de acuerdo a la importancia asignada por las Coordinadoras de los Programas de Disease Management de OMINT S.A., ya que es una manera de priorizar aquellos casos que efectivamente pueden ser gerenciados con los recursos dispuestos en los programas (ver tabla 4.6.4.2.1-2). Es decir, ciertas internaciones, como por ejemplo las complicaciones de parto, no son causas que merezcan ser atendidas por los programas, al menos actualmente.

Aditividad por ICD-9		
Diseases and Injuries Tabular Index	Gravedad	Ponderación
1. INFECTIOUS AND PARASITIC DISEASES (001-139)	B	0%
10. DISEASES OF THE GENITOURINARY SYSTEM (580-629)	B	0%
11. COMPLICATIONS OF PREGNANCY, CHILDBIRTH, AND THE PUERPERIUM	M	50%
12. DISEASES OF THE SKIN AND SUBCUTANEOUS TISSUE (680-709)	B	0%
13. DISEASES OF THE MUSCULOSKELETAL SYSTEM AND CONNEX. TISSUE	M	50%
14. CONGENITAL ANOMALIES (740-759)	A	100%
15. CERTAIN CONDITIONS ORIGINATING IN THE PERINATAL PERIOD (760-779)	B	0%
16. SYMPTOMS, SIGNS, AND ILL-DEFINED CONDITIONS (780-799)	M	50%
17. INJURY AND POISONING (800-999)	M	50%
2. NEOPLASMS (140-239)	A	100%
3. ENDOCRINE, NUTRIT. AND METABOL. DISEASES, AND IMMUNITY DISORD.	M	50%
4. DISEASES OF THE BLOOD AND BLOOD-FORMING ORGANS (280-289)	A	100%
5. MENTAL DISORDERS (290-319)	M	50%
6. DISEASES OF THE NERVOUS SYSTEM AND SENSE ORGANS (320-389)	A	100%
7. DISEASES OF THE CIRCULATORY SYSTEM (390-459)	A	100%
8. DISEASES OF THE RESPIRATORY SYSTEM (460-519)	A	100%

<sup>5</sup> ICD-9: International Classification of Diseases

9. DISEASES OF THE DIGESTIVE SYSTEM (520-579)	M	50%
SUPPLEMENTARY CLASSIFIC. OF EXT. CAUSES OF INJURY AND POIS.	M	50%
SUPPLEMENTARY CLASSIFICATION OF FACT. INFLUENCING HEALTH STATUS	B	0%

Tabla 4.6.4.2.1-2 – Porcentajes de aditividad por riesgo de ICD-9.

#### d - Tipo de intervención (Quirúrgica, parto, quimioterapia, etc.).

Al igual que el factor anterior, este riesgo se debe ponderar en razón de la utilidad para los recursos de los Programas. La puntuación establecida en la tabla 4.6.4.2.1-3 es la siguiente:

TIPO	Puntaje
Cesárea	0%
Cirugía ambulatorias	50%
Cirugía cardiovascular	100%
Cirugía cardiovascular periférica	100%
Clínica Adultos	0%
Clínica Pediátrica	0%
Endoscopías	0%
Hemodinamia Diagnóstica	0%
Hemodinamia Terapéutica	0%
Internaciones subagudas	0%
Neonatología	0%
Partos	0%
Quimioterapia	0%
Quirúrgica Adultos	100%
Quirúrgica Pediátrica	100%
Rehabilitación	0%
Transplante	100%
Unidad de Dolor	0%

Tabla 4.6.4.2.1-3 – Porcentajes de aditividad por riesgo de tipo de intervención.

#### 4.6.4.2.2 Ponderación del riesgo por Internación

La ponderación de los cuatro factores con objeto de extraer un único valor, se estableció de la siguiente manera (ver tabla 4.6.4.2.2):

Importancia relativa de cada factor			
Puntaje x Enfermedad	Puntaje x reinternación	Puntaje por ICD-9	Puntaje por Tipo de Intervención
15%	30%	30%	25%

Tabla 4.6.4.2.2 – Ponderación relativa de cada factor aditivo de internación.

La priorización de los factores fue establecida a partir de la importancia asignada por los coordinadores de los Programas. Criteriosamente el factor de riesgo por enfermedad tiene un peso menor al ser un valor incluido en el índice de “Riesgo



por combinación de Enfermedad Crónica”. Más allá del criterio médico, esta manera facilita la repuntuación de la importancia frente a cambios en del contexto.

#### **4.6.4.3 Interpretación del riesgo de Sobrevida**

El riesgo de sobrevida es una de las variables biométricas que este proyecto incorpora. Este supone la probabilidad que un individuo fallezca dado la edad y el sexo del mismo. El objeto de su importancia radica en que el modelo debe considerar la incidencia del riesgo de muerte dado que engloba ciertos parámetros inherentes del sujeto (sexo y edad). En este sentido, se considera que para una correcta utilización de los recursos se debe priorizar a aquellos individuos de más alto riesgo, y ciertamente como la aplicación de este modelo es el gerenciamiento en salud de individuos riesgosos, esta probabilidad no puede ser obviada.

##### **4.6.4.3.1 Ponderación del riesgo de Sobrevida**

Para este estudio se utilizaron las tablas GAM-71 y MI-85 sobre mortalidad en individuos comunes e individuos discapacitados respectivamente. Estas tablas son genéricas y se aplican directamente a cada individuo de acuerdo al sexo y edad.

#### **4.6.4.4 Interpretación del riesgo por combinación de Enfermedad Crónica**

El proyecto en cuestión aplica a individuos que padecen alguna de las enfermedades crónicas listadas previamente. Ciertamente el riesgo de poseer una u otra enfermedad es de suma importancia y debe ponderarse. Sin embargo, en este punto debe advertirse que existen otros factores que enmascaran este riesgo, como lo son la cantidad de enfermedades padecidas y la combinación de las mismas. El primero de ellos advierte el creciente riesgo al poseer más enfermedades. El segundo resalta el riesgo variable frente a las distintas combinaciones de enfermedades crónicas.

La ponderación de este índice presenta a su vez ciertas consideraciones a tener presentes.

### I - Efecto combinatorio variando la cantidad de cronicidades

Por un lado debe notarse que la cantidad de padecimientos no necesariamente implica un riesgo mayor. A modo de ejemplo, un sujeto que presente cronicidades cardiovasculares y diabetes, tiene un riesgo más alto que otro que presente cronicidades de alergias, dislipidemias y gastrointestinales. Esta consideración se debe al efecto combinatorio variando la cantidad de cronicidades.

### II – Efecto de igualdad en la cantidad de cronicidades y variación en la combinación

Así mismo, debe considerarse el efecto opuesto: igualdad en la cantidad de cronicidades y variación en la combinación. Ello implica que dentro de un subgrupo fijado por la igualdad en la cantidad de enfermedades, existe una diferencia en el riesgo de sus individuos dada por la combinación. Ejemplificando dentro del subgrupo de aquellos que presentan tres padecimientos, el riesgo de un sujeto que presente cronicidades cardiovasculares, diabetes y renales es mayor que aquel que presente cronicidades infecciosas, endocrinas y ocular.

#### 4.6.4.4.1 Ponderación del riesgo por combinación de Enfermedades Crónicas

El riesgo que presenta cada enfermedad crónica se estableció a partir de la importancia que esta tiene para su gerenciamiento por medio de los programas de Disease Management. Este riesgo se fijó a partir de la siguiente puntuación:

Id	severidad
1	sin gravedad
2	baja gravedad
3	media gravedad
4	alta gravedad
5	muy alta gravedad

Tabla 4.6.4.4.1 – Severidad de combinación de Enfermedades Crónicas

A fin de evaluar el riesgo propio de la combinación entre enfermedades crónicas, se puntuó la relación de gravedad basada el riesgo potencial a la salud que la combinación presente. Ello implica por ejemplo que la combinación de enfermedad cardiovascular - dislipidemias tiene una gravedad 5, y la combinación

cardiovascular – gastrointestinal tiene una gravedad 2. Tales combinaciones se encuentran en el Anexo 4.

A partir de ello, a cada paciente se le sumo el puntaje obtenido por las distintas combinaciones que presenten sus enfermedades entre sí. La obtención de un número que indique el riesgo es el paso siguiente. Para ello se estableció que el riesgo queda determinado por dos scores complementarios. El primero de ellos denominado “Score por cantidad de Enfermedades” que computa el cociente entre la puntuación obtenida y la puntuación total que pueda llegar a obtener dada su cantidad de enfermedades. Esto es:

$$R_1 = \frac{\sum \text{puntaje}}{\text{puntaje}_{\text{maximo}}} \quad (4.6.4.4.1-1)$$

donde el puntaje máximo viene dado por

$$\text{puntaje}_{\text{maximo}} = Q_{\text{combinaciones}} * 5 \quad (4.6.4.4.1-2)$$

, donde el 5 indica la mayor gravedad y el  $Q_{\text{combinaciones}}$  la cantidad de combinaciones no permutadas entre enfermedades. Ello sin embargo desfavorece a aquellos pacientes con más cantidad de enfermedades, ya que sus chances de lograr el 100% se ven perjudicadas por la mayor cantidad de enfermedades. Es decir, es más factible que logren el 100% si se tiene dos enfermedades:  $p = 0,5$ , que si se tiene tres:  $p = 0,333$ .

Por consiguiente debe establecerse otro score adicional denominado “Score Global”. Este toma en cuenta el % relativo de las chances de llegar al 100%, castigando entonces a los de menos enfermedades, y adhiriendo un riesgo creciente a los pacientes de más enfermedades. Matemáticamente esto supone que

$$R_2 = 1 - \frac{1}{Q_{\text{enfermedades}}} + 4\% \quad (4.6.4.4.1-3)$$

donde la adición del 4% se debe al ajuste de la curva para que esta llegue al 100% frente a una cantidad de veintitrés enfermedades, número total de enfermedades crónicas relevadas en la población. Este índice a su vez enmascara el riesgo potencial de los grupos con menor cantidad de enfermedades. Luego ambos scores son unificados en base a una asignación del 40% para el primero y del 60% para el segundo.

Así mismo como se busca actuar sobre los pacientes más graves, la forma de hallar que estos impacten de manera significativa, independientemente de su cantidad de enfermedades, es que a cada uno que tenga alguna combinación de muy alta gravedad (5), se le considere el 100% de riesgo. Aquí no importa tanto que se establezca un coeficiente que priorice a los que tienen más cinco, ya que el resto de los índices se encargan de ello, en particular el “Riesgo Cluster”. Para

este índice únicamente se evalúa la gravedad de la combinación de enfermedades de cada socio, maximizando aquellos con alguna combinación crítica, y priorizando el resto.

#### **4.6.4.5 Interpretación del riesgo por exceso y defecto de consultas**

##### **A – Riesgo por exceso de consultas**

Las fuentes de información en el mercado argentino de medicina prepaga no permiten determinar el estado de salud de un sujeto. Sin embargo, sí puede inferirse un estado de riesgo a partir de la excesiva utilización de servicios médicos, y en especial por el exceso de consultas. La lógica establece que un sujeto enfermo, y en particular un sujeto con padecimientos crónicos, realiza consultas a especialistas con cierta regularidad, y con mayor frecuencia si es que presenta un deterioro en su salud <sup>6</sup>. Por consiguiente, la ponderación de este índice infiere de cierto modo el estado de salud del sujeto, y es determinado en base a la cantidad de consultas realizadas de acuerdo a la cantidad de cronicidades que padezca. Es decir, de acuerdo a la cantidad de padecimientos, se determinó el número de consultas a partir del cual la utilización será considerada excesiva, y de allí en más se estableció el riesgo creciente que representa.

##### **B – Riesgo por defecto de consultas**

Respecto al defecto de consultas, se ha considerado necesario establecer una ponderación de su riesgo, ya que tal parámetro evidencia una falta de cuidado por parte del individuo. Si bien puede objetarse que la no concurrencia al médico puede ser evidencia de un estado de salud favorable, es necesario recordar que la población de aplicación de este proyecto es puramente crónica, y por lo cual, todos presentan al menos una enfermedad, y en consecuencia, el seguimiento médico es necesario fuere cual fuere el estado de salud. En consecuencia, la ponderación de este riesgo, al igual que el anterior, se establece a partir de la cantidad de visitas ajustadas por la cantidad de cronicidades.

---

<sup>6</sup> Lamentablemente las fuentes de información utilizadas para este trabajo no permitieron validar estadísticamente esta aseveración, aunque es afirmada ampliamente en círculos médicos.

#### 4.6.4.5.1 Criterio previo a la ponderación del riesgo por consultas

Previo a ponderar los índices de consultas se debe determinar la distribución de los datos. Si se evalúa la distribución de consultas médicas se observa que dado que la cantidad de datos es sumamente alta, puede ajustarse perfectamente a una distribución continua en lugar de una discreta. La prueba de bondad de ajuste realizada corresponde a una distribución Gamma, de parámetros igual a 5.73 y 2.69 de media y desvío respectivamente. Mediante el test de Kolmogorov-Smirnov su p-value es concluyente y próximo a 0.007. Gráficamente el resulta obtenido a partir de la figura 4.6.4.5.1 es el siguiente.

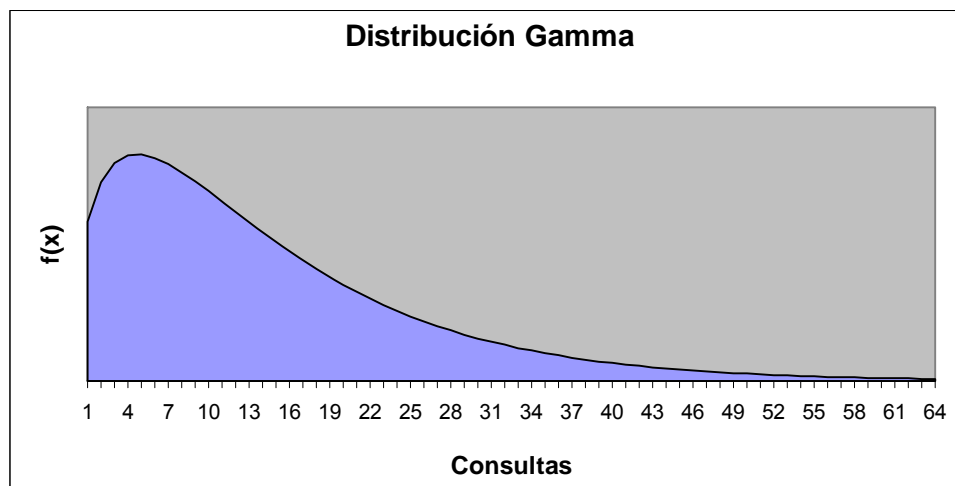


Figura 4.6.4.5.1 – Distribución de consultas médicas

#### 4.6.4.5.2 Ponderación del riesgo exceso de consultas

Esta ponderación puntúa la cantidad de consultas en exceso, ajustadas por la cantidad de enfermedades crónicas del paciente <sup>7</sup>. Esto último advierte la corrección que debe considerarse entre sujetos con distintas cantidades de enfermedades. En este sentido, como la cantidad creciente de cronicidades está correlacionada positivamente con el aumento de consultas ( $R^2=0.911$ ), debe ajustarse el mínimo de consultas deseables en base a la cantidad de enfermedades padecidas. A priori se puntualizó que el riesgo de este score debía ser monótonamente creciente, y en acuerdo a la distribución Gamma de la variable de consultas. El riesgo resultante puede concebirse a partir de la inversa

<sup>7</sup> Debe advertirse que pueden presentarse desvíos frente a sujetos hipocondríacos u aquellos que no realizan visitas por alguna razón particular, pero tales casos son sesgos propios del cálculo.

de la distribución gamma obtenida para las consultas. Gráficamente (figura 4.6.4.5.2-1) la función de densidad es la siguiente:

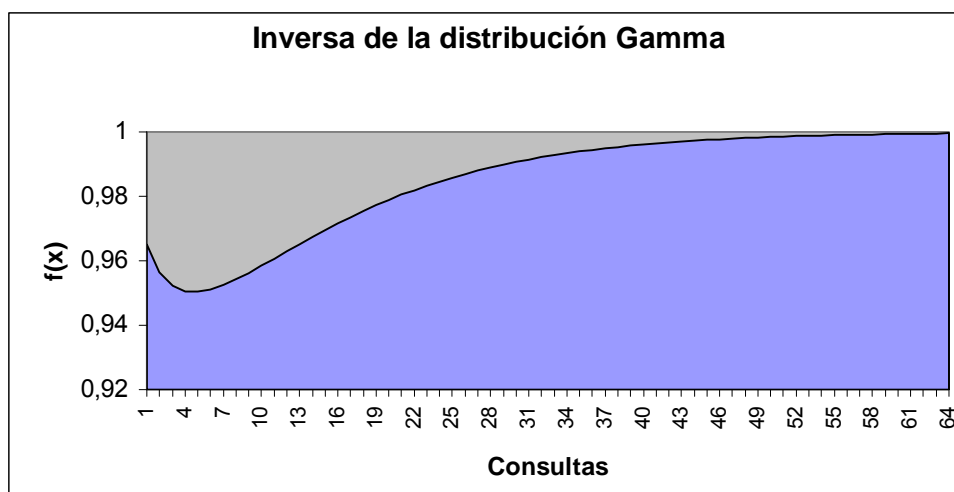


Figura 4.6.4.5.2-1 – Inversa de la distribución de consultas.

Sin embargo la pendiente debe ajustarse respecto a la cantidad de enfermedades padecidas. Es decir, el score debe puntuar más alto frente a una mayor cantidad de enfermedades. Por consiguiente, tomando la función a partir de las nueve consultas (valor a partir del cual se computará el exceso), se establece una pendiente exponencial monótonamente creciente y ajustada por los coeficientes de cantidad de consultas y enfermedades, según

$$R = \left[ 1 - \left( \frac{\lambda}{\Gamma(r)} (\lambda x_{\text{Consultas}})^{r-1} e^{-\lambda x_{\text{Consultas}}} \right) * Q_{\text{Enfermedades\_relevadas}} \right] \quad (4.6.4.5.2)$$

Estadísticamente, la fórmula establecida es la inversa de la función densidad de la Gamma, la cual es ajustada según la cantidad de enfermedades crónicas padecidas. Los parámetros  $\lambda$  y  $r$  se establecen en cada caso según la cantidad de enfermedades, es decir, representan los parámetros poblacionales de cada grupo con igual cantidad de padecimientos. Ello implica que el riesgo varía y se acrecienta frente a mayor cantidad de enfermedades. Por otro lado, a la función Gamma obtenida se la multiplica por un coeficiente de ajuste dado por la cantidad de enfermedades relevadas en el estudio. Este ajuste amplifica el riesgo creciente al aumentar la cantidad de enfermedades. Gráficamente el resultado es el siguiente (ver figura 4.6.4.5.2-2):

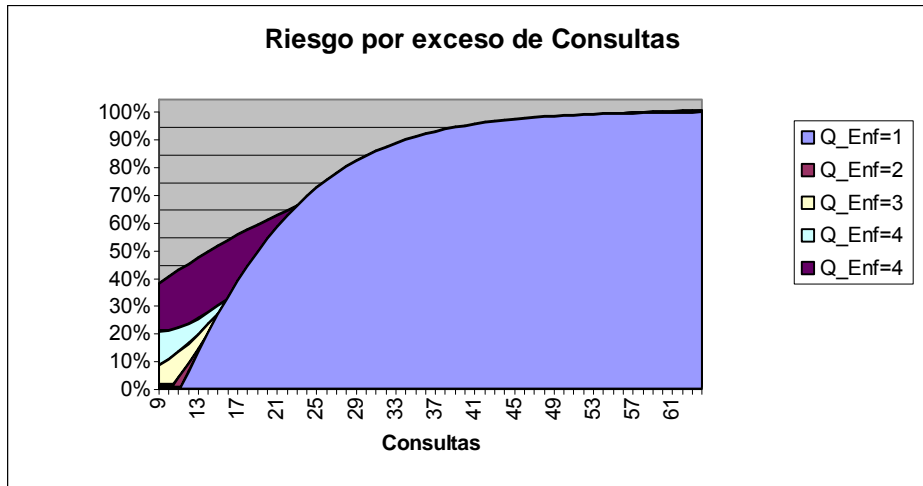


Figura 4.6.4.5.2-2 – Riesgo por exceso de consultas

### 4.6.4.5.3 Ponderación del riesgo defecto de consultas

Debido a que riesgo debe ser menor al aumentar la cantidad de consultas, y por otro lado debe ser mayor al aumentar la cantidad de padecimientos crónicos, se buscó un ajuste que, estadísticamente conforme a la distribución gamma obtenida para las consultas, presente un decaimiento moderado inicial. Esta propiedad la brinda la exponencial, la cual resulta asociada a la gamma dado que esta última es una suma de  $r$  variables exponenciales [García R., 2003]. Por consiguiente, se aplicó la siguiente fórmula dada por:

$$R = \frac{Q_{enfermedades} + e^{-Q_{consultas}}}{100} \quad (4.6.4.5.3).$$

La matriz 4.6.4.5.3 indica el resultado de ella.

Q enf	Defecto de consultas		
	0	1	2
1	2,0%	1,4%	1,1%
2	3,0%	2,4%	2,1%
3	4,0%	3,4%	3,1%
4	5,0%	4,4%	4,1%
5	6,0%	5,4%	5,1%
6	7,0%	6,4%	6,1%
7	8,0%	7,4%	7,1%
8	9,0%	8,4%	8,1%
9	10,0%	9,4%	9,1%
10	11,0%	10,4%	10,1%
11	12,0%	11,4%	11,1%
12	13,0%	12,4%	12,1%
13	14,0%	13,4%	13,1%
14	15,0%	14,4%	14,1%
15	16,0%	15,4%	15,1%

Matriz 4.6.4.5.3 – Ponderación del riesgo por defecto de consultas.

#### **4.6.4.6 Interpretación del riesgo por cuidados preventivos**

Otra de las componentes que el scoring presenta es la implicancia de los cuidados preventivos. La razón de ser de este índice radica en que desde el punto de vista del gerenciamiento de enfermos crónicos, el cuidado que el mismo sujeto pueda prestar a su salud es un aspecto sumamente relevante y que destaca la efectividad de las acciones de los programas de Disease Management. En definitiva, uno de los aspectos más importantes de estos, es la mejora en la salud del paciente a través de la educación en el autocontrol y control preventivo.

##### **4.6.4.6.1 Ponderación del riesgo por cuidados preventivos**

La ponderación en cuestión es en base a la mínima y máxima cantidad de consultas deseables. Tal índice establece que entre tres y ocho consultas durante un período anual aplican para el cuidado preventivo, y su riesgo ponderado es 0%. Estadísticamente esto se sustenta evaluando el ajuste obtenido, y siguiendo criterios médicos de prevención. Dada la media y el desvío, se establece el máximo y mínimo esperable de consultas en un  $\pm 1\sigma$  de la media, es decir, en 8 y 3 consultas respectivamente.

Puede objetarse el por qué de no ajustar tal índice por la cantidad de enfermedades crónicas que el paciente posee, siendo su razón que a fines del problema, este índice busca eximir de una componente a aquellos que presentan una utilización moderada, independientemente de su cantidad de enfermedades. Ello asume que un paciente estable, independientemente de sus padecimientos, realiza consultas periódicas de chequeo general, y algunas específicas, no debiendo ser entonces una utilización excesiva.

##### **4.6.4.7 Interpretación del riesgo del Costo relativo para OMINT.**

Otro de los aspectos relevantes que incluye este scoring, es la inclusión de la incidencia económica del set de enfermedades del paciente para la compañía de medicina prepaga. La razón de ello es que dado que económicamente, ciertas enfermedades son más caras que otras, su incidencia en la rentabilidad de la compañía es alta. En este sentido, es coincidente con el objetivo de maximización y priorización en la utilización de recursos.



#### 4.6.4.7.1 Ponderación del riesgo del Costo relativo para OMINT.

El costo relativo es calculado en base al % de incidencia de cada enfermedad en el gasto de la compañía. A continuación se detalla la tabla 4.6.4.7.1.

Enfermedad	% de Costo
Cardiovascular	16,5%
Neuronal	10,8%
Sangre – Estados Prehemorrágicos	10,4%
psiquiátricas	9,6%
Cáncer	8,6%
Endocrina - Glandular	7,6%
Discapacidad	7,5%
Diabetes	5,4%
Gastrointestinal	4,7%
Dislipidemias	3,7%
Músculo – Óseo - Esquelética	3,5%
Terapias de reemplazo hormonal	2,5%
Efermed Obstructivas vía aérea	2,3%
HIV	1,8%
Alergias	1,8%
Enfermedad ocular	1,6%
Infecciosa	1,1%
Renal	0,5%
Inmunosupresores	0,1%

Tabla 4.6.4.7.1 – Porcentaje de costo por enfermedad.

#### 4.6.4.8 Conclusión sobre los riesgos ponderados

Los riesgos ponderados pueden representarse gráficamente de la siguiente manera (ver figura 4.6.4.8):

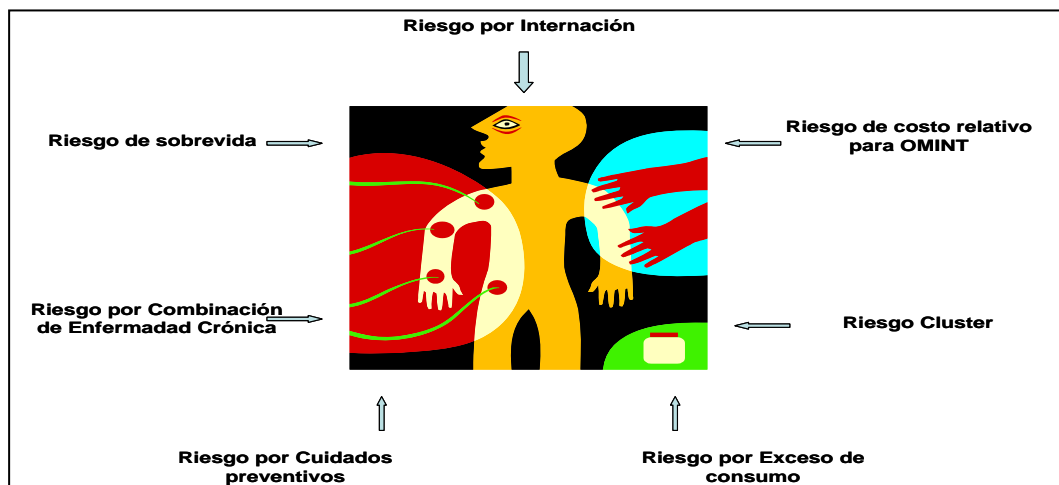


Figura 4.6.4.8 – Riesgos ponderados

Como se observa, la combinación de estos riesgos brinda un enfoque global sobre el paciente, relevando aspectos económicos, demográficos, de consumo, de salud y clínicos. El scoring final debe ser representado por un solo valor, por lo cual se pondera la importancia relativa de cada uno de estos índices. La asignación de esta se realiza respecto a la importancia considerada por los gerenciadore de los Programas, siendo los pesos estipulados,

Puntaje relativo de importancia de Índices					
Riesgo Cluster	Riesgo por Combinación E Crónica	Riesgo Consultas	Riesgo de sobrevida (%)	Riesgo de Costo OMINT	Riesgo de Reinternación
0,4	0,15	0,1	0,05	0,05	0,25

Tabla 4.6.4.8 – Puntaje relativo asignado a cada índice

De acuerdo a ello, los índices más significativos para los gerenciadore son el Riesgo Cluster, Riesgo de Reinternación y el de Combinación de Enfermedad Crónica. Ello evidencia que la información de principal interés para estos es aquella relacionada con el patrón de consumo y los estados clínicos y de salud.

## 4.7 Evaluación

### 4.7.1 Evaluación de los resultados respecto de los criterios de éxito del negocio

El scoring obtenido, y detallado en el anexo 5, engloba y combina los distintos aspectos que fueron determinados como criterios de éxito del proyecto. Tanto aspectos económicos, demográficos, clínicos, biométricos y del estado de salud, son puntualizados en particular, informándose por consiguiente el riesgo de cada uno de ellos, y no únicamente un riesgo ponderado que limite el tratamiento de la información por parte del usuario o gerenciador de los programas.

Respecto a las limitaciones que presentaba el procedimiento actual en cuanto a la calidad de la información alcanzada, estas son ampliamente superadas por el scoring desarrollado, especialmente en lo que se refiere a la diferencia de información sobre el estado de salud del paciente. Anteriormente los gerenciadore recibían información de gastos y en el caso de haber sido internados, el motivo de la misma. En base a esta vaga información debían gerenciar y priorizar los programas de Disease Management. El scoring resultante de este proyecto, no sólo utiliza esta información, sino que tal como fue descrito, adiciona otra serie de datos que complementan y profundizan el conocimiento que se tiene sobre cada paciente, permitiendo un mejor gerenciamiento.

Los resultados del scoring fueron validados por las Coordinadores de Cuidados Médicos de OMINT. Estas evaluaron el riesgo del paciente de acuerdo a su percepción sobre el estado de salud del mismo. Para ello se tomó una muestra de 140 casos, resultando el 75,6% de estos, con un riesgo acorde al percibido por las Coordinadores.

Por otro lado, los índices establecidos fueron ponderados en importancia de acuerdo a la necesidad actual de los programas. El diseño de estos permite que sus pesos relativos sean modificados, agregando una componente de flexibilidad al diseño para futuros cambios.

Por último, el scoring propuesto elimina por completo el delay que la información presentaba. La simpleza de extracción de la información de las distintas fuentes y de cálculo de sus índices, permite automatizar dinámicamente el scoring dentro de los sistemas para re-actualizar los valores diariamente.

Los resultados obtenidos ciertamente se adecuan y exceden los objetivos propuestos, considerándose por ende que el modelo generado es satisfactorio.

#### **4.7.2 Revisión del proceso de modelización**

Ciertas actividades merecen particular atención en cuando a que puedan afectar los resultados del modelo:

##### **A – Agrupación de enfermedades**

Ciertas enfermedades debieron ser reagrupadas durante el proceso de modelizado debido a que carecían de importancia en su cantidad de eventos. Tales enfermedades fueron: Sangre - Estados Prehemorrágicos; Músculo - Óseo – Esquelética.

##### **B – No apertura de las causas de rehabilitación**

Así como el evento de internación fue tabulado en sus causas y tipo de internación, los eventos de rehabilitación se encontraban dentro de los sistemas de la compañía.

##### **C – Determinación específica del índice de defecto de consultas**

Debe evaluarse su relación con el antónimo del mismo, es decir, el índice de exceso de consultas.

### **4.7.3 Pasos a seguir**

#### **4.7.3.1 Listado de posibles acciones**

##### **A – Agrupación de enfermedades**

Tal agrupación fue necesaria para el correcto modelizado de los datos obtenidos, pero no se asegura que sea así frente a futuros estudios. Es decir, dada la característica evolutiva de una población, la cantidad de padecientes de ciertas enfermedades pueden incrementarse o disminuir. Ello implica que en un futuro puede ser necesario reevaluar las agrupaciones establecidas, y ampliar las variables de consideración de los índices, en particular el riesgo cluster, ya que este hecho implicaría la adhesión o no de nuevas variables de segmentación.

##### **B – No apertura de las causas de rehabilitación**

Si bien tal información podía ser relevada, no se considera influyente su inclusión en cuanto a que no sean aspectos sobre los cuales los gerenciadore de los programas en OMINT S.A. puedan intervenir.

##### **C – Determinación específica del índice de defecto de consultas**

El índice en cuestión si bien estadísticamente es válido, debería estar asociado a la fórmula de cómputo del índice de exceso de consultas, el cual estadísticamente sí se encuentra más acorde a la distribución gamma establecida para las consultas. Sin embargo, si bien debería ser la cola izquierda de la inversa de la gamma, los resultados obtenidos con ella no son concretos en cuanto a que no reproducen la lógica del mayor riesgo frente a mayor número de enfermedades. Es por ello que debió aplicarse una salvedad, que aunque se encuentra estadísticamente justificada, es puede objetarse el por qué de tal fórmula y no otra. Así mismo, dado que el peso ponderado para este índice es bajo, al menos en lo que respecta a los pesos ponderados inicialmente, no se considera que deba modificarse el cálculo del mismo.

#### **4.7.3.2 Decisión**

A pesar de las excepciones encontradas en la revisión del modelo, se considera que estas no tienen un impacto significativo en el resultado obtenido, por lo cual se aprueba el proyecto para la siguiente etapa.

## 4.8 Entrega Final

### 4.8.1 Plan de Entrega

#### 4.8.1.1 Información a ser entregada

La entrega de información a los Coordinadores de los Programas de Disease Management consistirá en un listado que contenga datos mínimos sobre el riesgo del paciente dentro de las consultas del software de gestión (CRM) que se utiliza actualmente para el gerenciamiento de los programas (ver tabla 4.8.1.1). Este listado contendrá el Número de Socio, el Nombre del Socio, el riesgo abierto en cada una de sus categorías, y el Score Final ponderado (riesgo relativo final). El listado estará ordenado por riesgo descendente de manera que los coordinadores puedan priorizar su acción.

NroSoc	Nombre	edad	sexo	Riesgo Cluster	Riesgo x Combinación E Crónica	Riesgo Consultas	Riesgo de sobrevida	Riesgo de Costo OMINT	Riesgo de Internación	Riesgo relativo Final
XX	Anónimo	35	F	75,0%	43,4%	31,7%	0,0651%	27,3%	49,0%	53,3%
XX	Anónimo	54	M	12,5%	100,0%	43,4%	2,5073%	7,5%	45,6%	36,2%
XX	Anónimo	5	F	87,5%	100,0%	70,5%	0,9159%	18,3%	22,5%	63,6%
XX	Anónimo	60	M	55,0%	43,4%	49,4%	1,3119%	8,7%	63,5%	49,8%
XX	Anónimo	64	M	57,5%	100,0%	74,3%	1,9185%	5,4%	0,0%	45,8%
XX	Anónimo	54	F	87,5%	43,4%	57,7%	0,2982%	8,6%	56,8%	61,9%
XX	Anónimo	11	F	87,5%	14,5%	49,4%	0,9245%	16,1%	32,9%	51,2%
XX	Anónimo	64	M	100,0%	14,5%	54,7%	1,9185%	25,1%	54,5%	62,6%
XX	Anónimo	4	F	0,0%	100,0%	53,3%	0,9148%	7,5%	0,0%	20,8%
XX	Anónimo	67	M	100,0%	100,0%	66,7%	2,6316%	27,8%	0,0%	63,2%
XX	Anónimo	74	M	85,0%	100,0%	59,2%	5,1221%	29,0%	46,3%	68,2%
XX	Anónimo	0	M	5,0%	43,4%	66,9%	0,1672%	10,4%	20,5%	20,9%

Tabla 4.8.1.1 – Score final

Así mismo el sistema debe proveer la manera de que el coordinador pueda expandir la información inicial para poder evaluar los distintos datos crudos utilizados para el determinar el riesgo de cada socio, a saber, la cantidad de consultas de cierto tipo, gasto incurrido, enfermedades que presenta, medicación que consume, internaciones recientes con causa y tipo de la misma entre otras. De esta manera el coordinador podrá conocer en detalle el patrón de consumo y situación presente del paciente, permitiendo así un gerenciamiento más efectivo de acuerdo a su situación particular.

#### **4.8.1.2 Informatización del Proyecto**

En cuanto a la instalación de este proyecto dentro de OMINT S.A., es necesaria la participación del sector de Tecnología Informática (TI) para su aplicación dentro de los sistemas de gestión de la compañía. Este deberá diseñar el soporte específico que permita la correcta extracción automática de los datos desde las distintas fuentes de información estipuladas, y así también el cálculo de los índices y su posterior entrega a los usuarios finales, a saber, los gerencadores de los Programas. Por otro lado debe contemplar planes de contingencia frente a errores en el procesamiento de las actualizaciones de información.

#### **4.8.1.3 Propagación de la información a los usuarios**

El sistema debe estar configurado de manera tal que diariamente se actualicen los datos de los pacientes incluidos dentro de los programas de Disease Management, y por otro lado, recorra las fuentes de información a fin de detectar nuevos socios crónicos de acuerdo a la utilización de medicación, tal como fue descrito en el Inciso 4.4.1. Una vez completado dentro del sistema el proceso de actualización de los datos desde el Data Warehouse, el sistema debe recalculer para cada paciente los distintos índices estipulados, y ordenar descendentemente el riesgo relativo final, para luego alimentar el sistema de gestión (CRM) utilizado por los gerencadores. Luego los gerencadores observarán en el CRM el listado de pacientes ordenado por su riesgo, y sobre ellos priorizarán su labor de seguimiento.

#### **4.8.1.4 Medición de los beneficios del proyecto**

La medición de los beneficios luego de la implementación del proyecto es algo compleja. Oportunamente fue comentada la dificultad que presentan los programas de Disease Management para demostrar un valor económico que sustente su aplicación. Sin embargo, este proyecto puede ser validado de acuerdo a aspectos funcionales y a aspectos económicos. El primero de ellos surge del feedback propio por parte de los gerencadores, donde la verdadera funcionalidad y precisión del proyecto será determinada dentro del día a día. Estos serán los que en definitiva determinen si el proyecto es útil o no para su labor. Por otro lado, debe evaluarse la capacidad de inclusión de los pacientes dentro de los programas, tal como fue comentado en los incisos 3.2 y 4.3.

Respecto al aspecto económico, podrán recalcularse los ahorros obtenidos sobre la población incluida dentro de los programas con una metodología similar a la

utilizada en el “Informe sobre el impacto económico de los programas de Case Management”. Así mismo, podrá compararse el ahorro impartido con el gerenciamiento previo y posterior a este proyecto.

#### **4.8.1.5 Problemas y contingencias en la entrega de resultados**

Los problemas que puedan surgir en la entrega de resultados (listado del scoring), están ligados específicamente al funcionamiento de los sistemas informáticos, tanto para la actualización de los datos desde las distintas fuentes de información, como para el volcado correcto en los sistemas de gestión de los usuarios. De acuerdo a ello, el sistema deberá contar con alarmas de aviso que se disparen automáticamente frente a interrupciones del proceso de actualización, no debiendo bajo ningún concepto suspender el volcado de la información a los usuarios, ya que esta es necesaria para que efectúen su labor. Respecto a ello, la confección de los índices dentro del armado del proyecto fue concebida de una manera tal que permite calcular cada índice por separado, por lo cual, en caso de presentarse un corte en el procesamiento, no se inutilizará todo el proceso de actualizado, sino únicamente el cálculo de ese índice específico, por lo cual, se evita la parada total del proceso, y el sistema automáticamente debe recurrir al proceso de contingencia que supone la utilización del último valor correcto conocido del índice.

#### **4.8.2 Plan de monitoreo y mantenimiento**

##### **4.8.2.1 Aspectos dinámicos del entorno**

Debe considerarse la dinámica del entorno a fin de entender cuando el proyecto deja de tener validez y debe adaptarse.

##### **4.8.2.1.1 Nuevas fuentes de Información**

En este ámbito es necesario comprender que las fuentes de información del sistema de medicina prepaga de a poco van logrado una mejora sustancial en su calidad. Nuevas fuentes de información sumamente útiles deberán ser incorporadas, ya sea a partir de la incorporación de nuevos índices, o de la inclusión de nuevas variables para determinar los preexistentes.

#### **4.8.2.1.2 Nuevos programas de Disease Management**

Así mismo desde los niveles gerenciales de OMINT S.A. surgirán nuevas directivas para el gerenciamiento de otras enfermedades no incluidas en este estudio, por lo cual deberán readaptarse los índices.

#### **4.8.2.1.3 Nuevas enfermedades**

La constante evolución de la medicina indica el descubrimiento de nuevas enfermedades y tratamientos [Whellan *et al*, 2005], las cuales podrán ser incluidas dentro del marco del proyecto.

#### **4.8.2.1.4 Cambio en el costo de las enfermedades**

El tratamiento de enfermedades conlleva un costo asociado. Ciertamente el descubrimiento de nuevas terapias y de nuevas drogas impacta en el costo que cada enfermedad traslada a la compañía [Dickson and Jacobzone, 2003], por lo cual las prioridades de focalización pueden variar.

#### **4.8.2.1.5 Cambios en las prioridades de los programas**

Ciertos índices estipulados poseen una componente de importancia ad-hoc dada por OMINT S.A. De acuerdo a ello, las prioridades en la asignación pueden mutar con el tiempo y deberán modificarse los parámetros de cálculo.

#### **4.8.2.1.6 Apertura específica de enfermedades**

El desarrollo del modelo contempla variables genéricas para todas las enfermedades. En la medida que la calidad en las fuentes de información aumente, puede ser viable la inclusión de otras variables específicas a cada enfermedad, por lo cual debe modificarse la forma de cálculo de los índices.



#### **4.8.2.2 Medición de la precisión de los resultados**

Los resultados arrojados por el sistema deberán ser validados por el Analista de Programas Especiales por medio de un muestreo aleatorio de casos a fin de determinar la correspondencia con los datos del sistema. Así mismo el Analista deberá reunirse con los coordinadores a fin de recibir de ellos el feedback sobre la veracidad de los resultados. Cabe remarcar que los coordinadores aportan los aspectos subjetivos que infieren durante el contacto con el paciente, por lo cual a ellos corresponde determinar, dentro de su rango de subjetividad, la precisión de los resultados del scoring.

#### **4.8.2.3 Caducidad del proyecto**

El tiempo de vida del modelo estará determinado por la utilidad que este presente a sus usuarios frente a los cambios en el contexto detallados en el inciso 4.8.2.1. De todos modos, debe recalarse que la forma en la cual el modelo fue concebido es sumamente flexible. Esta permite adicionar sin menores complicaciones nuevas variables de importancia que surjan de una mejora en las fuentes de información.

Por otro lado, en cuanto a la detección de nuevas enfermedades, el índice cluster puede recalarse, determinando el riesgo de los nuevos grupos poblacionales.

Por último dado que el procedimiento de ponderación de los índices fue parametrizado de una manera simple, frente a cambios en la importancia de estos para la compañía su adaptación es sencilla.

#### **4.8.2.4 Plan de mantenimiento y chequeo**

A fin de validar el funcionamiento del modelo, periódicamente deberán reunirse los miembros del ECCM junto con el Analista de Programas Especiales a fin de validar el correcto funcionamiento del modelo, e indicar mejoras sobre la performance de los resultados. Posteriormente el Analista de Programas Especiales deberá analizar los comentarios de la reunión y buscará optimizar el modelo de acuerdo a los requerimientos.

## 5 CONCLUSIONES

Las compañías de medicina prepaga son uno de los sectores del mercado de la salud que hoy en día encuentran un futuro incierto. Tal como fue fundamentado, es indudable que el costo para estas será cada vez mayor, principalmente debido a que la detección y tratamiento de enfermedades crónicas es cada vez más eficiente, complejo y costoso, impactando directamente en la solidez financiera de las compañías.

Una de las estrategias más efectivas encaradas desde el sector a fin de paliar esta situación fue la implementación de programas de Disease Management. Estos programas consideran que el cuidado de la salud puede ser entregado más eficientemente si aquellos pacientes enfermos o crónicos, toman un rol activo en su cuidado y su evolución es guiada conjuntamente por expertos en salud, a fin de prevenir o demorar la progresión de su enfermedad. Ciertamente la prevención temprana de eventos adversos, por medio de una intervención adecuada, es un beneficio global tanto para la empresa como para el paciente, logrando una mejor sustentabilidad económica a largo plazo. Si el sistema es capaz de detectar aquellos sujetos en riesgo, el uso eficiente de recursos permitirá focalizarse en su cuidado, y de una forma más efectiva.

Como toda empresa, los recursos deben ser empleados eficazmente. En el caso particular de OMINT S.A., los recursos tecnológicos, económicos y humanos destinados a sus programas de Disease Management distaban de serlo, objeto de una ineficiente explotación de los sistemas de información. Ello generaba que la labor de los gerencadores fuese dificultosa en cuanto a la toma de decisiones, tanto para el establecimiento de prioridades, como para la detección y evaluación de la evolución del paciente durante el gerenciamiento, con el consecuente impacto en la reducción del beneficio económico que este tipo de programas presenta.

Por consiguiente, el Scoring de riesgo desarrollado imparte una nueva perspectiva para el gerenciamiento de los enfermos crónicos. La incorporación de las distintas variables clínicas, demográficas, económicas, biométricas y del estado de salud, conjuntamente con la detección de los patrones de consumo de los distintos subgrupos poblacionales, aportan un enfoque más acorde para la correcta toma de decisiones, y en consecuencia, para el mejor aprovechamiento de los recursos utilizados. Así mismo, el scoring desarrollado presenta una estructura de armado sumamente flexible. Ello implica que los cambios que puedan efectuársele para ampliar su rango de aplicación, ya sea para la consideración de nuevas variables, o bien la ponderación de otros índices, son fácilmente realizables. Por otro lado, la sistematización y automatización del scoring en los sistemas de información de la compañía, permiten incorporar la actualización continua de la información de cada

paciente, incorporando por ende una componente dinámica que permite priorizar la asignación de recursos.

## 6 FUTURAS LÍNEAS DE INVESTIGACIÓN

El scoring desarrollado surge a partir de la utilización de fuentes de información que en la actualidad OMINT S.A. considera que son de alta calidad. Ciertamente existen en los sistemas de la compañía otros datos que de lograrse una calidad superior pueden llegar a demostrar una buena aptitud que justifique su inclusión, tal como lo son los datos referentes a resultados de laboratorio. Por consiguiente, debe evaluarse en un futuro la posibilidad de incluir nuevas variables relevantes al score.

Otro proyecto que se desprende del propuesto, es el desarrollo de una estratificación del riesgo ad-hoc por tipo de enfermedad. Es decir, fácilmente puede adaptarse el scoring propuesto para enfermedades crónicas específicas, incluyendo por consiguiente ciertas variables clave para cada tipo de enfermedad medida.

Por otro lado, el proyecto propuesto supone un primer paso para la determinación del riesgo de afiliación. En este sentido, dado que el proyecto desarrollado determinó el riesgo combinado de los distintos estratos poblacionales a partir del riesgo que implica su patrón de consumo, es posible extrapolar este concepto a fin de predecir el riesgo inherente que un sujeto particular presente al momento de afiliación a la compañía, estableciendo así una prima más acorde al riesgo potencial del sujeto.

## 7 BIBLIOGRAFÍA

- Moroney, Sheila D. *Understanding Health Care Cost Drivers*, February 2003. National Institute of Health Policy.
- Kasper, Dennis L., Braunwald, Eugene, Anthony Fauci, 2004. *Harrison's Principles of Internal Medicine*. Ed. McGraw-Hill, 16th Edition
- Zighed D a, Komorowski J, 2004. *Principles of Data Mining and Knowledge Discovery*. Ed. Springer, 1st Edition.
- Bodenheimer T. and Fernandez A., July 2005. *High and Rising Health Care Costs*. Ann Intern Med, pp: 26 - 31.
- Thorpe Kenneth E., Curtis S. Florence, and Joski, Peter, 2004. *Which Medical Conditions Account For The Rise In Health Care Spending?* Health Affairs.
- Apuntes de la Cátedra de Sistemas Administrativos de Información 2005, ITBA.
- Wong, Jiahui, Gilbert, Julie and Fara-On, Maria, May 2005. *Rising Tide - Understanding Demand in Health Care*. The Change Foundation.
- Whellan, David J., Cohen, Elizabeth J., Matchar, David B, 2005. Disease Management in Healthcare Organizations: Results of In-Depth Interviews with Disease Management Decision Makers. THE AMERICAN JOURNAL OF MANAGED CARE, Vol 8 No 7, pp. 633-641.
- DataBank Group, 2005. Análisis de la competencia: medicina prepaga. Competitors.
- Whellan, David J., Cohen, Elizabeth J., Matchar, David B, 2005. *Disease Management in Healthcare Organizations: Results of In-Depth Interviews with Disease Management Decision Makers*. THE AMERICAN JOURNAL OF MANAGED CARE, Vol 8 No 7, pp. 633-641.
- Crawford Albert G., Fuhr Joseph P., Hubbs Brandon, Oct 2005. *Comparative Effectiveness of Total Population versus Disease-Specific Neural Network Models in Predicting Medical Costs*. Department of Health Policy of Jefferson Medical College, Vol. 8, No. 5, pp. 277-287.
- IBM Business Consulting Services, 2003. *HealthInsder*. Survey No.9.
- Lorig K., Ritter P., Stewart A., Sobel D., William Brown B., Bandura A., Gonzalez V., Laurent D., Holman H., Nov 2001. *Chronic Disease Self-Management Program: 2-Year Health Status and Health Care Utilization Outcomes*. Medical Care. Vol 39, pp.1217-1223.
- Kings Fund. 2005. *Predictive risk Project – Literature Review –*. Kings Fund

- Grimaldi Rey, D. 1996. *Marketing social para la promoción de la salud*. Gran Canaria: Universidad de las Palmas.
- PricewaterhouseCoopers, 2002. *The factors fueling rising healthcare costs*. Washington, American Association of Health Plans. [On-line] Available: <http://www.pwchealth.com/cgi-local/hcregister.cgi?link=pdf/fuel.pdf>
- National Advisory Council on Aging (1996). *Aging vignette #29*. Ottawa, Health Canada Division of Aging and Seniors.[On-line] Available: [http://www.phac-aspc.gc.ca/seniorsaines/pubs/vignette/pdf/vig21-33\\_e.pdf](http://www.phac-aspc.gc.ca/seniorsaines/pubs/vignette/pdf/vig21-33_e.pdf)
- Oermann Marilyn H., Templin, Thomas A., Jun 2000. *Important Attributes of Quality Health Care: Consumer Perspectives*. Journal of Nursing Scholarship. Vol 32, pp 167.
- [Internet 1]: Artículo sobre “*Machine Learning*” en Wikipedia, dirección: [http://en.wikipedia.org/wiki/Machine\\_learning](http://en.wikipedia.org/wiki/Machine_learning), página web vigente al 26/11/2006.
- [Internet 2]: Artículo sobre “*Support Vector Machine*” en Wikipedia, dirección: [http://en.wikipedia.org/wiki/Support\\_vector\\_machine](http://en.wikipedia.org/wiki/Support_vector_machine), página web vigente al 26/11/2006.
- [Internet 3]: Artículo sobre “*Bayesian Learning*” en “*Machine Learning (Theory)*”, dirección: <http://hunch.net/?p=65> , página web vigente al 26/11/2006.
- [Internet 4]: Artículo sobre “*Unsupervised learning*” en Wikipedia, dirección: [http://en.wikipedia.org/wiki/Unsupervised\\_learning](http://en.wikipedia.org/wiki/Unsupervised_learning), página web vigente al 26/11/2006.
- [Internet 5]: Artículo sobre “*Self-Organizing Maps*” en “*Neural Networks Information Homepage*”, dirección: <http://koti.mbnet.fi/~phodju/nenet/SelfOrganizingMap/General.html>, página web vigente al 26/11/2006
- Hyvärinen A. 2001. *Independent Component Analysis*. Wiley Interscience.
- Petersen LA., Pietz K. and Woodard LD. 2005. *Comparison of the predictive validity of diagnosis-based risk adjusters for clinical outcomes*. Medical Care 43[1], pp.61-67.
- Hughes AS., Averill RF. and Eisenhandler J. 2004. *Clinical Risk Groups (CRGs): a classification system for risk-adjusted capitation-based payment and health care management*. Medical Care 42[1], pp.81-90.
- Dove HG., Duncan I. and Robb A. 2003. *A prediction model for targeting low-cost, high risk members of managed care organizations*. American Journal of Managed Care. Vol 9. pp.381-389.

- Barber M. and Gregorie D. 2001. *Looking for the needle in the haystack: techniques for finding*. Ref.Type: Generic.
- Selby JV., Karter AJ., Ackerson LM., Ferrara A. and Lui, J. 2001. *Developing a prediction rule from automated clinical databases to identify high-risk patients in a large population with diabetes*. Ref.Type: Generic
- Parker JP., McCombs JS. and Graddy, EA. 2003. *Can pharmacy data improve prediction of hospital outcomes? Comparison with a diagnoses-based comorbidity measure*. Medical Care. Vol 41 [3]. pp.407-419.
- McCusker J., Cardin S., Bellavance F. and Belzile E. 2000. *Return to the emergency department among elders: patterns and predictors*. Academic Emergency Medicine. Vol 7[3]. Pp.249-259.
- Wahls TL., Barnett MJ. and Rosenthal GE. 2004. *Predicting resource utilization in a veteran Health Administration primary care population: comparison of methods based on diagnoses and medications*. Medical Care. Vol.42[2]. Pp.123-128.
- Dendukuri N., McCusker J. and Belzile E. 2004. *The identification of seniors at risk screening tool: further evidence of concurrent and predictive validity*. Journal of the American Geriatric Society. Vol.52[2]. Pp.290-296.
- Stergiou C. and Siganos D. *Neural Networks*. Dirección: [http://www.doc.ic.ac.uk/~nd/surprise\\_96/journal/vol4/cs11/report.html](http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html)
- Dickson M. and Jacobzone S. 2002. *Pharmaceutical Use and Expenditure for Cardiovascular Disease and Stroke: A Study of 12 OECD Countries*. Presentation at the 4th European Conference on Health Economics, Paris.
- CIHI (Canadian Institute for Health Information). 2003. *HL7 Canada 2002/2003 The year in review*. Pp.1-16.
- Interdepartmental Committee on Aging and Seniors Issues. 2001. *Canada's Aging Population*. Pp.1-52.
- OMINT S.A. 2007. Informe sobre el impacto económico de los programas de Case Management. Pp.1-13.
- Open Plan. 2005. *Chronic Condition List*. Medical Scheme. pp.1-2.

## 8 ANEXOS

### 8.1 - Anexo 1.1: Query de información demográfica.

Select

```
a.NroSoc,  
max(b.edad) as edad,  
max(b.sexo) as sexo,  
max(b.provincia) as Provincia,  
max(b.linea) as linea
```

Into AA\_Socios\_Cronicos\_Final\_Ampliada

```
from AA_Estratificacion_Cronicos_SALIDA_Socios_Cronicos a  
left join dw..dw_poblacion_hist b on left(b.nsoc,8)+left(b.nfa,2)=a.NroSoc  
where 1=2  
group by a.NroSoc, b.edad
```

Insert into AA\_Socios\_Cronicos\_Final\_Ampliada

Select

```
a.NroSoc,  
max(b.edad) as edad,  
max(b.sexo) as sexo,  
max(b.provincia) as Provincia,  
max(b.linea) as linea
```

from AA\_Estratificacion\_Cronicos\_SALIDA\_Socios\_Cronicos a

```
left join dw..dw_poblacion_hist b on left(b.nsoc,8)+left(b.nfa,2)=a.NroSoc  
group by a.NroSoc, b.edad
```

### 8.2 - Anexo 1.2: Query de información de consumo.

Select

```
a.Nsoc,
```



b.sexo,  
b.edad,  
b.provincia,  
b.linea,  
c.Cantidad as Q\_Consultas,  
d.Cantidad\_Prestaciones as Q\_Cirugias,  
e.Cantidad\_Prestaciones as Q\_Imagenes,  
f.Cantidad\_Prestaciones as Q\_Protesis,  
g.Cantidad\_Prestaciones as Q\_Rehabilitacion,  
h.Q\_Medic\_Cronica,  
i.Q\_Patolog\_Cronicas

```
Into AA_Estratificacion_SALIDA_FINAL_Socios_50k
from AA_Estratificacion_Listado_50k a
Left join AA_Poblacion_2005_Ampliada b on b.nrosoc=a.Nsoc
Left join AA_estratificcacion_SALIDA_Consultas_y_Visitas c on c.nrosoc=a.nsoc
Left join AA_Estratificacion_SALIDA_Cirugia d on d.nrosoc =a.nsoc
Left join AA_Estratificacion_SALIDA_Imagenes_Alta_Complejidad e on e.nrosoc
=a.nsoc
Left join AA_Estratificacion_SALIDA_Protesis_y_Ortesis f on f.nrosoc=a.nsoc
Left join AA_Estratificacion_SALIDA_Rehabilitacion g on g.nrosoc=a.nsoc
Left join AA_Estratificacion_SALIDA_Medicacion_Cronica_Agrupado h on
h.nrosoc=a.nsoc
Left join AA_Estratificacion_SALIDA_Socios_Cronicos i on i.Nrosoc=a.nsoc
Where 1=2
```

```
Insert Into AA_Estratificacion_SALIDA_FINAL_Socios_50k
```

```
Select
```

a.Nsoc,  
b.sexo,  
b.edad,  
b.provincia,  
b.linea,  
c.Cantidad as Q\_Consultas,

```

d.Cantidad_Prestaciones as Q_Cirugias,
e.Cantidad_Prestaciones as Q_Imagenes,
f.Cantidad_Prestaciones as Q_Protesis,
g.Cantidad_Prestaciones as Q_Rehabilitacion,
h.Q_Medic_Cronica,
i.Q_Patolog_Cronicas
from AA_Estratificacion_Listado_50k a
Left join AA_Poblacion_2005_Ampliada b on b.nrosoc=a.Nsoc
Left join AA_estratificcacion_SALIDA_Consultas_y_Visitas c on c.nrosoc=a.nsoc
Left join AA_Estratificacion_SALIDA_Cirugia d on d.nrosoc =a.nsoc
Left join AA_Estratificacion_SALIDA_Imagenes_Alta_Complejidad e on e.nrosoc
=a.nsoc
Left join AA_Estratificacion_SALIDA_Protesis_y_Ortesis f on f.nrosoc=a.nsoc
Left join AA_Estratificacion_SALIDA_Rehabilitacion g on g.nrosoc=a.nsoc
Left join AA_Estratificacion_SALIDA_Medicacion_Cronica_Agrupado h on
h.nrosoc=a.nsoc
Left join AA_Estratificacion_SALIDA_Socios_Cronicos i on i.Nrosoc=a.nsoc

```

### **8.3 - Anexo 1.3: Query de información de gasto.**

```

/*Grupo Socios Crónicos*/
Select NroSoc, Count(NroSoc) as Q_Patolog_Cronicas
Into AA_Estratificacion_Cronicos_SALIDA_Socios_Cronicos
from AA_Socios_Cronicos_Final
where 1=2
group by NroSoc

Insert Into AA_Estratificacion_Cronicos_SALIDA_Socios_Cronicos
Select NroSoc, Count(NroSoc) as Q_Patolog_Cronicas
from AA_Socios_Cronicos_Final
group by NroSoc
--(2127 row(s) affected)

```

/\*Creo tabla de gastos generales de Socios Crónicos 2005\*/

Select

c.NroSoc,  
a.prestacion,  
a.idnomenclador,  
a.especialidad,  
a.Cantidad,  
a.Importe,  
a.Rubro\_Jerarquía\_1,  
a.Rubro\_Jerarquía\_2

Into AA\_Estratificacion\_Cronicos\_Gastos\_Generales

From AA\_Estratificacion\_Etareo\_gastos\_generales a with (nolock)

inner join AA\_Estratificacion\_Cronicos\_SALIDA\_Socios\_Cronicos c on  
c.NroSoc=a.NroSoc

where 1=2

Insert into AA\_Estratificacion\_Cronicos\_Gastos\_Generales

Select

c.NroSoc,  
a.prestacion,  
a.idnomenclador,  
a.especialidad,  
a.Cantidad,  
a.Importe,  
a.Rubro\_Jerarquía\_1,  
a.Rubro\_Jerarquía\_2

From AA\_Estratificacion\_Etareo\_gastos\_generales a with (nolock)

inner join AA\_Estratificacion\_Cronicos\_SALIDA\_Socios\_Cronicos c on  
c.NroSoc=a.NroSoc

--(2343217 row(s) affected); 00.00.16

```
Select * from AA_Estratificacion_Cronicos_Gastos_Generales
```

```
--/*SACO CONSULTAS A ESPECIALISTAS*/
```

```
Select DISTINCT
```

```
    a.nrosoc,
```

```
    sum(a.cantidad) as Cantidad
```

```
Into AA_estratificacion_Cronicos_SALIDA_Consultas_y_Visitas
```

```
from AA_Estratificacion_Etareo_gastos_Consultas_y_Visitas a
```

```
inner join AA_Estratificacion_Cronicos_SALIDA_Socios_Cronicos b on b.NroSoc =  
a.NroSoc
```

```
where 1=2
```

```
group by a.nrosoc
```

```
Insert Into AA_estratificacion_Cronicos_SALIDA_Consultas_y_Visitas
```

```
Select DISTINCT
```

```
    a.nrosoc,
```

```
    sum(a.cantidad) as Cantidad
```

```
from AA_Estratificacion_Etareo_gastos_Consultas_y_Visitas a
```

```
inner join AA_Estratificacion_Cronicos_SALIDA_Socios_Cronicos b on b.NroSoc =  
a.NroSoc
```

```
group by a.nrosoc
```

```
order by a.nrosoc
```

```
--2005: (20532 row(s) affected)
```

```
Select * from AA_estratificacion_Cronicos_SALIDA_Consultas_y_Visitas
```

```
--/*SACO CIRUJIAS*/
```

```
Select DISTINCT
```

```
    a.NroSoc,
```

```
    sum(a.cantidad) as Cantidad_Prestaciones
```

```
Into AA_Estratificacion_Cronicos_SALIDA_Cirugia
from AA_Estratificacion_Etareo_gastos_generales a
inner join AA_Estratificacion_Cronicos_SALIDA_Socios_Cronicos b on b.NroSoc =
a.NroSoc
where 1=2
group by a.NroSoc
```

```
Insert Into AA_Estratificacion_Cronicos_SALIDA_Cirugia
Select DISTINCT
    a.NroSoc,
    sum(a.cantidad) as Cantidad_Prestaciones
from AA_Estratificacion_Etareo_gastos_generales a
inner join AA_Socios_Cronicos_Final b on b.NroSoc = a.NroSoc
where prestacion in ('07.02.1','07.02.2','07.02.3','07.02.4','07.02.5','07.02.6','07.02.7',
'07.02.8','07.02.9','07.02.10','07.02.11','07.02.12','07.03.01','07.03.02','07.05.01',
'07.06.01','07.06.02','07.06.03','07.06.04','90.07.12','07.80.14a','07.80.14g','07.80.15a','07.8
0.16a',
'07.80.17a','07.80.18','07.80.19','07.80.20a','07.80.21a','07.80.21g','07.80.22a','07.80.23','07.
80.24',
'07.80.80a','07.80.81a','11.04.01','10.04.02','11.02.01','11.02.12','11.02.03','06.01.01','06.01.
01',
'06.01.02','06.01.09')
or prestacion like ('07.01.%')
or prestacion like ('07.04.%')
group by a.NroSoc
--2005: (422 row(s) affected); 0.00.24
Select * from AA_Estratificacion_Cronicos_SALIDA_Cirugia
```

```
--/*SACO IMAGENES DE ALTA COMPLEJIDAD*/
```

```
Select Distinct
    a.NroSoc,
    sum(a.cantidad) as Cantidad_Prestaciones
```

```

Into AA_Estratificacion_Cronicos_SALIDA_Imagenes_Alta_Complejidad
From AA_Estratificacion_Etareo_gastos_imágenes a
inner join AA_Estratificacion_Cronicos_SALIDA_Socios_Cronicos b on b.NroSoc =
a.NroSoc
where a.Rubro_Jerarquía_2 in ('Radiología Intervencionista', 'Resonancias','Tomografías')
and 1=2
Group by a.nrosoc

```

```

Insert Into AA_Estratificacion_Cronicos_SALIDA_Imagenes_Alta_Complejidad
Select Distinct
      a.NroSoc,
      sum(a.cantidad) as Cantidad_Prestaciones
From AA_Estratificacion_Etareo_gastos_imágenes a
inner join AA_Estratificacion_Cronicos_SALIDA_Socios_Cronicos b on b.NroSoc =
a.NroSoc
where a.Rubro_Jerarquía_2 in ('Radiología Intervencionista', 'Resonancias','Tomografías')
Group by a.nrosoc

```

```

Select * from AA_Estratificacion_Cronicos_SALIDA_Imagenes_Alta_Complejidad
--2005: (4646 row(s) affected); 0.00.02

```

```

--/*SACO PROTESIS y ORTESIS*/
Select DISTINCT
      a.NroSoc,
      sum(a.cantidad) as Cantidad_Prestaciones
into AA_Estratificacion_Cronicos_SALIDA_Protesis_y_Ortesis
from AA_Estratificacion_Etareo_gastos_generales a
inner join AA_Socios_Cronicos_Final b on b.NroSoc = a.NroSoc
where 1=2
group by a.NroSoc

```

```
Insert into AA_Estratificacion_Cronicos_SALIDA_Protesis_y_Ortesis
Select DISTINCT
    a.NroSoc,
    sum(a.cantidad) as Cantidad_Prestaciones
from AA_Estratificacion_Etareo_gastos_generales a
inner join AA_Socios_Cronicos_Final b on b.NroSoc = a.NroSoc
where prestacion in
('80.01.01','80.01.02','80.01.03','80.01.04','80.01.05','80.01.06','80.01.07',
'80.01.08','80.01.09','80.01.10','80.01.11')
group by a.NroSoc
--2005: (220 row(s) affected); 0.00.07
Select * from AA_Estratificacion_Cronicos_SALIDA_Protesis_y_Ortesis
```

```
--/*SACO REHABILITACION*/
Select DISTINCT
    a.NroSoc,
    sum(a.cantidad) as Cantidad_Prestaciones
into AA_Estratificacion_Cronicos_SALIDA_Rehabilitacion
from AA_Estratificacion_Etareo_gastos_rehabilitacion a
inner join AA_Socios_Cronicos_Final b on b.NroSoc = a.NroSoc
where 1=2
group by a.nrosoc
```

```
Insert into AA_Estratificacion_Cronicos_SALIDA_Rehabilitacion
Select DISTINCT
    a.NroSoc,
    sum(a.cantidad) as Cantidad_Prestaciones
from AA_Estratificacion_Etareo_gastos_rehabilitacion a
inner join AA_Socios_Cronicos_Final b on b.NroSoc = a.NroSoc
group by a.nrosoc
--2005: (81 row(s) affected); 0.00.00
Select * from AA_Estratificacion_Cronicos_SALIDA_Rehabilitacion
```

```
--/*SACO MEDICACION CRONICA*/  
select  
    Nrosoc,  
    sum(Cantidad) as Q_Medic_Cronica  
into AA_Estratificacion_Cronicos_Salida_Medicamentos_Cronicos  
from AA_Estratificacion_Consumo_Medicamentos_Cronicos_x_Categoria  
where 1=2  
group by NroSoc  
  
Insert into AA_Estratificacion_Cronicos_Salida_Medicamentos_Cronicos  
select  
    Nrosoc,  
    sum(Cantidad) as Q_Medic_Cronica  
from AA_Estratificacion_Consumo_Medicamentos_Cronicos_x_Categoria  
group by NroSoc  
--(20351 row(s) affected)
```





### 8.4 – Anexo 2: Identificación de cada variable categórica.

Variable: Provincia	
Zona Geográfica	Id
Capital	0
BS As	1
Litoral	2
NOA	3
Sur	4
Centro	5

Variable: Linea	
Categoría	Id
O	0
F	1
Gerenciadas	2
Otras	3

Variable: Tipo de Internación	
TIPO	ID
Cesárea	1
Cirugía amb	2
Cirugía cardiovascular	3
Cirugía cardiovascular periférica	4
Clínica Adultos	5
Clínica Pediátrica	6
Endoscopías	7
Hemodinamia Diagnóstica	8
Hemodinamia Terapéutica	9
Internaciones subagudas	10
Neonatología	11
Partos	12

Quimioterapia	13
Quirúrgica Adultos	14
Quirúrgica Pediátrica	15
Rehabilitación	16
Transplante	17
Unidad de Dolor	18

ICD-9: Diseases and Injuries Tabular Index
1. INFECTIOUS AND PARASITIC DISEASES (001-139)
10. DISEASES OF THE GENITOURINARY SYSTEM (580-629)
11. COMPLICATIONS OF PREGNANCY, CHILDBIRTH, AND THE PUERPERIUM (630-677)
12. DISEASES OF THE SKIN AND SUBCUTANEOUS TISSUE (680-709)
13. DISEASES OF THE MUSCULOSKELETAL SYSTEM AND CONNECTIVE TISSUE (710-739)
14. CONGENITAL ANOMALIES (740-759)
15. CERTAIN CONDITIONS ORIGINATING IN THE PERINATAL PERIOD (760-779)
16. SYMPTOMS, SIGNS, AND ILL-DEFINED CONDITIONS (780-799)
17. INJURY AND POISONING (800-999)
2. NEOPLASMS (140-239)
3. ENDOCRINE, NUTRITIONAL AND METABOLIC DISEASES, AND IMMUNITY DISORDERS (240-279)
4. DISEASES OF THE BLOOD AND BLOOD-FORMING ORGANS (280-289)
5. MENTAL DISORDERS (290-319)
6. DISEASES OF THE NERVOUS SYSTEM AND SENSE ORGANS (320-389)
7. DISEASES OF THE CIRCULATORY SYSTEM (390-459)
8. DISEASES OF THE RESPIRATORY SYSTEM (460-519)
9. DISEASES OF THE DIGESTIVE SYSTEM (520-579)
SUPPLEMENTARY CLASSIFICATION OF EXTERNAL CAUSES OF INJURY AND POISONING (E800-E999)
SUPPLEMENTARY CLASSIFICATION OF FACTORS INFLUENCING HEALTH STATUS AND CONTACT WITH HEALTH SERVICES (V01-V85)

## 8.5 – Anexo 3.1: Características de cada Cluster

	Overall	Clust 1	Clust 2	Clust 3	Clust 4	Clust 5	Clust 6	Clust 7	Clust 8	Clust 9	Clust 10	Clust 11	Clust 12	Clust 13	Clust 14	Clust 15	Clust 16
Q_Cirugias	0,0	0,0	0,0	0,1	0,0	0,0	0,1	0,0	0,0	0,0	0,1	0,1	0,0	0,1	0,0	0,2	0,0
Q_Protesis	0,0	0,0	0,0	0,1	0,0	0,1	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Q_Rehab	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Q_Consultas	2,7	2,7	2,4	3,2	2,6	2,1	2,4	2,3	2,9	2,0	2,6	2,3	2,7	2,9	2,5	3,1	2,9
Q_Medic	1,4	1,4	1,0	1,8	1,5	1,0	2,0	2,0	1,1	1,0	2,0	1,8	1,5	1,6	1,6	2,0	1,4
Q_Internaciones	0,2	0,2	0,2	0,3	0,2	0,2	0,4	0,0	0,1	0,1	0,3	0,1	0,1	0,1	0,1	0,3	0,2
Cardiovascular	0,4	0,1	0,0	0,8	0,8	0,0	1,0	1,0	1,0	0,0	1,0	1,0	1,0	0,5	0,0	1,0	0,1
Inmunosupresores	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Terap reempl horml	0,0	0,0	0,1	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,1
Alergias	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,1
Dislipidemias	0,2	0,0	0,0	0,1	0,1	1,0	0,8	0,1	1,0	1,0	0,4	0,4	0,3	0,2	1,0	0,6	0,1
E. Obst via aerea	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,2	0,0	0,0	0,0	0,0	0,1	0,1
Enfermedad ocular	0,1	0,0	0,1	0,1	0,0	0,4	0,1	0,0	0,0	0,0	0,1	0,3	0,0	0,0	0,1	0,1	0,0
Diabetes	0,1	0,0	0,1	0,3	0,0	0,2	0,0	0,2	0,0	0,0	0,4	0,0	0,1	0,1	0,1	0,3	0,0
Gastrointestinal	0,1	0,1	0,1	0,6	0,0	0,0	0,0	0,0	0,1	0,0	0,1	0,3	0,0	0,1	0,1	0,2	0,1
psiquiatricas	0,3	0,1	0,1	0,4	0,4	0,0	0,0	0,0	0,0	0,0	0,4	0,0	0,0	0,3	0,2	0,9	0,5
Endocrina - Glandular	0,2	1,0	0,0	0,2	0,0	0,4	0,2	0,1	0,0	0,0	0,0	0,0	0,0	1,0	0,1	0,1	0,0
Músc - Oseo - Esq	0,1	0,1	0,2	0,2	0,1	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,1	0,1	0,1
Sangre - Estados Prehemorrágicos	0,1	0,0	0,1	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,2
Discapacidad	0,0	0,0	0,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Neuronal	0,2	0,1	0,1	0,2	0,2	0,0	0,1	0,3	0,0	0,0	0,1	0,4	0,0	0,1	0,1	0,3	0,4
HIV	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Cancer	0,0	0,0	0,1	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,1	0,0	0,0	0,1	0,0	0,0	0,0
Renal	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
Infecciosa	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
edad	3,6	3,3	3,0	4,4	4,0	3,7	4,3	4,8	4,6	3,6	4,5	4,3	4,0	4,1	4,1	4,5	3,2
sexo	0,6	0,9	0,4	0,8	0,8	0,0	0,0	0,0	1,0	0,0	0,0	0,1	0,1	0,8	0,5	0,3	1,0

### 8.6 - Anexo 3.2: Ponderación de gravedad de cada Cluster

Cluster PP	Cluster Secc	Mujeres				Hombres			
		Q consumida	Q Importe	Q Edad	Q enf	Q consumida	Q Importe	Q Edad	Q enf
1	1	High Risk	High Risk	Medium Risk	High Risk				
1	2	Medium Risk	Medium Risk	Medium Risk	Medium Risk				
1	3	Low Risk	Low Risk	Medium Risk	Medium Risk	Medium Risk	Medium Risk	Medium Risk	Medium Risk
1	4	High Risk	Medium Risk	High Risk	High Risk	Medium Risk	Low Risk	High Risk	Medium Risk
1	5	High Risk	High Risk	High Risk	Medium Risk				
1	6	Medium Risk	Medium Risk	Medium Risk	High Risk	High Risk	Medium Risk	High Risk	Medium Risk
1	7	Low Risk	Low Risk	Medium Risk	Medium Risk				
1	8	Medium Risk	Medium Risk	High Risk	High Risk				
1	9	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
2	1	High Risk	Low Risk	Low Risk	Low Risk	Medium Risk	High Risk	High Risk	Low Risk
2	2	Low Risk	High Risk	Low Risk	Low Risk				
2	3	Medium Risk	High Risk	Medium Risk	Low Risk	High Risk	Medium Risk	Medium Risk	Medium Risk
2	4	Low Risk	High Risk	Low Risk	Low Risk	Medium Risk	Low Risk	Low Risk	Low Risk
2	5	Medium Risk	Low Risk	High Risk	Medium Risk				
2	6	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk	Low Risk
2	7					High Risk	Low Risk	Low Risk	Low Risk
2	8	Low Risk	Low Risk	Medium Risk	Low Risk	Low Risk	Low Risk	Medium Risk	Medium Risk
2	9	Low Risk	Low Risk	High Risk	Medium Risk				
2	10					Low Risk	Low Risk	Medium Risk	Medium Risk
2	11	Low Risk	Low Risk	Medium Risk	Medium Risk				
2	12					Low Risk	Medium Risk	Medium Risk	Low Risk
2	13	High Risk	Medium Risk	Low Risk	Low Risk	High Risk	High Risk	Medium Risk	Low Risk
2	14	Low Risk	Low Risk	Low Risk	Low Risk	Medium Risk	Low Risk	Medium Risk	Low Risk
2	15	High Risk	Medium Risk	High Risk	High Risk				
2	16	Low Risk	Medium Risk	Medium Risk	Low Risk	Medium Risk	High Risk	High Risk	Low Risk
3	1	Low Risk	Low Risk	Medium Risk	High Risk				
3	2	High Risk	Medium Risk	High Risk	High Risk	High Risk	Medium Risk	High Risk	High Risk
3	3	Medium Risk	High Risk	Medium Risk	High Risk	Medium Risk	High Risk	Medium Risk	High Risk
3	4					Medium Risk	High Risk	Medium Risk	High Risk
3	5					Medium Risk	Medium Risk	Medium Risk	Medium Risk
3	6	High Risk	High Risk	Low Risk	Medium Risk				



3	7	High Risk	Low Risk	High Risk	High Risk				
3	8	High Risk	Low Risk	Medium Risk	Medium Risk				
3	9	Medium Risk	Low Risk	Medium Risk	Medium Risk				
4	1	High Risk	High Risk	High Risk	High Risk				
4	2	Medium Risk	Medium Risk	High Risk	High Risk				
4	3	High Risk	High Risk	Medium Risk	Medium Risk				
4	4	Medium Risk	High Risk	High Risk	Medium Risk				
4	5	Medium Risk	Low Risk	High Risk	High Risk				
4	6	Low Risk	Low Risk	Medium Risk	Low Risk				
4	7	Medium Risk	High Risk	Medium Risk	Low Risk				
4	8	Medium Risk	Low Risk	Medium Risk	High Risk				
4	9					Medium Risk	Medium Risk	Medium Risk	Medium Risk
5	1					Low Risk	Low Risk	Medium Risk	Medium Risk
6	1					Medium Risk	High Risk	High Risk	High Risk
6	2					High Risk	High Risk	High Risk	High Risk
6	3					Low Risk	High Risk	High Risk	Medium Risk
6	4					Medium Risk	High Risk	High Risk	High Risk
6	5					Medium Risk	Medium Risk	Medium Risk	High Risk
6	6					Medium Risk	High Risk	High Risk	High Risk
6	7					Low Risk	Low Risk	Medium Risk	Medium Risk
6	8					Low Risk	Medium Risk	High Risk	High Risk
6	9					High Risk	Low Risk	High Risk	High Risk
7	1					Low Risk	Low Risk	Medium Risk	High Risk
7	2					Medium Risk	Low Risk	High Risk	Low Risk
7	3					High Risk	High Risk	Medium Risk	High Risk
8	1	Low Risk	Low Risk	High Risk	Medium Risk				
9	1					Low Risk	High Risk	Medium Risk	Low Risk
9	2					Low Risk	High Risk	Medium Risk	Low Risk
9	3					Low Risk	Low Risk	Medium Risk	Low Risk
10	1					Low Risk	Low Risk	Medium Risk	Medium Risk
10	2					Low Risk	Medium Risk	High Risk	Medium Risk
10	3					High Risk	High Risk	High Risk	High Risk
10	4					Low Risk	Medium Risk	Medium Risk	Medium Risk
11	1	Medium Risk	Medium Risk	High Risk	High Risk	Medium Risk	Medium Risk	High Risk	High Risk
11	2	Low Risk	High Risk	High Risk	High Risk	High Risk	High Risk	High Risk	High Risk
11	3	Low Risk	Low Risk	Medium Risk	High Risk	Low Risk	Low Risk	Medium Risk	Medium Risk
11	4	Low Risk	Low Risk	High Risk	Medium Risk	High Risk	High Risk	High Risk	High Risk

11	5	Low Risk	Low Risk	High Risk	High Risk	Medium Risk	Medium Risk	High Risk	High Risk
11	6	Low Risk	High Risk	High Risk	High Risk	Medium Risk	High Risk	Medium Risk	High Risk
11	7					Medium Risk	Medium Risk	High Risk	High Risk
11	8	Low Risk	Medium Risk	High Risk	High Risk	Medium Risk	Medium Risk	High Risk	High Risk
11	9	High Risk	Medium Risk	High Risk	High Risk	High Risk	Medium Risk	High Risk	High Risk
12	1	Medium Risk	Medium Risk	High Risk	Medium Risk				
12	2	Medium Risk	High Risk	High Risk	Medium Risk				
12	3	Low Risk	Low Risk	High Risk	Medium Risk				
12	4	Medium Risk	High Risk	High Risk	Medium Risk				
12	5	Medium Risk	Low Risk	Medium Risk	Medium Risk				
12	6	Low Risk	Medium Risk	Medium Risk	Medium Risk				
12	7					Medium Risk	High Risk	Medium Risk	Medium Risk
13	1	Medium Risk	Medium Risk	High Risk	High Risk				
13	2	Low Risk	Medium Risk	Medium Risk	High Risk				
13	3	Low Risk	Low Risk	Low Risk	Medium Risk	High Risk	Medium Risk	Medium Risk	Medium Risk
13	4	High Risk	High Risk	Medium Risk	High Risk	Medium Risk	High Risk	Medium Risk	Medium Risk
13	5	Low Risk	Low Risk	Medium Risk	Medium Risk				
13	6	Low Risk	Low Risk	Medium Risk	Medium Risk	Low Risk	Medium Risk	Medium Risk	Low Risk
13	7	Medium Risk	Medium Risk	High Risk	High Risk				
13	8	Low Risk	Low Risk	Low Risk	Medium Risk				
13	9	Low Risk	High Risk	Medium Risk	High Risk	Low Risk	Low Risk	Low Risk	Medium Risk
14	1					Low Risk	Medium Risk	Medium Risk	Medium Risk
14	2					Medium Risk	Medium Risk	Medium Risk	Low Risk
14	3	High Risk	Medium Risk	Medium Risk	Medium Risk	High Risk	Low Risk	Medium Risk	Medium Risk
14	4					Medium Risk	High Risk	Medium Risk	Medium Risk
14	5					Low Risk	Medium Risk	Medium Risk	Medium Risk
14	6					High Risk	High Risk	High Risk	High Risk
14	7					Low Risk	Medium Risk	Medium Risk	Medium Risk
14	8	Medium Risk	High Risk	High Risk	Low Risk	Medium Risk	High Risk	Medium Risk	Medium Risk
14	9	Medium Risk	Medium Risk	Medium Risk	Low Risk	Low Risk	Low Risk	Low Risk	Medium Risk
15	1					Low Risk	High Risk	High Risk	Medium Risk
15	2					Low Risk	Low Risk	Medium Risk	Medium Risk
15	3	Medium Risk	High Risk	High Risk	High Risk	Medium Risk	High Risk	High Risk	High Risk
15	4					Medium Risk	Medium Risk	High Risk	High Risk
15	5					Low Risk	Low Risk	Medium Risk	Medium Risk
15	6					Medium Risk	High Risk	High Risk	High Risk
15	7					Low Risk	Medium Risk	High Risk	High Risk



15	8					Low Risk	Medium Risk	Medium Risk	High Risk
15	9	Medium Risk	High Risk	High Risk	High Risk	High Risk	High Risk	High Risk	High Risk
15	10	Medium Risk	High Risk	High Risk	High Risk	Medium Risk	High Risk	High Risk	High Risk
15	11	High Risk	Medium Risk	High Risk	High Risk	High Risk	High Risk	High Risk	High Risk
15	12	Low Risk	Low Risk	Medium Risk	High Risk	Low Risk	Medium Risk	Medium Risk	High Risk
15	13					Medium Risk	High Risk	Medium Risk	High Risk
15	14					Medium Risk	High Risk	High Risk	High Risk
15	15	Medium Risk	Medium Risk	High Risk	High Risk	Medium Risk	High Risk	High Risk	High Risk
15	16					Medium Risk	High Risk	High Risk	High Risk
16	1	Low Risk	Low Risk	Medium Risk	Medium Risk	Medium Risk	High Risk	Low Risk	Medium Risk
16	2	Medium Risk	High Risk	Medium Risk	High Risk				
16	3	Low Risk	High Risk	Medium Risk	Medium Risk				
16	4	Low Risk	Medium Risk	Low Risk	Low Risk				
16	5	Low Risk	Medium Risk	Medium Risk	Medium Risk				
16	6	Low Risk	Low Risk	Medium Risk	Medium Risk				
16	7	Low Risk	Medium Risk	High Risk	High Risk				
16	8	Low Risk	Low Risk	Medium Risk	Low Risk				
16	9	Medium Risk	High Risk	High Risk	High Risk				
16	10	Medium Risk	High Risk	High Risk	High Risk				
16	11	High Risk	High Risk	High Risk	High Risk				
16	12	Medium Risk	High Risk	Medium Risk	High Risk				
16	13	Low Risk	High Risk	Medium Risk	High Risk	High Risk	High Risk	Low Risk	High Risk
16	14	Medium Risk	High Risk	Medium Risk	Medium Risk				
16	15	Medium Risk	High Risk	High Risk	High Risk				
16	16	Medium Risk	High Risk	High Risk	High Risk				



**8.7 – Anexo 4: Puntuación de combinaciones de Enfermedades crónicas.**

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	
A	3																							
B	1	3																						
C	1	1	5																					
D	2	3	5	5																				
E	1	1	3	3	5																			
F	1	1	5	5	4	5																		
G	4	3	3	3	5	2	4																	
H	1	3	4	5	2	3	2	4																
I	2	2	4	5	2	4	1	3	2															
J	1	3	4	3	2	2	3	3	2	2														
K	1	3	2	2	2	2	1	3	1	5	2													
L	2	4	3	5	2	3	2	5	3	4	3	3												
M	1	4	3	5	3	2	5	3	5	3	5	4	2											
N	1	3	4	5	3	4	5	2	2	3	3	3	5	3										
O	1	3	2	5	3	2	5	2	2	3	3	4	5	5	4									
P	1	2	3	2	5	1	2	2	1	2	3	2	2	2	2	4								
Q	1	2	3	2	5	3	2	2	5	3	3	3	5	4	2	4	5							
R	1	2	2	2	5	1	1	2	1	2	3	2	2	4	3	5	4	4						
S	1	1	3	3	2	2	3	3	1	4	3	2	2	3	4	4	5	4	3					
T	1	3	4	5	2	3	1	4	2	4	3	3	4	5	4	4	3	3	3	5				
U	1	3	4	2	2	4	2	4	2	5	3	3	5	4	5	3	3	2	2	4	3			
V	1	4	2	3	1	4	2	4	3	3	2	4	2	2	3	3	2	2	3	3	4	2		
W	1	1	2	3	1	2	3	2	2	5	4	2	2	2	3	5	2	5	4	4	4	3	2	

**Id.**

Enfermedad	Id
Alergias	A
Cancer	B
Cardiovascular	C
Diabetes	D
Discapacidad	E
Dislipidemias	F
Efermed. Obstructivas via aerea	G
Endócrina	H
Enfermedad ocular	I
Estados prohemorragicos	J
Gastrointestinal	K
Glandular	L
HIV	M
Infecciosa	N
Inmunosupresores	O
Musculo-Esquelética	P
Neuronal	Q
Oseo-Articular	R
psiquiatricas	S
Renal	T
Sangre	U
Terapias de reemplazo hormonal	V
Tratamiento dolor crónico	W

### 8.8 – Anexo 5: Scoring resultante.

NroSoc	Nombre	edad	sexo	Riesgo Cluster	Riesgo x Combiación E Crónica	Riesgo Consultas	Riesgo de sobrevida	Riesgo de Costo OMINT	Riesgo de Internación	Riesgo relativo Final
1037230801	Anónimo	35	F	75,0%	43,4%	31,7%	0,0651%	27,3%	49,0%	53,3%
1046339600	Anónimo	54	M	12,5%	100,0%	43,4%	2,5073%	7,5%	45,6%	36,2%
1133868804	Anónimo	5	F	87,5%	100,0%	70,5%	0,9159%	18,3%	22,5%	63,6%
1158393700	Anónimo	60	M	55,0%	43,4%	49,4%	1,3119%	8,7%	63,5%	49,8%
1053437800	Anónimo	64	M	57,5%	100,0%	74,3%	1,9185%	5,4%	0,0%	45,8%
1071173700	Anónimo	54	F	87,5%	43,4%	57,7%	0,2982%	8,6%	56,8%	61,9%
1144613503	Anónimo	11	F	87,5%	14,5%	49,4%	0,9245%	16,1%	32,9%	51,2%
1094998000	Anónimo	64	M	100,0%	14,5%	54,7%	1,9185%	25,1%	54,5%	62,6%
1099624702	Anónimo	4	F	0,0%	100,0%	53,3%	0,9148%	7,5%	0,0%	20,8%
1050129400	Anónimo	67	M	100,0%	100,0%	66,7%	2,6316%	27,8%	0,0%	63,2%
1011543400	Anónimo	74	M	85,0%	100,0%	59,2%	5,1221%	29,0%	46,3%	68,2%
1021689301	Anónimo	0	M	5,0%	43,4%	66,9%	0,1672%	10,4%	20,5%	20,9%