



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

FUSIONES Y ADQUISICIONES DE EMPRESAS FARMACÉUTICAS Y LA RED DE
GENES-ENFERMEDADES

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA

MEMORIA PARA OPTAR AL GRADO DE INGENIERO CIVIL INDUSTRIAL

TOMÁS WILNER

PROFESOR GUÍA:
CARLOS NOTON NORAMBUENA

MIEMBROS DE LA COMISIÓN:
MANUEL HERMOSILLA COFFRÉ
MARCELO OLIVARES ACUÑA

Este trabajo ha sido parcialmente financiado por CONICYT

SANTIAGO DE CHILE
2017

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE MAGÍSTER EN ECONOMÍA APLICADA
Y AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL
POR: TOMÁS WILNER
FECHA: 2017
PROF. GUÍA: DR. CARLOS NOTON NORAMBUENA

FUSIONES Y ADQUISICIONES DE EMPRESAS FARMACÉUTICAS Y LA RED DE GENES-ENFERMEDADES

Un gran esfuerzo en I+D se ha desplegado durante los últimos diez años para el desarrollo de terapias genéticas, impulsado por la esperanza de combatir las enfermedades en su expresión más básica. Simultáneamente se observan diversas fusiones y adquisiciones en este mercado. ¿Cuál es la principal causa de las fusiones? ¿Son razones de eficiencia en la investigación justificadas por la red genética?

Para responder esta pregunta se elabora un modelo probit cuya variable dependiente es la dummy de fusión o adquisición entre firmas, al cual se le aplican gran cantidad de variantes y diversos tests de robustez. Se evalúa la importancia del target genético junto con otras variables como target de enfermedades y target de áreas terapéuticas. A su vez se adiciona una variable que intenta dilucidar la existencia de una relación continua entre la dummy y la presencia de target genéticos en ambas firmas, proveniente de la teoría de grafos.

El enfoque de este trabajo es novedoso por dos razones: hasta el momento nadie ha investigado si las fusiones o adquisiciones entre empresas farmacéuticas pueden explicarse a partir de las terapias genéticas que estas realizan (producto del reciente florecimiento de este tipo de terapias) y tampoco se ha explorado la estructura de red de los genes para explicar el comportamiento de las firmas (en términos de adquisición o fusión).

Esta investigación utiliza dos bases de datos. De Thomson Reuters Cortellis Life Sciences se extraen 1.417 proyectos con target genético, correspondientes a 514 firmas distintas, que ocurren desde el año 1989 hasta el año 2014. Mientras que de Deloitte RECAP IQ se obtienen 3.763 fusiones y adquisiciones de empresas farmacéuticas, desde 1988 hasta 2011.

Los datos sugieren que en presencia de dos firmas con target genético compartido, su probabilidad de fusión o adquisición es aproximadamente sesenta y cinco veces más alta, que cuando no hay genes en común. Las otras variables no son siempre significativas para el modelo, y no resisten los tests de robustez. Con respecto a la variable continua proveniente de la teoría de grafos, "Distancia de Wallis", se encuentra que es colineal a la dummy de target genético y pierde significancia cuando se coloca en el modelo junto a ella. A su vez se demuestra que la probabilidad de que una firma pequeña sea adquirida por una grande aumenta considerablemente si ambas firmas están trabajando en un gen en común y la firma pequeña ha desarrollado un fármaco relacionado a dicho gen que haya superado una etapa inicial ("Discovery").

Este trabajo de tesis deja abiertas preguntas de investigación, como por ejemplo evaluar si las externalidades en I+D presentadas por las fusiones o adquisiciones en esta nueva era se condicen con las que se presentaban antes de que la terapia génica tomase relevancia.

A mi hermano y a mis padres.

Agradecimientos

Quiero agradecer a mi familia, por el apoyo incondicional desde el principio hasta el final de mi carrera.

A mis amigos, Felipe Ugalde, Louis De Vos, José María Samaniego, Matías Cerda, Vicente Yañez, Eduardo Alcaina, Felipe Cerón, y a muchos más, por haberme acompañado y haber hecho este trayecto de siete años mucho más ameno.

A Carlos Noton, por darme la oportunidad de trabajar con él sin “peros” de por medio.

A Manuel Hermosilla y a Jorge Lemus, por aceptarme en este proyecto sin conocerme y darme la confianza para trabajar, siempre con buena voluntad.

A Marcelo Olivares, por introducirme en el área de investigación y haber sido el primero en enseñarme y motivarme para seguir mejorando en esta carrera.

Tabla de Contenido

Introducción	1
1. Datos	4
2. Modelos	8
2.1. Modelo Probit agregado	8
2.2. Modelo Probit con variación temporal	9
2.3. Test de Robustez: Probit Dinámico	10
2.3.1. Análisis por tamaño de firmas	10
2.3.2. Análisis dinámico	11
2.4. Modelo Probit con variación temporal, y métrica de distancia de Wallis	11
2.5. Test de Robustez e interacción con la evolución del proyecto	17
2.5.1. Análisis por tamaño de firmas: Probabilidad de adquisición	17
2.5.2. Interacción con la evolución del proyecto	17
3. Resultados	18
3.1. Modelo Probit agregado	18
3.2. Modelo Probit con variación temporal	20
3.3. Test de Robustez: Probit Dinámico	23
3.3.1. Análisis por tamaño de firmas	24
3.3.2. Análisis dinámico	24
3.4. Modelo Probit con variación temporal, y métrica de Distancia de Wallis	26
3.5. Test de Robustez e interacción con la evolución del proyecto	28
3.5.1. Análisis por tamaño de firmas: Probabilidad de adquisición	28
3.5.2. Interacción con la evolución del proyecto	28
Conclusión	32
Bibliografía	34
Anexo A	36

Índice de Tablas

1.1. Resumen de la Base de Datos	5
1.2. Estadísticas descriptivas de sumgene, meangene y mediangene	5
1.3. Estadísticas descriptivas de la Distancia de Wallis	6
2.1. Cantidad de fusiones y adquisiciones, y cantidad de target genéticos compartidos con una o más firmas, para cada una de las firmas grandes	10
3.1. Resultados Modelo Probit Agregado	18
3.2. Resultados Efectos Marginales Probit Agregado	19
3.3. Resultados Modelo Probit con Variación Temporal	21
3.4. Efectos Marginales Modelo Probit con Variación Temporal	22
3.5. Resultados Variables Genéticas Modelo Probit con Variación Temporal	23
3.6. Efectos Marginales Promedio Variables Genéticas Modelo Probit con Variación Temporal	23
3.7. Resultados Modelo Probit con combinaciones de firmas grandes y pequeñas	24
3.8. Resultados Modelo Probit Variables Rezagadas 1 año	25
3.9. Resultados Modelo Probit Variables Rezagadas 3 años	25
3.10. Resultados Modelo Probit Variación Temporal con Distancia de Wallis	26
3.11. Efecto Marginal Promedio de la Distancia de Wallis	27
3.12. Efectos Marginales Variables dicotómicas en modelo con Distancia de Wallis	27
3.13. Resultados Modelo Probit Distancia de Wallis con combinaciones de firmas grandes y pequeñas	28
3.14. Resultados Modelo Probit incluyendo efecto fijo por etapa de proyecto	29
3.15. Resultados Modelo Probit interacción evolución del proyecto y TargetGen	30
3.16. Resultados Modelo Probit interacción evolución del proyecto y distancia de Wallis	31
A.1. Corrección de nombres de las firmas para realizar el matching de las Bases de Datos	36

Introducción

Las fusiones y adquisiciones de las empresas farmacéuticas generan preocupación por el probable incremento de poder de mercado que implica un aumento de precios pagados por los consumidores para tratamientos y remedios. En Estados Unidos, el Departamento de Justicia y la Comisión de Comercio Federal inspeccionan los acuerdos de fusiones, e intentan bloquear aquellos que aminoren la competencia y no generen sinergias. Un ejemplo reciente de esto es el bloqueo del Departamento de Justicia a las fusiones de cuatro de las cinco aseguradoras del área de la salud más importantes de Estados Unidos: Aetna con Humana y Anthem con Cigna, las cuales provocarían que “la competencia entre aseguradoras que los obliga a bajar las primas, cuidados de más alta calidad y mejores beneficios se vea eliminada”, de acuerdo a la Fiscal General de Estados Unidos, Loretta E. Lynch¹. En la misma línea, la reducción en la competencia producto de la fusión de empresas farmacéuticas puede provocar precios más altos para pacientes y aseguradoras, disminuyendo el consumo y el bienestar del consumidor en general.

Por otra parte, las eficiencias en la producción o reducción de costos que aparecen gracias a las fusiones ayudan a contrarrestar esta disminución en la competencia. Economistas han descubierto evidencia que apoya un contraargumento: externalidades sustanciales en I+D podrían incrementar la productividad innovativa post-fusión (Devos, Kadapakkam and Krishnamurthy, 2009) posiblemente superando, por ejemplo con tratamientos más abundantes y efectivos, los efectos perjudiciales de un aumento de poder de mercado (Crandall and Winston, 2003; Williamson, 1968). Generalmente, las externalidades se pueden presentar de dos formas: externalidades de consumo (denominadas también de mercado) y externalidades de producción (denominadas también de tecnología). El tamaño y el signo de estas externalidades permiten inferir cuán deseable es invertir en investigación y desarrollo (I+D) desde un punto de vista social y privado. Se ha demostrado que las externalidades tecnológicas (conocimiento) tienen un efecto positivo mayor, que el negativo provocado por las externalidades de mercado (Bloom, Schankerman and Van Reenen, 2013). A su vez, se ha descubierto que las economías de alcance son más importantes que las economías de escala para el desarrollo de nuevos remedios (Cockburn and Henderson, 2001). Es más, en la industria farmacéutica las economías de alcance son importantes para determinar la frontera de producción de una empresa, y el incremento en la productividad innovativa de las firmas grandes puede justificar que valga la pena tolerar la pérdida de eficiencia producto del aumento de su poder de mercado (Henderson and Cockburn, 1996).

¹Noticia del New York Times, el día 22 de Julio de 2016, página B3 bajo el título “U.S. Sues to Stop Mergers of Giant Health Insurers”

Todo esto revela los efectos post-fusión, ¿pero que hay de los incentivos previos a ella? De acuerdo con Danzon, Epstein and Nicholson (2007) la razón por la cual las firmas grandes realizan fusiones es por exceso de capacidad provocada por la expiración de patentes o brechas en los ciclos de producción de diversos productos. Las firmas pequeñas, sin embargo, las realizan como una estrategia de salida en respuesta a problemas financieros. Complementando esto para la industria farmacéutica, Meder (2016) encuentra que las firmas son más propensas a fusionarse a medida que la cercanía entre sus portafolios de productos es mayor.

Existe ya literatura que se ha enfocado en demostrar la propensión a fusiones y adquisiciones basados en las patentes que poseen las firmas (Marco and Rausser, 2008; Danzon, Epstein and Nicholson, 2007). Sin embargo, impulsado por la esperanza de combatir las enfermedades en su expresión más básica, un gran esfuerzo en I+D se ha desplegado durante los últimos diez años para el desarrollo de terapias genéticas (i.e., remedios cuyo blanco son los genes relacionados a enfermedades). Aunque solo algunas de estas han alcanzado el mercado hasta ahora, gran cantidad de productos en desarrollo sugieren que en el mediano plazo muchas más lo harán (Potts, 2008, pp.180-182).

Enfocándose en la innovación de este tipo de terapias, el objetivo de esta tesis es contribuir al debate desarrollando un análisis novedoso de fusiones y adquisiciones de empresas farmacéuticas. La data recolectada para esta investigación proviene de Thomson Reuters Cortellis Life Sciences, e involucra 1.417 proyectos de terapia genética correspondientes a 514 farmacéuticas. Esto representa una parte del mercado de fusiones y adquisiciones dentro del conjunto de firmas que participan en este esfuerzo por desarrollar este tipo de terapias, lo que obliga a plantear la pregunta: ¿Permanecerán ciertas en esta era genética las implicancias de la fusión de empresas farmacéuticas identificadas por la literatura previa?

Antes de responder esta interrogante, es necesario preguntarse si las firmas, al momento de tomar la decisión de fusionarse o adquirir, evalúan, estudian o analizan los genes o enfermedades con los que la contraparte trabaja. La respuesta no es obvia; de hecho, de las empresas involucradas en este estudio que realizaron una adquisición o fusión, un 36 % lo hizo por terapias genéticas, un 32 % lo hizo por enfermedades, y el resto por distintas razones, como por ejemplo aumentar su poder de mercado o mejorar su coordinación en el desarrollo de productos². Sin embargo, a pesar de que el motivo subyacente pueda ser terapia genética, enfermedad u otro, generalmente una firma adquieren a otra en base a sus remedios o futuros remedios aprobados por la FDA³. Por tanto, en este trabajo se optó por elegir la motivación subyacente con mayor cantidad de adeptos en este subconjunto de firmas, la terapia genética y, como ya se mencionó anteriormente, el estudio se enfocará en este tipo de terapias. El estudio alternativo se deja planteado para futuras investigaciones.

Por tanto, para responder la pregunta de investigación planteada, se realiza un trabajo empírico para explicar fusiones y adquisiciones en función de la investigación genética de las firmas. En primer lugar se evalúa una variable dicotómica, la cual es 1 en caso de que dos firmas tengan terapias con algún gen en común. A su vez, se calcula una medida de

²Estos valores se obtuvieron al investigar las razones detrás de las 44 fusiones y adquisiciones de empresas farmacéuticas que se llevaron a cabo entre 1988 y 2011. La fuente de esta información son principalmente diarios como el New York Times, The Pharma Letter y Forbes.

³Food and Drug Administration: Agencia de Alimentos y Medicamentos o Agencia de Drogas y Alimentos.

cercanía entre genes, la cual se caracteriza por considerar a cuántas enfermedades en común cada par de genes está asociado. Gracias a esta medida se generan tres nuevas variables, que consideran para cada firma el promedio, la suma y la mediana de la cercanía entre sus genes. Posteriormente, todo esto se contrasta con una medida de distancia planteada por Wallis et al. (2001). Esta última proviene de la teoría de grafos, y permite observar si la probabilidad de fusión entre dos firmas tiene una variación continua respecto a la proximidad que tengan éstas en el grafo. Los nodos en este grafo son los genes, y los arcos son las enfermedades con las que éstos se relacionan. Por tanto, en caso de tener al menos una enfermedad en común, estos genes estarán conectados.

El enfoque de este trabajo es novedoso ya que hasta el momento nadie ha investigado si las fusiones o adquisiciones entre empresas farmacéuticas pueden explicarse a partir de las terapias genéticas que estas realizan y a su relación a la estructura de la red de genes. Este hecho se debe, como se señaló anteriormente, al reciente florecimiento de este tipo de terapias. A esto hay que agregarle que nadie a explorado la estructura de red de los genes para explicar el comportamiento de las firmas (en términos de adquisición o fusión).

El informe de la presente tesis se organiza como sigue: el capítulo 1 presenta y describe la data, el capítulo 2 detalla los modelos empleados, los test de robustez realizados y las variables independientes desarrolladas, el capítulo 3 analiza los resultados y el capítulo 4 concluye.

Capítulo 1

Datos

Para responder la pregunta de investigación de esta tesis se utilizan dos bases de datos: una que proporciona información sobre la relación de genes y enfermedades y otra que da cuenta de fusiones y adquisiciones.

La primera se obtiene de Thomson Reuters Cortellis Life Sciences. De ella se extraen 1.417 proyectos con target genético, correspondientes a 514 firmas distintas, que ocurren desde el año 1989 hasta el año 2014. Existen siete que no presentan fecha, por lo cual finalmente se trabajará con 1.410 proyectos. Cada uno de ellos tiene el target genético (`targetid`), el código del remedio (`drugid`), el código de la firma (`firmid`), el nombre de la firma (`companyname`), el código de la enfermedad que atacan (`indicationid`), el nombre del área terapéutica que atacan (`therapeuticarea`) y el año en el que alcanzaron su etapa más avanzada (`year`).

Los remedios en su etapa de desarrollo deben pasar por cinco etapas: Discovery, Phase I, Phase II, Phase III y Market. No todos llegan al mercado, debido a las estrictas pruebas de regulación a las que son sometidos. A veces, también, son comprados por otras firmas antes de alcanzar la última etapa. La variable `year` indica el año en el cual ese remedio alcanzó, de todas las fases que haya traspasado, la fase más avanzada con la empresa respectiva.

La segunda base se obtiene de Deloitte RECAP IQ. Ésta contiene las fusiones y adquisiciones de las empresas farmacéuticas, desde 1988 hasta 2011. Son 3.763 casos, en los cuales se detalla el comprador (`licensor`), el comprado (`licensee`) y la fecha (`date`), entre otras.⁴

Es necesario, con el fin de realizar los modelos probit, que cada observación esté conformada por una combinación de dos firmas, i y j , y que se indique si hubo fusión o adquisición, y si compartieron en algún momento un target genético, de área terapéutica y/o de enfermedad. De las $n = 514$ firmas pertenecientes a la base de Thomson Reuters Cortellis Life Sciences, se obtiene un total de $\frac{n \times (n-1)}{2}$ combinaciones posibles, equivalente a 131.841 pares de firmas.

⁴Debido a que las dos bases provienen de fuentes diferentes, el nombre de una misma firma podría haber estado escrito de forma distinta. Es por esta razón que en ambas se crea una nueva variable denominada `matchingname`, la cual homogeniza los nombres de las firmas. Este trabajo fue realizado manualmente en Microsoft Excel, y puede observarse en la tabla A.1 en el Apéndice, donde la primer columna indica el nombre que la firma tenía en la base de Thomson Reuters Cortellis Life Sciences, la segunda muestra el que tenían en Deloitte RECAP IQ y la tercera enseña aquel utilizado para unir las bases.

Estos pares de firmas luego son conectados con la base de Deloitte RECAP IQ, obteniendo así la cantidad de fusiones y adquisiciones total.

La tabla 1.1 muestra un resumen de la nueva base de datos generada. De un total de 131.841 combinaciones de firmas, solo 44 de ellas se fusionaron, 2.550 comparten target genético, 14.615 target de enfermedad y 47.742 de área terapéutica. En otras palabras, la cantidad de firmas que comparten target genético es considerablemente menor que las que comparten otro tipo de target (enfermedad y/o área terapéutica). Aproximadamente, de cada 6 pares de firmas que comparten target de enfermedad, 1 par comparte target genético; mientras que de cada 20 pares firmas que comparten área terapéutica, 1 par comparte gen.⁵

	F&A	TargetGen	Enfermedad	Área Terapéutica
Número de observaciones con	44	2.550	14.615	47.742

Tabla 1.1: Resumen de la Base de Datos

Posteriormente se desagregó la data por años, los cuales abarcaron desde 1995 hasta 2012. Allí, para cada año se consideraron las fusiones o adquisiciones realizadas hasta aquel en observación, y para crear las variables se consideraron los proyectos realizados hasta ese mismo año (incluido). A su vez, se actualizaban las firmas a medida que se iban fusionando, es decir, si la firma Schering-Plough adquiría a Canji en el año 1995, entonces para el año 1996 esta última ya no se consideraría, y de allí en adelante solo aparecería la primera en representación de ambas.

Junto con ello se realizó un análisis de cercanía entre genes presentes en la base de Thomson Reuters Cortellis Life Sciences, en donde se observaba cuántas enfermedades tenía en común cada par. Mayor cantidad de enfermedades implicaba mayor cercanía. Esto permitió crear tres nuevas variables: una que calculaba el promedio de la cercanía entre los genes de las firmas (meangene), otra la mediana (mediangene) y por último la suma (sumgene). En la tabla 1.2 se puede apreciar un resumen descriptivo de estas tres variables, considerando la variación temporal desde 1995 a 2012. Claramente en los tres casos las variables tienen un valor cercano a cero, lo que implica que las firmas trabajan con genes que, generalmente, atacan a distintas enfermedades.

Variable	Obs	Promedio	Desv. Est.	Min	Max
sumgene (suma)	2.238.346	0,2847384	2,733751	0	199
meangene (promedio)	2.238.346	0,142405	1,494909	0	159
mediangene (mediana)	2.238.346	0,1264081	1,455339	0	159

Tabla 1.2: Estadísticas descriptivas de sumgene, meangene y mediangene

Para finalizar se genera una última variable de cercanía entre las firmas, proveniente de la teoría de grafos. Esta contempla los conjuntos de intersección y de unión de nodos en común, y se define como Distancia *Wallis* = $1 - \frac{\text{intersección}}{\text{unión} - \text{intersección}}$. Para este caso en particular, los nodos serían genes y estarían unidos por un arco en caso de compartir al menos una enfermedad. De acuerdo con Bunke y Messmer (1997), el matching de grafos a partir de la corrección de

⁵Notar que los conjuntos no son disjuntos, i.e., puede que hayan pares de firmas que cumplan con los tres casos.

errores (o error correcting graph matching) puede ser utilizado como una medida de distancia indicando la diferencia entre los grafos. Consecuentemente, Bunke y Shearer (1998) definen una métrica de distancia basada en el máximo común subgrafo⁶, que se demuestra que está relacionada al matching de grafos a partir de la corrección de errores (Bunke, 1997). Por lo tanto, el máximo común subgrafo puede ser utilizado para indicar una medida de distancia. Para hallarlo, se utiliza una función biyectiva que encuentra estructuras comunes entre las dos redes, sin embargo en cuanto a genes respecta, no es importante que tengan la misma estructura, sino que sean los mismos genes. Por ende la función debe ser la identidad, lo que como consecuencia genera que el máximo común subgrafo sea la intersección entre ambos grafos. Por otra parte, el denominador de esta medida lo que hace es medir el tamaño de las redes en comparación. Si hay mayor intersección entre dos conjuntos pequeños, entonces esta distancia debería ser menor a la misma intersección entre dos redes de gran tamaño.

A modo de ejemplo, al año 2002, las firmas Bayer y Collateral Therapeutics trabajaban con dos genes en común; sin embargo, la primera tenía un total de cinco terapias genéticas mientras que la segunda solo dos. Por lo tanto, la Distancia Wallis entre estas firmas, al año 2002, sería de $1 - \frac{2}{(5+2)-2} = 0,6$ (notar que si ambas firmas intersectaran en todas sus terapias genéticas, la distancia sería igual a 0). La sección 2.4 de este trabajo explica en detalle esta variable, mientras que la tabla 1.3 presenta sus estadísticas descriptivas, contemplando las 131.841 combinaciones de firmas sin variación temporal y las 2.238.346 combinaciones con variación temporal. Consecuente con las otras variables, esta también tiene un valor muy cercano a cero, mostrando la baja intersección de genes entre firmas.

Distancia Wallis	Obs	Promedio	Desv. Est.	Min	Max
Sin variación temporal	131.841	0,9923994	0,0713896	0	1
Con variación temporal	2.238.346	0,9970104	0,0471609	0	1

Tabla 1.3: Estadísticas descriptivas de la Distancia de Wallis

Al graficar la red completa, la cantidad de nodos es tal que no se alcanzan a distinguir unos de otros. Por tanto, con el fin de mostrar cómo funciona este grafo la figura 1.1 muestra un subconjunto de los nodos. Cada nodo corresponde a un gen, y están unidos a través de un arco en caso de que compartan al menos una enfermedad en común. Si existiesen dos firmas, A y B, donde A trabajara con los genes 7157, 5743, 3569 y 6648; y B trabajara con 7124, 353 y 3569, su Distancia Wallis sería de $1 - \frac{1}{(4+3)-1} = 0,83$.

⁶El máximo común subgrafo es aquella subred isomorfa entre dos grafos que contiene la mayor cantidad de nodos posible.

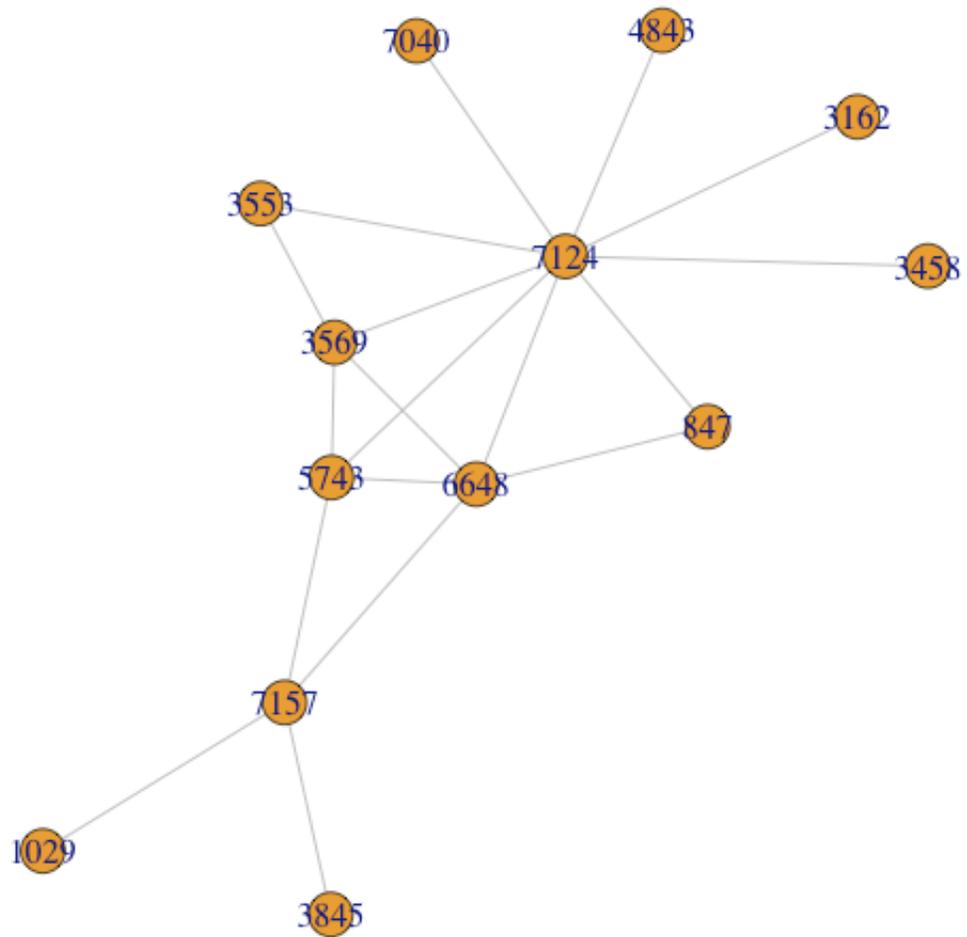


Figura 1.1: Subconjunto de genes, a modo de ejemplificar la idea de la red. Nodos son los genes, y los arcos son las enfermedades que tienen en común.

Capítulo 2

Modelos

La nueva era genética se caracteriza por firmas que trabajan en desarrollar remedios cuyo objetivo (o target) ahora es el gen relacionado a la enfermedad a tratar. Con el fin de comprobar si las fusiones o adquisiciones pueden ser explicadas por intersecciones en el target genético de los remedios, se estima un modelo probit con diversos tests de robustez.

2.1. Modelo Probit agregado

El primer análisis está enfocado en obtener los primeros indicios acerca del target genético. La idea es, mediante un modelo sin muchas restricciones, comprobar que el target genético podía explicar las fusiones o adquisiciones entre empresas. Éste contempla todos los proyectos sin restringir el año en el que fueron realizados, donde y_{ij} es la variable dependiente, que representa la fusión o adquisición entre las firmas i y j ⁷, y viene descrita por:

$$y_{ij} = \beta_0 + \beta_1 \text{TargetGen}_{ij} + \beta_2 \text{ÁreaTerapéutica}_{ij} + \beta_3 \text{Enfermedad}_{ij} + \varepsilon_{ij}$$

donde TargetGen_{ij} es una variable binaria igual a 1 en el caso en que las firmas i y j , en cualquier momento, hayan tenido algún proyecto con el mismo target genético. La lógica es la misma para las variables $\text{ÁreaTerapéutica}_{ij}$ y Enfermedad_{ij} , solo que aplicado al área terapéutica y a la enfermedad atacadas por las firmas, respectivamente. Debido a la escasa variabilidad que existe para cada empresa (son muy pocas las que se fusionan por lo que, en general, casi todas tienen $y_{ij} = 0$ en todas sus filas) no se agrega efecto fijo por firma. Se espera que β_1 , β_2 y β_3 sean positivos y significativos, ya que es lógico intuir que una firma sea más propensa a adquirir a otra cuando comparten mismos targets en genes, enfermedades y/o áreas terapéuticas. A su vez, se cree que el target genético tendrá el mayor módulo de los tres, siendo el más importante de los factores.

⁷Para efectos de este modelo y todos los que siguen, es irrelevante si la firma i absorbe a la j o si es j quien absorbe a i .

Debido a que este modelo contempla todas las fusiones y adquisiciones desde 1988 hasta 2011, y todos los proyectos genéticos desde 1989 hasta 2014, se pueden presentar casos en los cuales primero ocurrió una fusión y posterior a este hecho ambas firmas comenzaron a trabajar el mismo target genético, de enfermedad o de área terapéutica. Esto, por ende, puede generarle problemas de endogeneidad al modelo. Para poder solucionarlo, se añade variación temporal al modelo, desagregando los datos por año. De esta forma se evita el problema enunciado anteriormente.

2.2. Modelo Probit con variación temporal

El modelo anterior podía dilucidar cierta tendencia del target genético para con las fusiones o adquisiciones, pero no refleja el impacto que ésta tiene en realidad. En este modelo se construye la data de tal forma que represente la variación temporal que afectó a las firmas a lo largo de los años. Ahora, por tanto, se considera el año en que los proyectos fueron realizados y también cuándo las firmas decidieron fusionarse. La variable dependiente es y_{ijt} representa la fusión o adquisición entre las firmas i y j en el año t , y se describe:

$$y_{ijt} = \beta_0 + \beta_1 \text{TargetGen}_{ijt} + \beta_2 \text{ÁreaTerapéutica}_{ijt} + \beta_3 \text{Enfermedad}_{ijt} \\ + \beta_4 \text{Sumgene}_{ijt} + \beta_5 \text{Meangene}_{ijt} + \beta_6 \text{Mediangene}_{ijt} + \varepsilon_{ijt}$$

donde la especificación de las variables es la misma que en el modelo anterior, pero se agregan Sumgene_{ijt} , Meangene_{ijt} y Mediangene_{ijt} que corresponden a la suma, promedio y mediana de la distancia entre los genes que posee cada par de firmas, respectivamente. Ahora, a su vez, se agrega el subíndice t que representa la variación temporal, desde el año 1995 hasta el año 2012. Al igual que en el caso anterior, no se incluye en el modelo efecto fijo por firma, ni tampoco por año (esto último producto que en cada año la variabilidad de fusiones es muy baja). La hipótesis correspondiente a β_1 , β_2 y β_3 es análoga a la descrita para el caso anterior. Con respecto a β_4 , β_5 y β_6 se espera que sean positivas debido a la misma lógica empleada con las variables binarias: cuanto mayor sea la cantidad de targets genéticos en común, mayor será la probabilidad de fusión. Con respecto a su significancia, se espera que lo sean cuando expliquen por si solas a la variable dependiente, pero en el caso de estar incluidas en el modelo junto con la variable binaria TargetGen , se cree que no serán significativas debido a la colinealidad con ella.

2.3. Test de Robustez: Probit Dinámico

2.3.1. Análisis por tamaño de firmas

Se consideran las 23 firmas con mayor cantidad de ingresos, de acuerdo con MEPS⁸, en el área farmacéutica como las firmas grandes. Todas aquellas que no son grandes se consideran pequeñas. Se determina, de este grupo, cuáles tienen participación en la terapia genética. La tabla 2.1 muestra cuáles son estas firmas grandes (en total 16), la cantidad de fusiones y adquisiciones y la cantidad de targets genéticos compartidos por este subconjunto de firmas.

	F&A	TargetGen
Abbott	0	26
AstraZenaca	0	46
Boehringer Ingelheim	1	22
Bristol-Myers Squibb	1	14
Daiichi Sankyo	1	13
Eisai	1	3
Eli Lilly & Co	0	21
GlaxoSmithKline	3	3
Johnson & Johnson	1	15
Merck	3	31
Novartis	1	15
Pfizer	2	23
Roche	1	16
Sanofi	3	19
Shire	0	10
Teva Pharmaceutical	2	2

Tabla 2.1: Cantidad de fusiones y adquisiciones, y cantidad de target genéticos compartidos con una o más firmas, para cada una de las firmas grandes

Este modelo solo evalúa las fusiones o adquisiciones entre firmas grandes y pequeñas, ignorando todos los otros casos posibles. Por lo tanto se puede describir como:

$$y_{ijt} = \beta_0 + \beta_1 \text{TargetGen}_{ijt} + \beta_2 \text{ÁreaTerapéutica}_{ijt} + \beta_3 \text{Enfermedad}_{ijt} \\ + \beta_4 \text{Sumgene}_{ijt} + \beta_5 \text{Meangene}_{ijt} + \beta_6 \text{Mediangene}_{ijt} + \varepsilon_{ijt}$$

con i siendo firma grande y j firma pequeña, o viceversa.

⁸Medical Expenditure Panel Survey, como se definen ellos, es “un conjunto de encuestas a gran escala de familias e individuos, sus proveedores médicos y empleadores en los Estados Unidos. MEPS es la fuente más completa de datos sobre el costo y el uso de la cobertura de seguro médico y de salud”.

2.3.2. Análisis dinámico

Debido a que la data de panel está segmentada anualmente, puede ocurrir que dos firmas tengan el mismo target genético, de enfermedad y/o de área terapéutica producto de que se fusionaron ese año. Por ejemplo, quizás la firma i y la firma j tienen el mismo target genético desde mediados de año producto de una fusión entre ellas aunque antes no lo tenían, sin embargo para la data, ambas firmas se fusionan y a su vez comparten el target genético desde comienzos de ese mismo año. Es directo ver que, en consecuencia, puede que exista un problema de endogeneidad. Para solucionarlo se realiza un modelo con variables rezagadas en un año, para lo cual las firmas pueden que se fusionen en un año, pero sus targets genéticos, de enfermedad y de área terapéutica serán los del año anterior. Este modelo contiene las mismas variables que los casos detallados en las secciones 2.2 y 2.3.1, y se especifica como sigue:

$$y_{ijt} = \beta_0 + \beta_1 \text{TargetGen}_{ij,t-1} + \beta_2 \text{ÁreaTerapéutica}_{ij,t-1} + \beta_3 \text{Enfermedad}_{ij,t-1} \\ + \beta_4 \text{Sumgene}_{ij,t-1} + \beta_5 \text{Meangene}_{ij,t-1} + \beta_6 \text{Mediangene}_{ij,t-1} + \varepsilon_{ijt}$$

Se realiza también un análisis análogo, pero con un modelo que examina tres años de rezago con el fin de hacer un estudio aún más robusto:

$$y_{ijt} = \beta_0 + \beta_1 \text{TargetGen}_{ij,t-3} + \beta_2 \text{ÁreaTerapéutica}_{ij,t-3} + \beta_3 \text{Enfermedad}_{ij,t-3} \\ + \beta_4 \text{Sumgene}_{ij,t-3} + \beta_5 \text{Meangene}_{ij,t-3} + \beta_6 \text{Mediangene}_{ij,t-3} + \varepsilon_{ijt}$$

Es de esperar que, de ser el modelo robusto, se mantengan los signos y significancia de los coeficientes observados en el probit con variación temporal para todos estos tests de robustez.

2.4. Modelo Probit con variación temporal, y métrica de distancia de Wallis

Los modelos empíricos previos establecen que la probabilidad de fusión de dos firmas pueda estar relacionada con que ambas posean investigaciones genéticas en común. Sin embargo, TargetGen es dicotómica, por lo que solo captura relaciones discretas. El siguiente paso es explorar especificaciones que permitan una relación funcional entre una medida continua de distancia en la estructura de la red y las decisiones de fusiones. Las variables Sumgene, Meangene, Mediangene pueden mostrar alguna propensión en la cantidad de genes en común, pero son meramente intuitivos, no tienen ninguna investigación (o justificación) detrás que los avale como buenos indicadores.

Lo que falta, entonces, es una variable con sustento teórico y que permita modelar la relación continua. Para ello, es necesario introducir a la investigación la teoría de grafos. La idea es estructurar los datos como una gran red, en la cual los genes serían los nodos y las enfermedades que tienen en común determinarían los arcos de esta. La figura 1.1 muestra un subconjunto de genes y su conexión a través de las enfermedades con las que se relacionan.

Un detalle no menor es que basta con que los genes tengan al menos una enfermedad en común para que el arco exista.

Esta estructura (genes como nodos y enfermedades como links) puede apreciarse en Goh et al. (2007), en donde se ordena de esta forma para aprovechar la asociación fenotípica que un link representa para dos genes. De esta manera, esta disposición ayuda a calcular la relación fenotípica entre los genes, y así poder analizar su cercanía. Una vez teniendo la correspondencia entre genes, puede determinarse una relación entre las firmas, como se explica más adelante en esta sección.

Las firmas, entonces, serían subgrafos dentro de esta gran red de genes y enfermedades, en donde cada una está delimitada por los genes con los que trabaja. Es necesario por lo tanto calcular la cercanía entre estas firmas, es decir, se requiere la distancia entre dos subgrafos.

Una métrica de distancia debe cumplir con cuatro propiedades. Sean G_1, G_2 y G_3 tres grafos, entonces se debe cumplir que:

1. $0 \leq d(G_1, G_2) \leq 1$,
2. $d(G_1, G_2) = 0 \iff G_1$ y G_2 son isomorfos⁷ entre ellos,
3. $d(G_1, G_2) = d(G_2, G_1)$,
4. $d(G_1, G_3) \leq d(G_1, G_2) + d(G_2, G_3)$.

Diversas medidas de distancia se han definido a lo largo de los años. Para poder comprenderlas a cabalidad, primero es necesario definir algunos conceptos relativos a la teoría de grafos¹⁰.

Definición 2.4.1. Sea L un alfabeto finito de etiquetas para nodos y arcos. Un grafo es una tupla $G = (N, A, \alpha)$, donde

- N es un conjunto finito de nodos,
- $A \subseteq N \times N$ es un conjunto finito de arcos,
- $\alpha : N \cup A \rightarrow L$ es la función de etiquetado de nodo y arco.

Definición 2.4.2. Un grafo $G_1 = (N_1, A_1, \alpha_1)$ es un subgrafo de un grafo $G_2 = (N_2, A_2, \alpha_2)$, denotado como $G_1 \subseteq G_2$, si

- $N_1 \subseteq N_2$,
- $A_1 \subseteq A_2$,
- $\alpha_1(x) = \alpha_2(x) \forall x \in N_1 \cup A_1$.

El grafo G_2 es denominado supergrafo de G_1 si $G_1 \subseteq G_2$.

Definición 2.4.3. Sea $G = (N, A, \alpha)$ un grafo, N_2 un conjunto de nodos, y $f : N_1 \rightarrow N_2$ una función biyectiva, donde $N_1 \subseteq N$ y $N_1 \cap N_2 = \emptyset$. Un renombramiento de G a partir de f , denotado por $f(G)$, es el grafo $G_f = (N_f, A_f, \alpha_f)$ definido por

⁹Para saber qué es un isomorfismo, remitirse a la definición 2.4.4 enunciada más abajo.

¹⁰Estas definiciones se obtienen de Wallis et al. (2001).

- $N_f = N_2 \cup (N \setminus N_1)$,
- $A_f = \{(f'(u), f'(v)) \mid (u, v) \in A\}$,
- $\alpha_f(f'(n)) = \alpha(n) \forall n \in N$,
- $\alpha_f(a) = \alpha(a) \forall a \in A$.

donde $f' : N \rightarrow N_f$ se define como $f'(n) = \begin{cases} f(n) & \text{si } n \in N_1 \\ n & \text{cualquier otro caso} \end{cases}$

Definición 2.4.4. Dos grafos G_1 y G_2 son isomorfos, denotados por $G_1 \cong G_2$, si existe un renombramiento f de G_1 tal que $f(G_1) = G_2$. En este caso se dice que $f : G_1 \rightarrow G_2$ es un isomorfismo de grafo.

Definición 2.4.5. Un grafo \hat{G} es un común subgrafo de dos grafos G_1 y G_2 si existen subgrafos $\hat{G}_1 \subseteq G_1$ y $\hat{G}_2 \subseteq G_2$ tal que $\hat{G} \cong \hat{G}_1 \cong \hat{G}_2$. Es el máximo si no existe otro común subgrafo de G_1 y G_2 más grande que \hat{G} .

Definición 2.4.6. Un grafo \check{G} es un común supergrafo de dos grafos G_1 y G_2 si existen grafos $\check{G}_1 \subseteq \check{G}$ y $\check{G}_2 \subseteq \check{G}$ tal que $\check{G}_1 \cong G_1$ y $\check{G}_2 \cong G_2$. Es el mínimo si no existe otro común supergrafo de G_1 y G_2 más pequeño que \check{G} .

Fernández y Valiente (2001) definen la distancia de forma bastante particular. Sea MCS el mínimo común supergrafo y mcs el máximo común subgrafo, se tiene que la distancia entre dos grafos es:

$$d(G_1, G_2) = |MCS(G_1, G_2)| - |mcs(G_1, G_2)|$$

Por otra parte, Bunke y Shearer (1998) definen la distancia como:

$$d(G_1, G_2) = 1 - \frac{|mcs(G_1, G_2)|}{\max(|G_1|, |G_2|)}$$

De acuerdo con Wallis, Shoubridge, Kraetz, Ray (2001), la métrica de distancia basado en el máximo común subgrafo es de la forma:

$$d(G_1, G_2) = 1 - \frac{m(G_1, G_2)}{M(G_1, G_2)}$$

donde $m(G_1, G_2)$ es una medida de la similaridad entre los grafos G_1 y G_2 y $M(G_1, G_2)$ es una medida del tamaño del problema. Utilizando como ejemplo el caso de Bunke y Shearer (1998), se tiene que $m(G_1, G_2) = |mcs(G_1, G_2)|$ y $M(G_1, G_2) = \max(|G_1|, |G_2|)$.

Por su parte, Wallis et al. (2001) derivan una fórmula que utiliza la cardinalidad del máximo común subgrafo y la cardinalidad de la unión de los subgrafos en cuestión, para determinar la distancia entre ellos. Observando la definición 2.4.5, se puede apreciar que el máximo común subgrafo se determina a partir de una función biyectiva que encuentra

estructuras comunes entre estas dos subredes; sin embargo para comparar distancia con respecto a genes no es suficiente que dos subgrafos tengan la misma estructura, sino que es imperioso que tengan los mismos genes. Por tanto, la función para el caso particular de este trabajo de tesis sería la identidad y, debido a que cada gen aparece una sola vez en esta red, entonces el máximo común subgrafo será la intersección entre los dos subgrafos.

La métrica que demuestra Wallis et al. (2001) en su paper, para determinar la distancia entre los grafos G_1 y G_2 , es:

$$d(G_1, G_2) = 1 - \frac{|mcs(G_1, G_2)|}{|G_1| + |G_2| - |G_{12}|}$$

donde $|G|$ es la cardinalidad del grafo (cantidad de nodos pertenecientes a él), y por ende $|G_{12}|$ sería la cardinalidad de la intersección entre los grafos G_1 y G_2 . Es necesario recalcar nuevamente que para el caso específico que se estudia en esta investigación $|mcs(G_1, G_2)| = |G_{12}|$.

A su vez, se puede ver que Wallis et al. (2001) definen una medida en la que, cuantos más genes tengan en común dos firmas, la operación $\frac{|mcs(G_1, G_2)|}{|G_1| + |G_2| - |G_{12}|}$ será más cercana a 1, por lo que la distancia de Wallis entre estas firmas se aproximará más a 0. Aprovechando esta definición, se denomina la métrica empleada en este trabajo de tesis para las firmas i y j como

$$\text{DistanciaWallis}_{ij} = 1 - \frac{|mcs(G_i, G_j)|}{|G_i| + |G_j| - |G_{ij}|} = 1 - \frac{|\cap_{ij}|}{|\cup_{ij}| - |\cap_{ij}|}$$

donde G_i es el subgrafo que contiene los genes (nodos) con los que trabaja la firma i y G_j es el subgrafo análogo para la firma j . \cap_{ij} representa el conjunto intersección de los nodos entre estos dos subgrafos, mientras que \cup_{ij} corresponde al conjunto unión de los nodos de estos mismos.

Esta última se elige por sobre las otras dos por diversas razones. La distancia de Fernández y Valiente (2001) no establece una diferencia clara entre dos firmas que tengan genes en común y dos que no; ya que su valor siempre será (para este caso en particular) la cardinalidad de la unión del conjunto de genes con el que trabaja cada firma, menos la cardinalidad de la intersección. En cambio, las otras dos métricas son siempre distintas de 1 en caso de que intersecten los conjuntos, y serán siempre 1 en caso de que no. Por otra parte, la métrica definida por Bunke y Shearer (1998) no penaliza por el tamaño de los dos subgrafos, sino que solo por el más grande. A modo de ejemplo, suponiendo que existan dos grupos, A y B, de dos firmas cada uno que intersectan en la misma cantidad de genes, y cuyo subgrafo más grande sea del mismo tamaño, y el subgrafo más pequeño de A tenga el doble de genes que el más pequeño de B, entonces para esta métrica la distancia será la misma; sin embargo, para Wallis et al. (2001), la distancia entre las firmas del grupo A será mayor que la del B. En consecuencia, se prefiere trabajar con la distancia de Wallis et al. (2001).

Resta, por tanto, demostrar que esto es una métrica. Observando las cuatro propiedades

que debe cumplir una métrica, se puede apreciar que los primeros tres puntos son triviales para la Distancia de Wallis, y solo el cuarto pareciera mostrar cierta dificultad¹¹.

Considerando tres grafos G_1, G_2 y G_3 , se tiene que:

$$\begin{aligned} M(G_1, G_3) &= |G_1| + |G_3| - |G_{13}| \\ &\leq |G_1| + |G_3| = |G_1 \cup G_3| \end{aligned}$$

Por lo que

$$\begin{aligned} d(G_1, G_3) &= 1 - \frac{|G_{13}|}{|G_1| + |G_3| - |G_{13}|} \\ &\leq 1 - \frac{|G_1 \cap G_3|}{|G_1 \cup G_3|} \end{aligned}$$

Luego reemplazando en $d(G_1, G_2) + d(G_2, G_3) \geq d(G_1, G_3)$, se tiene que

$$1 - \frac{|G_1 \cap G_2|}{|G_1 \cup G_2|} + 1 - \frac{|G_2 \cap G_3|}{|G_2 \cup G_3|} \geq 1 - \frac{|G_1 \cap G_3|}{|G_1 \cup G_3|} \quad (2.1)$$

Ahora bien, el tamaño de cada conjunto se representará como:

- Hay x_{123} nodos en $G_1 \cap G_2 \cap G_3$,
- Hay x_{12} nodos en $G_1 \cap G_2$ pero no en G_3 ,
- Hay x_1 nodos en G_1 pero no en $G_2 \cup G_3$,

y así sucesivamente. Se define T como el numero total de nodos,

$$T = x_1 + x_2 + x_3 + x_{12} + x_{13} + x_{23} + x_{123}$$

Luego, la ecuación 2.1 sería:

$$1 - \frac{x_{12} + x_{123}}{T - x_3} + 1 - \frac{x_{23} + x_{123}}{T - x_1} \geq 1 - \frac{x_{13} + x_{123}}{T - x_2}$$

o bien

$$1 + \frac{x_{13} + x_{123}}{T - x_2} - \frac{x_{12} + x_{123}}{T - x_3} - \frac{x_{23} + x_{123}}{T - x_1} \geq 0$$

Despejando los denominadores se tiene:

$$\begin{aligned} (T - x_1)(T - x_2)(T - x_3) &+ (x_{13} + x_{123})(T - x_1)(T - x_3) \\ &- (x_{12} + x_{123})(T - x_1)(T - x_2) \\ &- (x_{23} + x_{123})(T - x_2)(T - x_3) \geq 0 \end{aligned} \quad (2.2)$$

¹¹Esta demostración se obtiene de Wallis, W. D., Shoubridge, P., Kraetz, M., and Ray, D. Graph distances using graph union. Pattern Recognition Letters 22 (2001), p. 703.

El lado izquierdo de esta ecuación se puede escribir como un polinomio cúbico de siete variables con todos términos positivos, por ejemplo:

$$(2T - x_1 + x_3)(x_{13}T + x_2x_{123}) + T[(x_2 + x_3)x_{23} + (x_1 + x_2)x_{12} + x_1x_3] \\ + (T - x_{12} - x_3)x_1x_2 + (T - x_{23})x_2x_3 + x_1x_3(x_{13} + x_{123})$$

Luego la ecuación 2.2, y en consecuencia la ecuación 2.1 siempre se cumplen. Por ende, la Distancia de Wallis es una métrica.

El modelo que incluye a este variable es:

$$y_{ijt} = \beta_0 + \beta_1 \text{TargetGen}_{ijt} + \beta_2 \text{ÁreaTerapéutica}_{ijt} + \beta_3 \text{Enfermedad}_{ijt} \\ + \beta_4 \text{DistanciaWallis}_{ijt} + \varepsilon_{ijt}$$

Donde todas las variables restantes mantienen lo explicado anteriormente en la sección 2.2.

Para este caso en particular, se espera que β_4 tenga signo negativo. Esto debido a que cuantos más genes en común trabajen las firmas, la Distancia de Wallis será menor. Por tanto, es de esperar a que tenga signo contrario a las demás covariables. A su vez, se espera que sea significativa siempre y cuando no esté en el modelo incluida la variable TargetGen, debido a la inminente colinealidad que puede haber entre estas dos covariables.

Finalmente, con el fin de comprender un poco mejor la Distancia de Wallis se presenta la figura 2.1. El grafo G_1 muestra dos subgrafos, g y g_{12} , los cuales intersectan en todos los nodos del primero pero no en todos los nodos del segundo, por lo tanto la Distancia de Wallis se encontraría entre 0 y 1. Así mismo sucede con el grafo G_2 , donde g_{12} y g_{23} intersectan en un subconjunto de nodos denominado g , por lo que la distancia también se encontraría entre 0 y 1. Por último, se tiene el grafo G_3 , el cual es un caso extremo debido a que los subgrafos no se intersectan. En este escenario, la Distancia de Wallis tendría valor igual a 0.

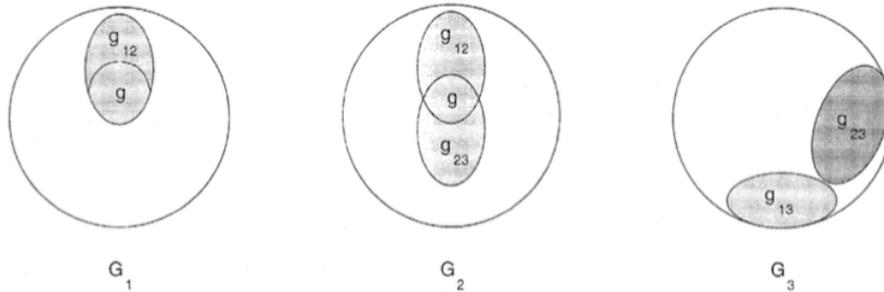


Figura 2.1: Ejemplos de subgrafos con un caso extremo y dos casos intermedios de Distancia de Wallis¹⁰

¹²Fuente: Bunke, H., and Shearer, K. A graph distance metric based on the maximal common subgraph. Pattern Recognition Letters, 19 (1998), p. 258.

2.5. Test de Robustez e interacción con la evolución del proyecto

2.5.1. Análisis por tamaño de firmas: Probabilidad de adquisición

Este modelo es muy parecido al de la sección 2.3.1; la diferencia radica en que ahora solo se evalúan las adquisiciones (ya que la probabilidad de que se fusionen una firma grande y una chica es muy baja. Solo representan 3 de 44 casos) y se agrega la distancia de Wallis, para evaluar su robustez.

El modelo se describe como sigue:

$$y_{ijt} = \beta_0 + \beta_1 \text{TargetGen}_{ijt} + \beta_2 \text{ÁreaTerapéutica}_{ijt} + \beta_3 \text{Enfermedad}_{ijt} \\ + \beta_4 \text{DistanciaWallis}_{ijt} + \varepsilon_{ijt}$$

2.5.2. Interacción con la evolución del proyecto

Los fármacos, para llegar al mercado, pasan por cinco etapas distintas: Discovery (etapa 0), Phase I (etapa 1), Phase 2 (etapa 2), Phase 3 (etapa 3) y Market (etapa 4). Lo que se busca con este modelo, es permitir la interacción entre las variables ya descritas, y las etapas de evolución que el proyecto va teniendo. De esta forma, se puede evaluar no solo si las firmas eligen adquirir a otra debido al target genético, enfermedad o área terapéutica, sino que también se puede evaluar si la etapa en la cual está el proyecto (que genera el interés de una firma para adquirir a otra) es relevante al momento de tomar la decisión. Para este modelo, al igual que para el anterior, se considera solo la interacción entre las firmas grandes y pequeñas. La etapa evaluada es la correspondiente al proyecto de la firma pequeña. En primera instancia, se coloca la etapa como efecto fijo, junto con las otras variables, y posteriormente se desarrollan dos modelos en los cuales interactúan las etapas junto con el target genético y otro en el que interactúan con la distancia de Wallis.

Los modelos son los siguientes:

$$y_{ijt} = \beta_0 + \beta_1 \text{TargetGen}_{ijt} + \beta_2 \text{ÁreaTerapéutica}_{ijt} + \beta_3 \text{Enfermedad}_{ijt} \\ + \beta_4 \text{DistanciaWallis}_{ijt} + \beta_5 \text{EtapaProyecto}_{ijt} + \varepsilon_{ijt}$$

$$y_{ijt} = \beta_0 + \beta_1 \text{TargetGen}_{ijt} \times \text{EtapaProyecto}_{ijt} + \varepsilon_{ijt}$$

$$y_{ijt} = \beta_0 + \beta_1 \text{DistanciaWallis}_{ijt} \times \text{EtapaProyecto}_{ijt} + \varepsilon_{ijt}$$

Capítulo 3

Resultados

Los modelos empíricos vistos en la sección anterior buscaban dilucidar si existía una incidencia del target genético de los remedios que producen las firmas en la decisión de fusionarse o adquirir a otra empresa. A continuación se presentan los resultados del primer modelo.

3.1. Modelo Probit agregado

Para este caso se analizó un modelo con cada variable por separado y luego se realizó un probit con todas ellas juntas. La idea era ver las diferencias entre los efectos individuales que tenía cada una y posteriormente observar cómo variaba la significancia si las tres eran consideradas para explicar las fusiones o adquisiciones. La tabla 3.1 muestra estos cuatro escenarios, solamente para cuando las variables explicativas tienen valor 1, donde una estrella implica que la variable es significativa con una confianza del 90 %, dos estrellas una confianza del 95 % y tres estrellas una confianza del 99 %.

	(1)	(2)	(3)	(4)
	Probit target	Probit indication	Probit therapeutic	Probit all variables
MAs				
1.target	1,230*** (0,0956)			1,109*** (0,104)
1.indication		0,512*** (0,0875)		0,162 (0,117)
1.therapeutic			0,421*** (0,0902)	0,166 (0,116)
<i>N</i>	131.841	131.841	131.841	131.841
pseudo <i>R</i> ²	0,153	0,037	0,031	0,163

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.1: Resultados Modelo Probit Agregado

Como se puede apreciar, las tres variables son significativas con un p-valor menor al 0,01 cuando explican por separado las fusiones o adquisiciones. Sin embargo, cuando se juntan las tres para explicar la variable dependiente, solo el target genético permanece significativo. Es válido notar también que el pseudo R^2 es considerablemente mayor para el modelo en el que se encuentra solo el target genético, en comparación a los modelos en los que se encuentran únicamente el target por enfermedad y el target por área terapéutica. Los coeficientes β_1 , β_2 y β_3 resultaron ser positivos, como se había predicho y efectivamente aquel que acompaña a TargetGen posee el mayor módulo de los tres.

Debido a que este es un modelo no lineal, más importante que ver la significancia de las variables en cada caso, es observar sus efectos marginales. La tabla 3.2 muestra los efectos marginales para el modelo en el que están las tres variables juntas, y debido a que son dicotómicas se analizan los escenarios donde tienen valor 0 y también cuando tienen valor 1. Se puede observar que la probabilidad de que exista una fusión o adquisición entre dos empresas aumenta casi treinta y nueve veces en caso de que haya al menos un target genético en común entre ellas, mientras que para los casos de enfermedad y área terapéutica este aumento es de casi el doble, respectivamente. Yendo estrictamente a los números, si dos empresas comparten al menos un target genético, entonces su probabilidad de fusionarse es de un 0,6%, por otra parte, la probabilidad es de un 0,05% y de un 0,04% en el caso de target por enfermedades y área terapéutica. Por lo tanto la probabilidad inducida por el target genético es doce veces más alta que para el target de enfermedades y quince veces más alta que para el de área terapéutica.

	(1)
	Efectos Marginales Probit agregado
0.target	0,000164*** (0,0000359)
1.target	0,00637*** (0,00159)
0.indication	0,000279*** (0,0000575)
1.indication	0,000475*** (0,000139)
0.therapeutic	0,000236*** (0,0000708)
1.therapeutic	0,000409*** (0,0000834)
<i>N</i>	131.841

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.2: Resultados Efectos Marginales Probit Agregado

Este análisis es importante para tener una primera impresión de cómo influyen, por lo

menos parcialmente, las variables explicativas en la variable dependiente. A pesar de tener resultados positivos con respecto al target genético, es necesario darle una estructura más realista a la base de datos y observar si existe algún cambio en los efectos de las variables hacia las fusiones al agregar variación temporal. Para ello es que se crea el segundo modelo, cuyos resultados se analizan a continuación.

3.2. Modelo Probit con variación temporal

Basado en la sección 1 de Datos, ahora se realiza el análisis de fusiones y adquisiciones en el tiempo. Las empresas que son adquiridas van “desapareciendo” y solo quedan las firmas compradoras; a medida que avanza el tiempo, la cantidad de combinaciones entre firmas es menor. Al mismo tiempo, se consideran los proyectos hasta la fecha, por lo que puede que aparezcan targets genéticos, de enfermedades o de áreas terapéuticas en común entre firmas cuando antes no había.

La tabla 3.3 muestra la significancia de siete versiones de este modelo, donde se tienen las tres variables principales en un principio por separado [(1),(2) y (3)], todas juntas [(4)] y luego agregándole las tres variables intuitivas de continuidad para target genético [sumgene (5), meangene (6), y mediangene (7)]. Se puede observar que cuando los targets (genético, por enfermedad y por área terapéutica) se evalúan por separado, son todos muy significativos, con un p-valor menor al 0,01. Este comportamiento es análogo a lo que ocurría en el modelo agregado, sin variación temporal. Sin embargo, ahora se aprecia que al tener estas tres variables juntas, target genético mantiene su alto nivel de significancia y el target por enfermedad es significativo con un 10 % de confianza. El target por área terapéutica, al igual que en el caso anterior, no es significativo para el modelo con todas las variables principales. Nuevamente, la intuición era correcta para este modelo, obteniendo coeficientes positivos, y se mantiene el módulo del coeficiente que acompaña a TargetGen como el mayor para cualquiera de los escenarios evaluados.

Como se señaló en la sección Modelos, no se incluyeron efectos fijos por firma ni por año debido a la poca variabilidad presente dentro de cada grupo. Efectivamente se probó incluyendo estos efectos fijos (por separado) y el resultado fue que ninguno era significativo para el modelo, y los coeficientes que acompañaban a los targets genéticos, de enfermedad y de área terapéutica no variaron su significancia ni signo, solo decrecieron levemente en módulo.

Al agregar ahora las variables “auxiliares” para target genético, se ve claramente cómo meangene y mediangene son significativas con p-valor de 0,1 y 0,05, respectivamente. Es necesario recalcar que el signo de estas últimas es negativo, lo que indica cierta colinealidad con el target genético. Para poder analizar con mayor precisión la implicancia de ellas en la variable dependiente se presenta la tabla 3.4 con sus efectos marginales.

Tal y como se veía en el modelo anterior, el efecto marginal del target genético genera que la probabilidad de fusión o adquisición incremente en este caso 65 veces más, aproximadamente, que cuando no hay genes en común. Si se observan los otros modelos, la probabilidad de fusión

	(1)	(2)	(3)	(4)
	Probit target	Probit indication	Probit therapeutic	Probit all 3
MAs				
1.target	1,223*** (0,0841)			1,063*** (0,111)
1.indication		0,718*** (0,0828)		0,334* (0,175)
1.therapeutic			0,470*** (0,0750)	0,0134 (0,164)
<i>N</i>	2.238.346	2.238.346	2.238.346	2.238.346
pseudo <i>R</i> ²	0,119	0,048	0,030	0,128
	(6)	(7)	(8)	
	Probit all+sum	Probit all+mean	Probit all+median	
MAs				
1.target	1,080*** (0,124)	1,103*** (0,117)	1,100*** (0,115)	
1.indication	0,339* (0,175)	0,335* (0,176)	0,331* (0,176)	
1.therapeutic	0,0133 (0,165)	0,0124 (0,166)	0,0139 (0,165)	
sumgene	-0,00153 (0,00265)			
meangene		-0,00819* (0,00450)		
mediangene			-0,00979** (0,00455)	
<i>N</i>	2.238.346	2.238.346	2.238.346	
pseudo <i>R</i> ²	0,128	0,129	0,129	

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.3: Resultados Modelo Probit con Variación Temporal

dado la presencia del target genético en común es mayor, debido a que debe contrarrestarse con el efecto negativo de las variables como meangene y mediangene, que son colineales. Esta colinealidad es lógica, ya que sumgene, meangene y mediangene existirán siempre y cuando haya al menos un target genético en común. Por lo tanto, cuando el target sea 1, ellas serán siempre mayores que 0 pero menores o iguales a 1. En consecuencia, se infiere que es considerablemente mayor el efecto dicotómico que el continuo para la variable dependiente.

Observando únicamente el caso (1), sin las tres variables continuas, el efecto marginal del target de enfermedad no es significativo, pero sí lo es el del área terapéutica. Este último indica que cuando se está en presencia de dos firmas que atacan al menos un mismo área terapéutica, entonces su probabilidad de fusión aumenta en un 0,002 %, mientras que si atacaran al mismo gen, su probabilidad aumentaría en un 0,077 %; lo que es considerablemente mayor.

Con respecto a los efectos marginales para las tres variables continuas en target genético, debido a que existía gran cantidad de valores para cada una de ellas, se optó por calcular

	(1)	(2)	(3)	(4)
	Margins all 3	Margins all+sum	Margins all+mean	Margins all+median
0.target	0,0000117*** (0,00000231)	0,0000117*** (0,00000231)	0,0000117*** (0,00000231)	0,0000117*** (0,00000231)
1.target	0,000765*** (0,000264)	0,000811*** (0,000313)	0,000876*** (0,000316)	0,000867*** (0,000307)
0.indication	0,0000153*** (0,00000307)	0,0000153*** (0,00000305)	0,0000153*** (0,00000307)	0,0000154*** (0,00000308)
1.indication	0,0000590 (0,0000361)	0,0000600 (0,0000367)	0,0000593 (0,0000365)	0,0000583 (0,0000359)
0.therapeutic	0,0000193*** (0,00000513)	0,0000193*** (0,00000514)	0,0000193*** (0,00000516)	0,0000193*** (0,00000513)
1.therapeutic	0,0000204** (0,00000933)	0,0000204** (0,00000935)	0,0000203** (0,00000937)	0,0000204** (0,00000940)
sumgene		-0,000000119 (0,000000205)		
meangene			-0,000000634* (0,000000358)	
mediangene				-0,000000758** (0,000000366)
<i>N</i>	2.238.346	2.238.346	2.238.346	2.238.346

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.4: Efectos Marginales Modelo Probit con Variación Temporal

el efecto marginal promedio. El efecto de sumgene no es significativo, pero meangene y mediangene sí lo son. El signo de ellos es negativo, por la colinealidad explicada anteriormente y sus valores, como muestra la tabla 3.4, son extremadamente pequeños.

La tabla 3.5 muestra tres modelos distintos, en donde se tienen por separado las variables sumgene, meangene y mediangene. Aquí se puede apreciar que las tres son altamente significativas y positivas, lo que reafirma la hipótesis de colinealidad. Al no estar presente la variable dicotómica de target genético, no tienen ningún efecto que contrarrestar por lo que su signo cambia y su significancia también. Sin embargo, un pseudo R^2 mucho menor indica que estas variables no explican tan bien como la dicotómica la varianza de la variable dependiente. Para poder apreciar bien su incidencia, la tabla 3.6 enuncia los efectos marginales promedio de estas tres variables en sus modelos respectivos.

Se puede apreciar cómo los signos y la significancia de los efectos marginales también cambia; ahora son positivos. A pesar de ello, como se había predicho, el módulo sigue siendo cuantiosamente menor que el de la variable dicotómica. Estos hechos muestran que efectivamente hay colinealidad entre estas variables y a su vez pareciera no haber continuidad a la hora de explicar la dummy de fusiones entre empresas, sino que basta con una coincidencia

	(1)	(2)	(3)
	Probit sumgene	Probit meangene	Probit mediangene
MA's			
sumgene	0,0128*** (0,00120)		
meangene		0,0156*** (0,00182)	
mediangene			0,0135*** (0,00195)
N	2.238.346	2.238.346	2.238.346
pseudo R^2	0,016	0,006	0,003

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.5: Resultados Variables Genéticas Modelo Probit con Variación Temporal

	(1)	(2)	(3)
	Margin sumgene	Margin meangene	Margin mediangene
sumgene	0,00000108*** (0,000000194)		
meangene		0,00000133*** (0,000000259)	
mediangene			0,00000115*** (0,000000244)
N	2.238.346	2.238.346	2.238.346

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.6: Efectos Marginales Promedio Variables Genéticas Modelo Probit con Variación Temporal

en los genes que atacan para que su probabilidad de fusión sea mucho mayor.

Lo que resta ahora es evaluar la robustez del modelo, y para ello se crearon los siguientes dos modelos que se explican en las secciones 3.3 y 3.4.

3.3. Test de Robustez: Probit Dinámico

El objetivo de esta subsección es netamente comprobar la robustez del modelo, en particular de la variables dicotómica de target genético.

3.3.1. Análisis por tamaño de firmas

En este caso, se observan solo las combinaciones entre firmas grandes y pequeñas. La tabla 3.7 muestra claramente cómo la única variable significativa para este modelo es la que representa el target genético. En el caso anterior, el target por enfermedad era significativo al 0,1, sin embargo, ahora es estadísticamente cero. TargetGen, por otra parte, es significativa con un nivel de confianza del 99%.

	(1)	(2)	(3)	(4)
	Probit all 3	Probit all+sum	Probit all+mean	Probit all+median
MAs				
1.target	0,961*** (0,154)	0,955*** (0,175)	0,959*** (0,163)	0,978*** (0,156)
1.indication	0,159 (0,197)	0,157 (0,191)	0,159 (0,197)	0,156 (0,198)
1.therapeutic	0,153 (0,143)	0,153 (0,142)	0,153 (0,142)	0,151 (0,144)
sumgene		0,000550 (0,00453)		
meangene			0,000710 (0,0126)	
mediangene				-0,00856 (0,0147)
<i>N</i>	132620	132620	132620	132620
pseudo R^2	0,116	0,116	0,116	0,116

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.7: Resultados Modelo Probit con combinaciones de firmas grandes y pequeñas

3.3.2. Análisis dinámico

Aquí se presentan dos modelos similares. Uno tiene variables rezagadas un año, y el otro variables rezagadas tres años. Se puede apreciar en las tablas 3.8 y 3.9 que nuevamente la única variable estadísticamente significativa (con un nivel de significancia de 0,01 para ambos casos) es TargetGen. Este resultado es importante, ya que eliminando el posible escenario endógeno del modelo, la variable en estudio continúa siendo altamente significativa.

Como es de esperar el pseudo R^2 es menor para el caso con tres rezagos en todos los modelos, pero la significancia de la variable se mantiene, lo que da constancia de su robustez.

Es clara la significancia de la variable en modelos que solucionan los posibles escenarios endógenos, por lo que ahora es importante determinar una medida continua con el fin de

	(1)	(2)	(3)	(4)
	Probit all 3	Probit sumgene	Probit meangene	Probit mediangene
MAs				
1.Ltarget	1,013*** (0,117)	1,012*** (0,132)	1,040*** (0,122)	1,042*** (0,120)
1.Lindication	0,226 (0,173)	0,225 (0,172)	0,226 (0,173)	0,223 (0,173)
1.Ltherapeutic	0,0917 (0,155)	0,0917 (0,155)	0,0913 (0,156)	0,0921 (0,156)
Lsumgene		0,000132 (0,00259)		
Lmeangene			-0,00458 (0,00404)	
Lmediangene				-0,00615 (0,00443)
<i>N</i>	2.106.505	2.106.505	2.106.505	2.106.505
pseudo <i>R</i> ²	0,098	0,098	0,099	0,099

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.8: Resultados Modelo Probit Variables Rezagadas 1 año

	(1)	(2)	(3)	(4)
	Probit all 3	Probit sumgene	Probit meangene	Probit mediangene
MAs				
1.L3target	1,005*** (0,121)	1,013*** (0,136)	1,022*** (0,126)	1,034*** (0,122)
1.L3indication	0,0514 (0,171)	0,0541 (0,168)	0,0522 (0,171)	0,0484 (0,171)
1.L3therapeutic	0,0288 (0,134)	0,0284 (0,134)	0,0281 (0,134)	0,0283 (0,134)
L3sumgene		-0,000674 (0,00366)		
L3meangene			-0,00251 (0,00432)	
L3mediangene				-0,00549 (0,00537)
<i>N</i>	1.844.873	1.844.873	1.844.873	1.844.873
pseudo <i>R</i> ²	0,049	0,049	0,049	0,049

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.9: Resultados Modelo Probit Variables Rezagadas 3 años

establecer si existe alguna tendencia en las fusiones y adquisiciones que se explique mejor que con una variable binaria. Los resultados de este modelo, con la inclusión de esta nueva variable, se muestran en la sección 3.4.

3.4. Modelo Probit con variación temporal, y métrica de Distancia de Wallis

Hasta este momento, se ha comprobado que existe una variable binaria que es capaz de explicar con alta significancia estadística la variable dependiente de fusiones o adquisiciones, ya sea con un panel fijo o con variación en el tiempo. Se ha visto también que esta variable es robusta para el modelo, manteniendo su significancia en tres especificaciones generales.

A su vez, tres variables “auxiliares” entregan cierta lucidez respecto a la relación continua que pueda tener el target genético con la dummy de fusiones o adquisiciones, pero no poseen ningún sustento que compruebe su veracidad. Lo que resta ahora, que es el propósito de este modelo, es comprobar si esta relación continua explica mejor la variable dependiente que una relación dicotómica, mediante una variable obtenida a partir de la teoría de redes. Para ello se creó la Distancia de Wallis, cuyos resultados pueden observarse en la tabla 3.10.

	(1) Probit Wallis	(2) Probit noTarget+Wallis	(3) Probit all+Wallis
MAs			
wallis	-1,376*** (0,0906)	-1,142*** (0,0979)	-0,213 (0,177)
1.target			0,950*** (0,159)
1.indication		0,485*** (0,178)	0,351** (0,176)
1.therapeutic		0,0888 (0,165)	0,0111 (0,164)
<i>N</i>	2.238.346	2.238.346	2.238.346
pseudo <i>R</i> ²	0,076	0,101	0,129

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.10: Resultados Modelo Probit Variación Temporal con Distancia de Wallis

Es claro ver que la Distancia de Wallis es negativa producto de lo que se ha mencionado anteriormente, cuanto más distante estén las firmas menor es su probabilidad de fusión. A su vez se aprecia que esta es significativa con un p-valor menor a 0,01 cuando explica por si sola la variable dependiente. Inclusive, si se observa el segundo modelo, se puede comprobar que al no incluir la dummy de target genético la Distancia de Wallis mantiene su significancia con un 99 % de confianza. Sin embargo, una vez que esta es agregada al modelo de probabilidad no lineal junto con la variable dicotómica TargetGen, esta pierde su significancia para el mismo, como se había predicho, lo que da cuenta de la colinealidad existente entre ellas. Este hecho era de esperarse, debido a que ocurre algo análogo a lo que sucedía con meangene y mediangene: toma valores en el intervalo $[0, 1[$ cuando la dummy es 1.

Ahora bien, con respecto a cómo varía la probabilidad de fusión o adquisición, para esta

variable es necesario observar el efecto marginal promedio, ya que debido a su continuidad hay muchos valores de la misma. En la tabla 3.11 puede verse la magnitud de estos efectos.

	(1)	(2)	(3)
	Margins Wallis	Margins noTarget+Wallis	Margins all+Wallis
wallis	-0,000111*** (0,0000180)	-0,0000904*** (0,0000162)	-0,0000165 (0,0000158)
<i>N</i>	2.238.346	2.238.346	2.238.346

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.11: Efecto Marginal Promedio de la Distancia de Wallis

La tabla 3.11 muestra que, si la Distancia de Wallis se encuentra sola en el modelo entonces la probabilidad de fusión disminuye, en promedio, en un 0,01 % con cada unidad en la que incremente la distancia entre las firmas. Este valor es significativo con una confianza del 99 %, y a su vez es aproximadamente siete veces menor (en módulo) que el efecto marginal que tiene TargetGen cuando explica por si sola la dummy de fusión (ver tabla 3.4). En caso de agregar todas las otras variables, exceptuando TargetGen, mantiene su significancia pero disminuye levemente su módulo, lo que es de esperarse debido al efecto positivo que tienen las otras variables para con la dummy de fusión. Por otra parte, si Wallis intenta explicar la variable dependiente junto con la variable dicotómica de terapia genética, su efecto marginal pierde su significancia, mientras que, como se observa en la tabla 3.12, TargetGen es significativa al 0,1.

	(1)
	Margins all 3
0.target	0,0000118*** (0,00000233)
1.target	0,000518* (0,000276)
0.indication	0,0000152*** (0,00000303)
1.indication	0,0000626 (0,0000386)
0.therapeutic	0,0000194*** (0,00000514)
1.therapeutic	0,0000202** (0,00000925)
<i>N</i>	2238346

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.12: Efectos Marginales Variables dicotómicas en modelo con Distancia de Wallis

No quedan dudas de que es mejor explicar la dummy de fusiones o adquisiciones con una

variable dicotómica que represente el target genético, ya que las continuas son claramente colineales con esta y pierden su significancia cuando intentan explicar la variable dependiente junto a ella.

3.5. Test de Robustez e interacción con la evolución del proyecto

3.5.1. Análisis por tamaño de firmas: Probabilidad de adquisición

Para complementar, se realiza el test de robustez de acuerdo al tamaño de las firmas, pero ahora incluyendo la distancia de Wallis. Los resultados se presentan en la tabla 3.13.

	(1)	(2)
	Probit all 3	Probit noTarget+Wallis
MA's		
1.target	0,881*** (0,173)	
1.indication	0,182 (0,214)	0,418** (0,205)
1.therapeutic	0,106 (0,157)	0,105 (0,166)
wallis		1,613*** (0,224)
<i>N</i>	132.620	132.620
pseudo R^2	0,088	0,086

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.13: Resultados Modelo Probit Distancia de Wallis con combinaciones de firmas grandes y pequeñas

Como se puede apreciar, la distancia de Wallis mantiene su significancia al considerar solo adquisiciones de firmas pequeñas por parte de las firmas grandes.

3.5.2. Interacción con la evolución del proyecto

Esta sección presenta tres modelos que incluyen la etapa del proyecto de la firma pequeña, para cada combinación de firma grande - pequeña. El primero agrega las etapas del proyecto como efecto fijo, mientras que los otros dos incluyen la interacción entre la etapa del proyecto y el target genético y la distancia de Wallis, respectivamente.

La tabla 3.14 muestra los resultados del primer modelo probit.

	(1)	(2)	(3)	(4)
	Probit Target	Probit Indication	Probit Therapeutic	Probit Wallis
MAs				
1.target	0,912*** (0,211)			
1.stage	-0,0160 (0,336)	-0,00793 (0,314)	0,0173 (0,309)	-0,00645 (0,332)
2.stage	0,0362 (0,230)	0,149 (0,229)	0,145 (0,222)	0,0253 (0,227)
3.stage	0,331 (0,362)	0,446 (0,343)	0,473 (0,346)	0,299 (0,369)
1.indication		0,306 (0,196)		
1.therapeutic			-0,260 (0,311)	
wallis				1,423*** (0,201)
<i>N</i>	19.596	19.596	19.596	19.596
pseudo <i>R</i> ²	0,136	0,028	0,016	0,116
	(5)	(6)		
	Probit all 3	Probit NoTarget+Wallis		
MAs				
1.target	0,977*** (0,210)			
1.stage	-0,107 (0,318)	-0,0640 (0,329)		
2.stage	-0,00834 (0,206)	0,00291 (0,209)		
3.stage	0,253 (0,349)	0,209 (0,358)		
1.indication	0,170 (0,222)	0,423* (0,237)		
1.therapeutic	0,304 (0,310)	0,172 (0,425)		
wallis		1,651*** (0,335)		
<i>N</i>	19.596	19.596		
pseudo <i>R</i> ²	0,150	0,142		

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.14: Resultados Modelo Probit incluyendo efecto fijo por etapa de proyecto

Como se puede apreciar, las etapas por si solas no son relevantes para el modelo, y las variables relacionadas al target genético (dicotómica y distancia de Wallis) se mantienen

altamente significativas a pesar de los efectos fijos presentes. Las etapas 0, Discovery, se omitió en la tabla debido a que su valor era cero por colinealidad, mientras que la 4, Market, entregaba valor cero debido a que no representaba ningún caso de éxito para el modelo (nunca se adquirió una firma cuyo proyecto ya estaba en el mercado).

Debido a que ya se conoce que los genes con los que trabajan las firmas son el hecho más relevante para explicar las fusiones y adquisiciones, sería interesante observar si es que la interacción de ellos con la evolución del proyecto es relevante. A continuación se presentan los resultados para la interacción entre la evolución del proyecto y la variable dicotómica que representa al target genético.

(1)	
Probit Stage Target	
MAs	
1.target#0.stage	0,439 (0,353)
1.target#1.stage	0,847** (0,385)
1.target#2.stage	0,846*** (0,270)
1.target#3.stage	1,144*** (0,414)
<i>N</i>	10.858
pseudo R^2	0,110
* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$	

Tabla 3.15: Resultados Modelo Probit interacción evolución del proyecto y TargetGen

La tabla 3.15 muestra como la variable dicotómica TargetGen es altamente significativa para explicar la dummy de adquisiciones cuando el proyecto de la empresa pequeña se encuentra en las etapas 1, 2 o 3. La etapa 4 se omitió porque, como se señaló previamente, en ningún caso en que el proyecto haya llegado al mercado, la firma había sido adquirida por otra. Si el proyecto se encuentra en la etapa 0, o Discovery, no es relevante para la firma grande ese proyecto. Esto es normal ya que una empieza a ser más atractiva a medida que avanza en las etapas para llegar al mercado. Estando en Discovery, todavía no traspasó ninguna barrera de la FDA, por lo que aún no tiene valor para los potenciales compradores.

Ahora bien, restaría observar este efecto de interacción pero con la variable continua, la distancia de Wallis. La tabla 3.16 muestra estos resultados.

(1)	
Probit Stage Wallis	
MAs	
0.stage#c.wallis	0,732 (0,494)
1.stage#c.wallis	1,222*** (0,468)
2.stage#c.wallis	1,647*** (0,322)
3.stage#c.wallis	1,815*** (0,544)
N	19.690
pseudo R^2	0,125

* $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$

Tabla 3.16: Resultados Modelo Probit interacción evolución del proyecto y distancia de Wallis

Se observa que la interacción de las etapas con la distancia de Wallis es análoga a lo que ocurre con la variable dicotómica TargetGen.

Nuevamente se puede concluir que las dos variables, dicotómica y continua, explican con gran similitud y significancia la dummy de adquisiciones; sin embargo, por todo lo enunciado anteriormente en este informe, es la primera la que posee más robustez para el modelo.

Conclusión

Se estimaron varios modelos probit con distintas especificaciones y submuestras. El objetivo principal era determinar la relevancia de los targets genéticos de las firmas con respecto a la probabilidad de fusión o adquisición entre ellas.

El modelo estático mostró un alto nivel de significancia de la variable dummy TargetGen, que tomaba el valor uno en caso de que las firmas compartieran al menos un target genético. Posteriormente al segundo modelo, se le agregó una especificación dinámica que involucraba variación en el tiempo, la cual no afectó la alta significancia de la misma. Se realizaron, por tanto, tests de robustez con el fin de estudiar esta variable en más detalle. El primero de estos, que involucraba solo las combinaciones entre firmas grandes y pequeñas, no afectó que TargetGen sea significativa para el modelo con un 99 % de confianza. El segundo tenía por objetivo descartar cualquier problema de endogeneidad presente entre la variable dependiente y la variable explicativa en estudio. Con rezagos de uno y tres años, se comprobó que el target genético era muy robusto para el modelo, y eliminando cualquier sospecha de endogeneidad igualmente mantenía su significancia. Restaba, por tanto, evaluar si la relación entre la dummy de fusión o adquisición de firmas mantenía una relación continua con la presencia de targets genéticos en común, o si en realidad se explicaba mejor con la dummy TargetGen. Para ello se introduce la variable denominada Distancia de Wallis, la cual toma valores entre 0 y 1 dependiendo de la cantidad de genes en común y la cantidad de genes en total con los que trabajan las firmas. El resultado mostró colinealidad entre TargetGen y la Distancia de Wallis, siendo la segunda menos representativa para el modelo (pierde su significancia cuando se coloca junto a la dummy) y, al trabajar con estas por separado la primera mostró un pseudo R^2 mayor.

A esto habría que agregarle que, si el remedio de una firma pequeña supera la etapa de Discovery, entonces en caso de tener al menos un gen en común con la firma grande, su probabilidad de ser adquirida aumenta en forma considerable y significativamente.

Se puede concluir entonces que la decisión de fusión o adquisición de firmas en la industria farmacéutica puede ser explicada por la presencia de targets genéticos en común entre las firmas. En caso de que exista al menos un gen en común entre ellas, entonces la probabilidad de que se fusionen es hasta sesenta y cinco veces mayor.

Así como Marco and Rauser (2008) y Danzon et al. (2007) mostraron la importancia de las patentes para con las fusiones y adquisiciones, este trabajo muestra que en esta nueva era genética es fundamental tener en cuenta los pipelines de las firmas un par de “escalones” más arriba. No es tan determinante la enfermedad que ataquen o el área terapéutica donde

se desempeñen, sino que lo más sustancial a la hora de evaluar la probabilidad de fusiones entre empresas son los target genéticos con los que trabajan.

En esta época de transición cada vez más firmas se están inclinando hacia la terapia genética, como lo describe Potts (2008). Por ende, es de esperarse que en el futuro sea mejor enfocarse en los genes con los que trabaja la empresa y su relación con las enfermedades, que poner el foco únicamente en estas últimas.

A esto se le agrega que las firmas pequeñas son frecuentemente adquiridas por firmas grandes, mediante el desarrollo de terapias de interés para estas últimas. De las 44 F&A estudiadas en este trabajo un 77% fueron adquisiciones en las que las firmas tenían genes en común, por lo tanto, las firmas pequeñas deberían enfocarse en desarrollar terapias con genes con los que ya estén trabajando firmas grandes si su objetivo es aumentar su probabilidad de adquisición. Por otra parte, el interés de las firmas grandes por adquirir a otras que trabajen en nuevos remedios con los mismos targets genéticos que ellas utilizan, podría explicarse por el aprovechamiento de las economías de alcance y el ahorro del costo de las primeras etapas de desarrollo que los nuevos medicamentos implican. De hecho, Cockburn y Henderson (2001), como se ha señalado previamente en este informe, encuentran que las economías de alcance son más importantes que las de escala para el desarrollo de medicamentos, y este trabajo lo que consigue es dilucidar que las economías de alcance conviene aprovecharlas de acuerdo al target genético con el cual trabajan las firmas grandes, más que los targets de enfermedad o de área terapéutica que puedan tener.

Siguiendo con la línea de Henderson y Cockburn (1996) podría inferirse que el aprovechamiento de estas economías de alcance para mejorar la productividad innovativa, aún justifica la pérdida de eficiencia producto del aumento de poder de mercado de las firmas grandes.

Sin embargo, de acuerdo a la data, a pesar de las economías de alcance que pueden aprovecharse solo 11 drogas alcanzaron la etapa de mercado, y todas ellas pertenecen a firmas grandes. El número de observaciones es muy pequeño para poder realizar un test significativo. Lo ideal, sería tener un dataset con gran cantidad de drogas en las que se indique cuáles llegaron al mercado, y cuáles fueron desarrolladas en principio por una firma pequeña y luego adquiridas por una firma grande. De esta forma se podría realizar un modelo de conteo, como una regresión de Poisson, con el cual determinar si el número de proyectos exitosos aumenta de acuerdo a si son comprados o no por firmas grandes. El análisis se deja planteado para futuras investigaciones.

A su vez, este trabajo de tesis deja abiertas preguntas de investigación, como por ejemplo evaluar si las externalidades en I+D presentadas por las fusiones o adquisiciones en esta nueva era se condicen con las que se presentaban antes de que la terapia génica tomase relevancia. A su vez, podría estudiarse si estas nuevas fusiones presentan más beneficios para las firmas como spillovers en I+D o si, por ejemplo, se mantiene en la mejora en la asignación de recursos como encuentra Devos et al. (2008).

Como se puede ver, aún queda mucho por estudiar en este ámbito de terapias genéticas: una etapa que recién comienza pero que tiene gran potencial de desarrollo por delante, y en donde el uso del análisis de redes puede jugar un rol importante dada la rápida velocidad de descubrimientos en la red entre genes y enfermedades.

Bibliografía

- [1] BLOOM, N., SCHANKERMAN, M., AND REENEN, J. V. Identifying Technology Spillovers and Product Market Rivalry. *Econometrica* 81, 4 (2013), 1347–1393.
- [2] BUNKE, H. On a relation between graph edit distance and maximum common subgraph. *Pattern Recognition Letters*, 18 (1997), 689–694.
- [3] BUNKE, H., AND MESSMER, B. T. Recent Advances in Graph Matching. *International Journal of Pattern Recognition and Artificial Intelligence* 11, 01 (1997), 169–203.
- [4] BUNKE, H., AND SHEARER, K. A graph distance metric based on the maximal common subgraph. *Pattern Recognition Letters*, 19 (1998), 255–259.
- [5] COCKBURN, I. M., AND HENDERSON, R. M. Scale and scope in drug development: unpacking the advantages of size in pharmaceutical research. *Journal of Health Economics* 20 (2001), 1033–1057.
- [6] CRANDALL, R. W., AND WINSTON, C. Does Antitrust Policy Improve Consumer Welfare? Assessing the Evidence. *Journal of Economic Perspectives* 17, 4 (2003), 3–26.
- [7] DANZON, P. M., EPSTEIN, A., AND NICHOLSON, S. Mergers and Acquisitions in the Pharmaceutical and Biotech Industries. *Managerial and Decision Economics* 28 (2007), 307–328.
- [8] DEVOS, E., KADAPAKKAM, P., AND KRISHNAMURTHY, S. How Do Mergers Create Value? A Comparison of Taxes, Market Power, and Efficiency Improvements as Explanations for Synergies. *The Review of Financial Studies* 22, 3 (2008), 1179–1211.
- [9] FERNANDEZ, M., AND VALIENTE, G. A graph distance metric combining maximum common subgraph and minimum common supergraph. *Pattern Recognition Letters*, 22 (2001), 753–758.
- [10] GOH, K., CUSICK, M. E., VALLE, D., CHILDS, B., VIDAL, M., AND BARABÁSI, A. The human disease network. *Proceedings of the National Academy of Sciences* 104, 21 (2007), 8685–8690.
- [11] HENDERSON, R. M., AND COCKBURN, I. M. Scale, Scope, and Spillovers: The Determinants of Research Productivity in Drug Discovery. *RAND Journal of Economics* 27, 1 (1996), 32–59.

- [12] MARCO, A. C., AND RAUSSER, G. C. The Role of Patent Rights in Mergers: Consolidation in Plant Biotechnology. *American Journal of Agricultural Economics* 90, 1 (2008), 133–151.
- [13] MEDER, H. *The Incentives for M&As in the Pharmaceutical Industry*. PhD thesis, KU Leuven, Belgium, hendrik.meder@kuleuven.be, 3 2016. Unpublished thesis.
- [14] POTTS, M. Pharmaceutical Mergers and Genetic Technology: A Problematic Combination. In *The Ethics of Genetic Commerce*, R. W. Kolb, Ed. Wiley-Blackwell, 2007, ch. 11, pp. 177–189.
- [15] WALLIS, W. D., SHOUBRIDGE, P., KRAETZ, M., AND RAY, D. Graph distances using graph union. *Pattern Recognition Letters* 22 (2001), 701–704.
- [16] WILLIAMSON, O. Economies as an Antitrust Defense: The Welfare Tradeoffs. *The American Economic Review* 58, 1 (1968), 18–36.

Anexo A

Tabla A.1: Corrección de nombres de las firmas para realizar el matching de las Bases de Datos

Thomson Reuters Cortellis Life Sciences	Deloitte RECAP IQ	matchingname
ALS Therapy Development Institute	ALS Distribuidora Limitada	ALS
ANI Pharmaceuticals Inc	ANI Pharmaceuticals	ANI Pharmaceuticals
Abbott Laboratories	Abbott	Abbott
Abbott Laboratories	Abbott Medical Optics	Abbott
ActivX Biosciences Inc	ActivX Biosciences	ActivX Biosciences
Acuity Pharmaceuticals	Acuity Pharmaceuticals, Froptix	Acuity Pharmaceuticals
Aegera Therapeutics Inc	Aegera Therapeutics	Aegera Therapeutics
Allergan Inc	Allergan	Allergan
Alnylam Pharmaceuticals Inc	Alnylam Pharmaceuticals	Alnylam Pharmaceuticals
Alseres Pharmaceuticals Inc	Alseres Pharmaceuticals	Alseres Pharmaceuticals
Amsterdam Molecular Therapeutics BV	Amsterdam Molecular Therapeutics	Amsterdam Molecular Therapeutics
AngioGenex Inc	AngioGenex	AngioGenex
Antisoma plc	Antisoma	Antisoma
Aptamera Inc	Aptamera	Aptamera
ArQule Inc	ArQule	ArQule
AstraZeneca plc	AstraZeneca	AstraZeneca
Avalon Pharmaceuticals Inc	Avalon Pharmaceuticals	Avalon Pharmaceuticals
Aventis Pharma AG	Aventis	Aventis

Thomson Reuters Cortellis Life Sciences	Deloitte RECAP IQ	matchingname
Aventis SA	Aventis Behring	Aventis
Avigen Inc	Aventis, Schering AG	Avigen
Avigen Inc	Avigen	Avigen
Bayer AG	Bayer	Bayer
Bayer Corp	Bayer CropScience	Bayer
Bayer Schering Pharma AG	Bayer Diagnostics	Bayer
Benitec Biopharma Ltd	Benitec	Benitec
Bio-Path Holdings Inc	Bio-Path Holdings	Bio-Path Holdings
BioDelivery Sciences International Inc	BioDelivery Sciences	BioDelivery Sciences
BioMarin Pharmaceutical Inc	BioMarin Pharmaceuticals	BioMarin Pharmaceuticals
BioVex Inc	BioVex	BioVex
Biocon Ltd	Biocon Ltd.	Biocon Ltd
Biogen Inc	Biogen	Biogen
Biogen Idec Inc	Biogen Idec	Biogen
Boehringer Ingelheim Corp	Boehringer Ingelheim	Boehringer Ingelheim
Boehringer Ingelheim Corp	Boehringer Mannheim	Boehringer Ingelheim
Boehringer Ingelheim Corp	Boehringer Mannheim, Corange	Boehringer Ingelheim
Boston Biomedical Research Inst	Boston Analytical	Boston Biomedical
Boston University	Boston Biochem	Boston Biomedical
Boston Biomedical Inc	Boston Biomedical	Boston Biomedical
Boston Biomedical Inc	Boston Scientific	Boston Biomedical
Bristol-Myers Squibb Co	Bristol-Myers Squibb	Bristol-Myers Squibb
Bristol-Myers Squibb Co	Bristol-Myers Squibb Medical Imaging	Bristol-Myers Squibb
Calando Pharmaceuticals Inc	Calando Pharmaceuticals	Calando Pharmaceuticals
Canji Inc	Canji	Canji
Celgene Corp	Celgene	Celgene
Cell Genesys Inc	Cell Genesys	Cell Genesys
Cengent Therapeutics Inc	Cengent Therapeutics	Cengent Therapeutics
Cephalon Inc	Cephalon	Cephalon

Thomson Reuters Cortellis Life Sciences	Deloitte RECAP IQ	matchingname
ChemGenex Pharmaceuticals Ltd	ChemGenex Pharmaceuticals ChemGenex Therapeutics, ChemGenex Pharmaceuticals, Cephalon	ChemGenex Pharmaceuticals ChemGenex Pharmaceuticals
ChemGenex Pharmaceuticals Ltd	Chiesi Farmaceutici SpA	Chiesi Farmaceutici
Chiron Corp Chiron Viagene Inc	Chiron	Chiron Chiron
Chiroscience R&D Inc Chiroscience R&D Inc	Chiroscience Chiroscience, Celltech	Chiroscience Chiroscience
Chugai Pharmaceutical Co Ltd	Chugai	Chugai
Clinical Data Inc Colby Pharmaceutical Co	Clinical Data Colby Pharmaceutical	Clinical Data Colby Pharmaceutical
Collateral Therapeutics Inc	Collateral Therapeutics	Collateral Therapeutics
Cordis Corp	Cordis	Cordis
Crucell NV	Crucell	Crucell
Cyclacel Pharmaceuticals Inc	Cyclacel Pharmaceuticals	Cyclacel Pharmaceuticals
Cyclis Pharmaceuticals Inc	Cyclis Pharmaceuticals	Cyclis Pharmaceuticals
CytRx Corp CytRx Laboratories Inc	CytRx	CytRx CytRx
Daiichi Sankyo Co Ltd	Daiichi Pharmaceutical	Daiichi Sankyo
Daiichi Seiyaku Co Ltd	Daiichi Pharmaceutical Korea	Daiichi Sankyo
Daiichi Seiyaku Co Ltd	Daiichi Sankyo	Daiichi Sankyo
Diamyd Inc Discovery Laboratories Inc	Diamyd Medical Discovery Laboratories	Diamyd Inc Discovery Laboratories
Edwards Lifesciences Corp	Edwards Lifesciences	Edwards Lifesciences
Eisai Co Ltd	Eisai	Eisai
Enzon Pharmaceuticals Inc	Enzon Pharmaceuticals	Enzon Pharmaceuticals

Thomson Reuters Cortellis Life Sciences	Deloitte RECAP IQ	matchingname
Epitomics Inc	Epitomics	Epitomics
Ercole Biotech Inc	Ercole Biotech	Ercole Biotech
Excaliard Pharmaceuticals Inc	Excaliard	Excaliard
Fate Therapeutics Inc	Fate Therapeutics	Fate Therapeutics
Fujisawa Pharmaceutical Co Ltd	Fujisawa	Fujisawa
Galena Biopharma Inc	Galena Biopharma	Galena Biopharma
GenVec Inc	GenVec	GenVec
GeneMedicine Inc	GeneMedicine	GeneMedicine
Genentech Inc	Genentech	Genentech
Genetic Therapy Inc	Genetic Therapy	Genetic Therapy
Genta Inc	Genta	Genta
Genzyme Corp	Genzyme	Genzyme
Genzyme Corp	Genzyme Biosurgery, Genzyme	Genzyme
Genzyme Corp	Genzyme Diagnostics	Genzyme
Genzyme Corp	Genzyme General, Genzyme	Genzyme
Genzyme Corp	Genzyme Genetics	Genzyme
Geron Corp	Geron	Geron
Gilead Palo Alto Inc	Gilead	Gilead
GlaxoSmithKline plc	GlaxoSmithKline	GlaxoSmithKline
Green Cross Corp	Green Cross	Green Cross
H3 Biomedicine Inc	H3 Pharma	H3 Biomedicine Inc
Harbor Therapeutics Inc	Harbor BioSciences	Harbor BioSciences
Harvard Medical School	Harvard Bioscience	Harvard Medical School
Helix Therapeutics Inc	Helix BioPharma	Helix BioPharma
Hoechst Marion Roussel Inc	Hoechst Marion Roussel	Hoechst Marion Roussel
Hospira Inc	Hospira	Hospira
Human Genome Sciences Inc	Human Genome Sciences	Human Genome Sciences
IDEC	IDEC	IDEC
Pharmaceuticals Corp	Pharmaceuticals	Pharmaceuticals
Inovio	Inovio	Inovio
Pharmaceuticals Inc	Pharmaceuticals	Pharmaceuticals
Intradigm Corp	Intradigm	Intradigm
Intrexon Corp	Intrexon	Intrexon

Thomson Reuters Cortellis Life Sciences	Deloitte RECAP IQ	matchingname
Isis Pharmaceuticals Inc	Isis Pharma GmbH, Schwarz Pharma	Isis Pharmaceuticals Inc
Johnson & Johnson Pharmaceutical Research & Development LLC	Johnson & Johnson	Johnson & Johnson
Johnson & Johnson Pharmaceutical Research & Development LLC	Johnson & Johnson CBS	Johnson & Johnson
Kyowa Hakko Kirin Co Ltd	Kyowa Hakko Kirin	Kyowa Hakko Kirin
Kyowa Hakko Kogyo Co Ltd	Kyowa Hakko Kogyo	Kyowa Hakko Kogyo
Lexicon Pharmaceuticals Inc	Lexicon Pharmaceuticals	Lexicon Pharmaceuticals
Lorus Therapeutics Inc	Lorus Therapeutics	Lorus Therapeutics
MGI Pharma Inc	MGI Pharma	MGI Pharma
Marina Biotech Inc	Marina Biotech	Marina Biotech
Mayne Pharma Pty Ltd	Mayne Pharma	Mayne Pharma
Mayne Pharma Pty Ltd	Mayne Pharma International	Mayne Pharma
Medistem Inc	Medistem	Medistem
Medtronic Inc	Medtronic	Medtronic
Merck & Co Inc	Merck	Merck
Merck & Co Inc	Merck KGaA	Merck
Merck & Co Inc	Merck-Medco	Merck
Mitsubishi Tanabe Pharma Corp	Mitsubishi Pharma	Mitsubishi Pharma
Morphochem Inc	Morphochem AG	Morphochem AG
Morphochem Inc	Morphochem AG	Morphochem AG
Novartis AG	Novartis	Novartis
Novartis Vaccines & Diagnostics	Novartis Generics	Novartis
Nuvelo Inc	Nuvelo	Nuvelo
Nuvelo Inc	Nuvelo, Affymetrix	Nuvelo
OXIS International Inc	OXIS International	OXIS International
OncoGenex Pharmaceuticals Inc	OncoGenex Pharmaceuticals	OncoGenex Pharmaceuticals
OncoGenex Technologies Inc	OncoGenex Technologies	OncoGenex Technologies

Thomson Reuters Cortellis Life Sciences	Deloitte RECAP IQ	matchingname
Onyx Pharmaceuticals Inc	Onyx Pharmaceuticals	Onyx Pharmaceuticals
Opexa Therapeutics Inc	Opexa Therapeutics	Opexa Therapeutics
OriGenix Technologies Inc	OriGene Technologies	OriGene Technologies
Oxford BioMedica plc	Oxford BioMedica	Oxford BioMedica
Pathfinder Management Inc	Pathfinder LLC	Pathfinder LLC
Pfizer Inc	Pfizer	Pfizer
Pfizer Inc	Pfizer Animal Health	Pfizer
Piramed Pharma	Piramed	Piramed
Plexxikon Inc	Plexxikon	Plexxikon
Procter & Gamble Pharmaceuticals Inc	Procter & Gamble	Procter & Gamble
Protiva Biotherapeutics Inc	Protiva	Protiva
Questcor Pharmaceuticals Inc	Questcor Pharmaceuticals	Questcor Pharmaceuticals
Radiant Pharmaceuticals Corp	Radiant Pharmaceuticals	Radiant Pharmaceuticals
Repligen Corp	Repligen	Repligen
RheoGene Inc	RheoGene	RheoGene
Rhone-Poulenc Rorer Inc	Rhone-Poulenc Rorer	Rhone-Poulenc Rorer
Roche Holding AG	Roche	Roche
Roche Madison Inc	Roche Biomedical Laboratories	Roche
Roche Madison Inc	Roche Diagnostics	Roche
SRI International	SRI	SRI
SangStat Medical Corp	SangStat	SangStat
Sangamo BioSciences Inc	Sangamo BioSciences	Sangamo BioSciences
Sanofi Pasteur	Sanofi Pasteur	Sanofi
Sanofi-Aventis US LLC	Sanofi-Aventis	Sanofi
Sanofi	Sanofi-Synthelabo	Sanofi
Schering-Plough Corp	Schering-Plough	Schering-Plough
Scripps Research Institute	Scripps Research Institute	Scripps Research Institute
Shenzhen SiBiono Gene Technologies Co Ltd	Shenzhen SiBiono GeneTech	Shenzhen SiBiono GeneTech

Thomson Reuters Cortellis Life Sciences	Deloitte RECAP IQ	matchingname
Shire plc	Shire	Shire
Shire plc	Shire	Shire
Shire plc	Pharmaceuticals	Shire
Shire plc	Shire Regenerative Medicine	Shire
Silence Therapeutics AG	Silence Therapeutics	Silence Therapeutics
Sirna Therapeutics Inc	Sirna Therapeutics	Sirna Therapeutics
Sirtris	Sirtris	Sirtris
Pharmaceuticals Inc	Pharmaceuticals	Pharmaceuticals
Skinetics Biosciences Inc	Skinetics Biosciences	Skinetics Biosciences
Solexa Inc	Solexa	Solexa
Somatix Therapy Corp	Somatix Therapy	Somatix Therapy
Stanford University	Somatix Therapy, Cell Genesys	Stanford University
StemCells Inc	Stanford Surgical Technology	StemCells
StemCells Inc	StemCells	StemCells
Symphony GenIsis	Symphony Technology Group	Symphony GenIsis
Synvista	Synvista	Synvista
Therapeutics Inc	Therapeutics	Therapeutics
TG Therapeutics Inc	TG Therapeutics	TG Therapeutics
TargeGen Inc	TargeGen	TargeGen
Tekmira	Tekmira	Tekmira
Pharmaceuticals Corp	Pharmaceuticals	Pharmaceuticals
Teva Pharmaceutical Industries Ltd	Teva Pharmaceutical	Teva Pharmaceutical
Tissue Repair Co	Tissue Repair	Tissue Repair
Toyama Chemical Co Ltd	Toyama Chemical	Toyama Chemical
Unitech	Unitech Pharma	Unitech Pharma
Pharmaceuticals Inc	Unitech Pharma	Unitech Pharma
University of Alabama at Birmingham	University of Alabama	University of Alabama
VGX	VGX	VGX
Pharmaceuticals LLC	Pharmaceuticals	Pharmaceuticals
VIRxSYS Corp	VIRxSys	VIRxSys
Valentis Inc	Valentis	Valentis
Vernalis plc	Vernalis	Vernalis

Thomson Reuters Cortellis Life Sciences	Deloitte RECAP IQ	matchingname
Vernalis plc	Vernalis Group	Vernalis
Warner-Lambert Co	Warner-Lambert	Warner-Lambert
Xenon Pharmaceuticals Inc	Xenon	Xenon
Yamanouchi Pharmaceutical Co Ltd	Yamanouchi	Yamanouchi
Zeneca Group plc	Zeneca	Zeneca
Zeneca Yakuhin KK		Zeneca
deCODE genetics ehf	deCODE Genetics	deCODE Genetics
deCODE genetics ehf	deCODE genetics ehf	deCODE Genetics
iCo Therapeutics Inc	iCo Therapeutics	iCo Therapeutics
uniQure BV	uniQure	uniQure