

UNIVERSIDAD NACIONAL DEL COMAHUE

Maestría en Estadística Aplicada



**DISTRIBUCIÓN DE LA MORTALIDAD POR CÁNCER DE
MAMA EN LA PROVINCIA DE NEUQUÉN, ARGENTINA**

Autor

Caro, Patricia Janet

Profesora en Matemática

Neuquén, 2019

COMISIÓN ASESORA DE TESIS

Director

Mg. Braicovich, Teresa

Co-director

Dra. Bruno, Cecilia

DEDICATORIA

A Mis hijos,

Gastón y Valentín

A mi compañero de vida,

Patricio

Que hacen que todos los esfuerzos valgan la pena.

Los amo profundamente con todo mi corazón.

AGRADECIMIENTOS

A mi directora, profesora, compañera y amiga, Teresa Braicovich por creer en mí y acompañarme en todo este proceso, pero más aún por enseñarme el apasionante camino de la docencia y de la investigación.

A mi co-directora, Cecilia Bruno, que me acompañó en este proceso y compartiendo su valiosa experiencia conmigo.

Al Profesor Sergio Bramardi, como director de la maestría me motivo e incentivó a iniciarme en el maravilloso mundo de la investigación.

A la Dra. Adriana Novoa y al Dr. Santiago Hasdeu por creer en mí y brindarme todo su apoyo para contribuir mi granito de arena en investigaciones de Salud Pública.

A mi amiga y compañera Laura Lamfre con estar siempre presente y por alentarme a terminar este trabajo.

A mis compañeros de proyecto de investigación, Raquel, Lorena, Yobran, Marcia, Fernando, Jaime, Micaela, por acompañarme en mis locuras y participando conmigo en jornadas de trabajos y congresos.

A mis padres, Blanca y Daniel, por enseñarme la importancia del esfuerzo, de la perseverancia y que todo es posible.

A mis compañeros de trabajo por alivianarme el camino.

A mis hermanas Beatriz y Graciela, a mi sobrina Romina, que recorrieron conmigo este proceso y siempre estuvieron presentes.

A las Universidades Nacionales de Río Negro y del Comahue, y en especial, a los departamentos de Matemática y Estadística de la facultad de Economía y Administración, por brindarme el ámbito para desarrollarme profesional y humanamente.

RESUMEN

El cáncer de mama es una enfermedad gradual altamente relacionada con condiciones ambientales, factores biológicos (por ejemplo la carga genética, el sexo, la edad, factores sociales (la ocupación, el nivel de instrucción) y hábitos (entre otros, tabaco, alcohol) (Riboli 2002). El patrón de la incidencia del cáncer de mama, por lo tanto, está influenciado por la región geográfica, la situación cultural y las características de la población (Stewart y Kleihues 2003). Esta enfermedad muestra variaciones espaciales y conocerlas resulta útil para establecer la ocurrencia de patrones de la variación al azar, ya que el conocimiento del patrón de la ocurrencia del cáncer y de su distribución espacial es muy importante para identificar a grupos más vulnerables de la población, lo que permitiría definir políticas sanitarias más convenientes para la prevención, la supervisión y el control (Viniegra, 2010). Esto no es ajeno a la situación de la provincia del Neuquén y es por ello que con esta investigación se propone contribuir en este sentido, trabajando con los datos proporcionados por Salud Pública, en particular, en el cálculo de tasas de mortalidad, estimadores del riesgo relativo (REMs), tasas de pacientes con cáncer de mama en salud pública y accesibilidad al mamógrafo público para las 28 áreas programáticas en la que se encuentra dividida sanitariamente la provincia de Neuquén, con el objetivo de identificar patrones espaciales de indicadores del cáncer de mama. El método estadístico propuesto para integrar las variables anteriormente detalladas es un análisis de componentes principales espaciales (ACPe), para la detección de la existencia de la autocorrelación espacial se utilizó el coeficiente de Moran y se realizó un análisis exhaustivo de las redes de contigüidad o estructura espacial, recurriendo a conceptos centrales de la Teoría de Grafos que resultaron de suma utilidad, ya que cierto índices de grafos muestran la topología de las redes utilizadas en este estudio. En función de los resultados obtenidos se identificó que desde el punto de vista de “estructura espacial” la distribución de las variables analizadas en la Provincia de Neuquén, siguen un patrón no aleatorio. Es decir que, de acuerdo al criterio de contigüidad seleccionado, se observaron diferencias entre las distintas áreas programáticas reflejándose una asimetría en cuanto a lo espacial, respecto a las oportunidades de acceder al mamógrafo, registros de mortalidad y detección de nuevos casos de cáncer de mama como también en su vulnerabilidad socioeconómica.

Palabras Claves: cáncer, distribución espacial, autocorrelación espacial, redes de conexión, análisis de componentes principales espacial.

ABSTRACT

Breast cancer is a gradual disease highly related to environmental conditions, biological factors (such as genetics, gender, age, race), social factors (such as occupation, education level) and life style (tobacco and alcohol consumption among others); obviously its causes are complex (Riboli 2002). It is to say, that diverse contextual factors exist and undoubtedly affect the human condition as well as the human health. Hence, breast cancer incidence pattern is influenced by geographical region, cultural situation and the characteristics of the population (Stewart and Kleihues 2003). This disease shows spatial variations, and knowing about them results useful to establish the occurrence of patterns of this variations at random, since knowledge of cancer occurrence pattern and its spatial distribution is extremely necessary to identify the population most vulnerable sectors, what would enable to establish more convenient sanitary policies so as to prevent, supervise and control the mentioned disease (Viniestra,2010). Such matter is not alien to Neuquén province, therefore the present research aims to contribute in this sense, working with the figures provided by the Public Health system, particularly, in mortality rates, Risk Evaluation and Mitigation strategies (REMs), rates of patients with breast cancer in the public health system and access to the public mammography for the 28 programmatic areas in which the province is healthily divided, with the objective of identifying spatial patterns of breast cancer indicators. The statistics method proposed so as to integrate the previously detailed variables is a spatial principal components analysis (PCA), for the detection of the existence of the spatial autocorrelation, the Moran coefficient was used and a thorough research of the contiguity networks or spatial structure was conducted, resorting to central concepts of the Grafos Theory which came in deeply useful, since certain Grafos´rates show the typology of the used networks in the present study. From the obtained results, it was identified that, since the “spatial structure” point of view, the distribution of the analysed variables in Neuquén province follow a not random pattern. This is to say, that according to the contiguity criteria selected, differences between the different programmatic areas were observed, showing asymmetry regarding the spatiality, about the opportunities to have access to a mammography, mortality rates and detection of new breast cancer cases, as well as its socioeconomic vulnerability.

INDICE GENERAL

RESUMEN	4
ABSTRACT	6
CAPÍTULO 1	12
CAPÍTULO 2	22
CAPÍTULO 3	53
CAPITULO 4	64
CAPÍTULO 5	76
CAPÍTULO 6	105
BIBLIOGRAFÍA	107

INDICE DE FIGURAS

FIGURA 1. Estimación de la Incidencia de CM a nivel mundial en 2012.	14
FIGURA 2. Estimación tasas estandarizadas por edad según población mundial por 100.000 mujeres de CM en la Argentina año 2014	16
FIGURA 3. Grafo G no dirigido de 5 vértices y Matriz de Adyacencia asociada.....	31
FIGURA 4. Grafo G no dirigido de 6 vértices y medidas de conexión asociadas.....	34
FIGURA 5. Número Asociado de cada vértice un grafo G de 6 vértices	38
FIGURA 6. Grafo G no dirigido de 6 vértices y Matriz de Accesibilidad Topológica asociada.	39
FIGURA 7. Triangulación de Delaunay en un grafo de 6 nodos.	45
FIGURA 8. Relación entre el Riesgo Relativo y Odds Ratio.	63
FIGURA 9. Indicador socioeconómico según área programática. Provincia de Neuquén. 2016.....	66
FIGURA 10. Redes de Conexión de la Provincia de Neuquén según criterio de contigüidad.....	70
FIGURA 11. Grafos según diferentes Redes de Conexión Provincia de Neuquén.	73
FIGURA 12. Grafos bipartitos de accesibilidad según ubicación geográfica de las áreas programáticas. Provincia de Neuquén.	74
FIGURA 13. Mapeo de los Razón Estandarizada de Mortalidad estandarizados.....	78
FIGURA 14. Grado de Centralidad (Degree) según los Grafos de las estructuras de vecindades. Provincia de Neuquén	85
FIGURA 15. Biplot del grado de centralidad (Degree) según los Grafos de las estructuras de vecindades. Provincia de Neuquén	86
FIGURA 16. Índice de Intermediación según los Grafos de las estructuras de vecindades. Provincia de Neuquén.....	87
FIGURA 17. Biplot de los índices de intermediación según los Grafos de las estructuras de vecindades. Provincia de Neuquén	89
FIGURA 18. Índice de Cercanía según los Grafos de las estructuras de vecindades. Provincia de Neuquén.....	90
FIGURA 19. Biplot de los índices de Cercanía según los Grafos de las estructuras de vecindades. Provincia de Neuquén	91
FIGURA 20. Índices de centralidad por EigenVector según los Grafos de las estructuras de vecindades. Provincia de Neuquén	92

FIGURA 21. Biplot de los índices de centralidad por EigenVector según los Grafos de las estructuras de vecindades. Provincia de Neuquén	93
FIGURA 22. Estilos de conversión según criterios de contigüidad. Provincia de Neuquén	94
FIGURA 23. Tasas de nuevos casos de cáncer de mama en Salud Pública, de las áreas programáticas según coordenadas espaciales. Provincia de Neuquén. Período 2001-2015.	98
FIGURA 24. Razón de Mortalidad Estandarizada de las áreas programáticas según coordenadas espaciales. Provincia de Neuquén. Período 2001-2012.	98
FIGURA 25. Accesibilidad al Mamógrafo de las áreas programáticas según coordenadas espaciales. Provincia de Neuquén.....	98
FIGURA 26. Índice Socio Económico de las áreas programáticas según coordenadas espaciales. Provincia de Neuquén.....	98
FIGURA 27. Biplot de las áreas programáticas generado según ACP. Provincia de Neuquén	100
FIGURA 28. Test de Mantel: Correlación entre las matrices distancia geográfica y distancia Euclidea de los datos. Provincia de Neuquén	101
FIGURA 29. Test de Mantel: Correlación entre las matrices distancia geográfica y CP1_ CP2 (ACP). Provincia de Neuquén	102
FIGURA 30. Biplot de las áreas programáticas generado según ACPe. Provincia de Neuquén	103

INDICE DE TABLAS

TABLA 1. Medidas de Centralidad para Grafo G no dirigido de 6 vértices	37
TABLA 2. Medidas de Accesibilidad para Grafo G no dirigido de 6 vértices	40
TABLA 3. Comparación de tasas según tipo de estandarización.	58
TABLA 4. Tasas de mortalidad específicas por grupo de edad para Argentina y Provincia de Neuquén de Cáncer de Mama (por cada 100.000). Periodo 2007-2011.	68
TABLA 5. Número de vecinos de cada área programa según criterio de contigüidad.	71
TABLA 6. Tasas de mortalidad estandarizadas por edad (TME), desvío estándar (DE(TME)) e intervalos de confianza (IC) para las tasas estandarizadas (por 100.000) de Cáncer de Mama. Provincia de Neuquén. Período 2001-2012.	76
TABLA 7. Casos observados (o_i), casos esperados (e_i), Razón Estandarizada de Mortalidad (REM) e intervalo de confianza (IC) para los Razones Estandarizadas para el Cáncer de Mama. Provincia de Neuquén. Período 2001-2012.	77
TABLA 8. Nuevos Casos de Cáncer de Mama en Salud Pública según edad . Provincia de Neuquén. Período 2001-2015.	78
TABLA 9. Nuevos Casos de Cáncer de Mama en Salud Pública según factor de riesgo. Provincia de Neuquén. Período 2001-2015.	79
TABLA 10. Nuevos Casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria I. Provincia de Neuquén. Período 2001-2015.	80
TABLA 11. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria II. Provincia de Neuquén. Período 2001-2015.	80
TABLA 12. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria III. Provincia de Neuquén. Período 2001-2015.	80
TABLA 13. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria IV. Provincia de Neuquén. Período 2001-2015.	81
TABLA 14. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria V. Provincia de Neuquén. Período 2001-2015.	81
TABLA 15. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Metropolitana. Provincia de Neuquén. Período 2001-2015.	81
TABLA 16. Nuevos casos de Cáncer de Mama en Salud Pública según año. Provincia de Neuquén. Período 2001-2015.	83

TABLA 17. Índices globales según estructura de vecindades. Provincia de Neuquén.....	84
TABLA 18. Pesos Espaciales según estilo de conversión, para el criterio de contigüidad de Tipo Reina (CCTR). Provincia de Neuquén	96
TABLA 19. Estadísticos descriptivos según estilo de conversión, para el criterio de contigüidad de 4 vecinos más cercanos (CC4V). Provincia de Neuquén	96
TABLA 20. Estadísticos descriptivos según estilo de conversión, para el criterio de contigüidad Delaunay Restringido (CCTDR). Provincia de Neuquén	97

CAPÍTULO 1

EPIDEMIOLOGÍA ESPACIAL EN LA DISTRIBUCIÓN DEL CÁNCER DE MAMA

1.1. Introducción a la Estadística Espacial en Epidemiología

La Estadística espacial se ocupa de la exploración, la descripción, la visualización y el análisis de los datos considerando sus características de distribución en el espacio, que suelen expresarse a través del uso de coordenadas geográficas. La característica distintiva del análisis estadístico de datos espaciales es que el patrón espacial de las localizaciones (objetos espaciales), la asociación espacial entre los valores observados en diferentes localizaciones (dependencia espacial) y la variación sistemática del fenómeno en las distintas localizaciones (heterogeneidad espacial) se convierte en el mayor foco de investigación. Estas técnicas no sólo son relevantes en Geografía, sino que son aplicables, además, a un amplio rango de áreas científicas, entre ellas la Epidemiología (Dray *et al.*, 2008, Jombart 2010, Córdoba *et al.*, 2014, Macchiavelli 2013).

Todo puede ser ubicado en una referencia geográfica y los eventos relacionados con la salud, como enfermedad, muertes, nacimientos, exposición a riesgos o cualquier otro suceso, pueden relacionarse con el sitio donde ocurren mediante un par de coordenadas, una dirección o bien un área. El estudio de la distribución geográfica y la asociación espacial de los eventos en salud se denomina epidemiología espacial. La incorporación de esta perspectiva contribuye en el entendimiento de los procesos de salud, enfermedad e incluso puede conducir a diferentes resultados de aquellos obtenidos sin considerar el espacio geográfico (Pina, 2010).

Una característica, asociada específicamente a la Epidemiología espacial, es que los datos son frecuentemente discretos. A diferencia de otras áreas del análisis espacial, que trabajan con datos continuos (Geoestadística), los datos encontrados en Epidemiología espacial toman frecuentemente la forma de

conteos de casos de enfermedad y/o muerte en determinadas regiones (Lawson, 2001).

Las técnicas de representación geográfica de datos en Epidemiología suelen ser un paso previo para la visualización de la estructura espacial y constituyen una herramienta potente para el análisis exploratorio de los datos antes de proceder a su modelización. Además, en las últimas décadas, el desarrollo de técnicas estadísticas avanzadas y la potencia de los recursos informáticos han promovido el desarrollo de modelos con diferentes enfoques como el frecuentista, bayesiano y multinivel (Lawson, 2003).

Para analizar la visualización de la estructura espacial en los datos se recurre a herramientas de la estadística espacial, como también de la Matemática Discreta. La matemática discreta no es una rama tan definida de la matemática como puede serlo álgebra, geometría, trigonometría o cálculo, en esta se trabaja con conjuntos finitos de objetos lo que incluye una gran variedad de tópicos y técnicas que surgen en cada día de la vida. La Teoría de Grafos es un tema avanzado a nivel universitario que forma parte de Matemática Discreta y permite abordar cuestiones muy distintas, algunas de ellas tienen origen en juegos o pasatiempos, otras en preguntas muy importantes y diversas de la ciencia y la técnica, siendo este uno de los motivos del gran desarrollo que tuvo en las últimas décadas como también el auge que posee actualmente (Braicovich, 2009).

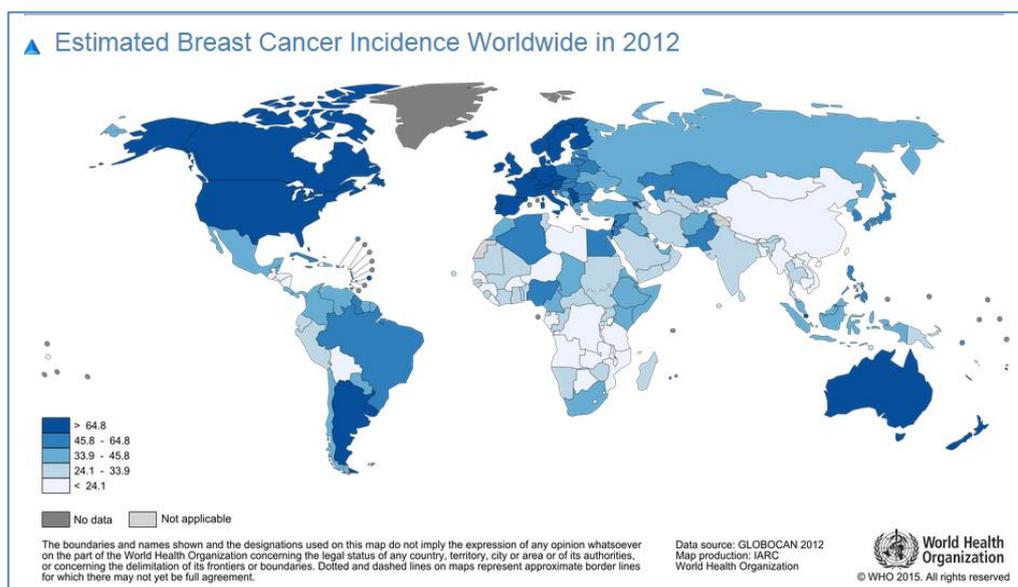
Muchas situaciones pueden ser modelizadas, aunque sea en una primera aproximación, mediante diagramas llamados grafos, constituidos por puntos y líneas, tal que estas últimas conectan pares de esos puntos o uno consigo mismo. Los puntos son denominados *vértices* o *nodos* y las líneas *aristas* o *arcos*, según el grafo no sea dirigido o lo sea respectivamente. Los grafos aparecen en diversas disciplinas bajo distintos nombres, por ejemplo: redes en ingeniería y en salud, sociogramas en psicología, organigramas en economía y planificación, diagramas de flujo en programación, diagramas de estado en informática, estructura molecular en química, etc. Según indica Wilson (1979) el primero en designar a estos diagramas “grafos” fue Sylvester en 1878 al publicar sus resultados sobre Teoría de Invariantes en Química.

Distribución Espacial del Cáncer de Mama (CM)

El cáncer es una enfermedad gradual altamente relacionada con condiciones ambientales, factores biológicos (por ejemplo la carga genética, el sexo, la edad, la raza), factores sociales (como puede ser la ocupación, el nivel de instrucción) y forma de vida (entre otras, tabaco, alcohol), sus causas son complejas (Riboli, 2002), es decir factores contextuales que tienen efectos en la condición humana y en la salud. El patrón de la incidencia del cáncer, por lo tanto, está influenciado por la región geográfica, la situación cultural y las características de la población (Stewart y Kleihues, 2003). El cáncer muestra variaciones espaciales y conocerlas resulta útil para establecer si ocurren patrones de la variación al azar, ya que el conocimiento del patrón de la ocurrencia del cáncer y de su distribución espacial es muy importante para identificar a grupos más vulnerables de la población, lo que permitirá definir políticas sanitarias más convenientes para la prevención, la supervisión y el control.

A nivel mundial el CM es el cáncer más frecuente y la primera causa de muerte en mujeres. Existen diferencias entre países, como se puede observar en la figura 1, Argentina presenta una elevada tasa de incidencia de CM.

FIGURA 1. Estimación de la Incidencia de CM a nivel mundial en 2012.



Fuente: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx

Cabe mencionar, en relación a la figura anterior, que en el año 2012 se diagnosticaron 1,67 millones de nuevos casos de CM (25% de todos los

cánceres), siendo el más común en mujeres, tanto en las regiones más como en las menos desarrolladas, con un poco más de casos en estas últimas (883.000 casos) que en las primeras (794.000). Las tasas de incidencia varían casi cuatro veces a través de las distintas regiones del mundo, con tasas que van del 27 por 100.000 en África Medio y Asia del Este a 92 por 100.000 en América del Norte.

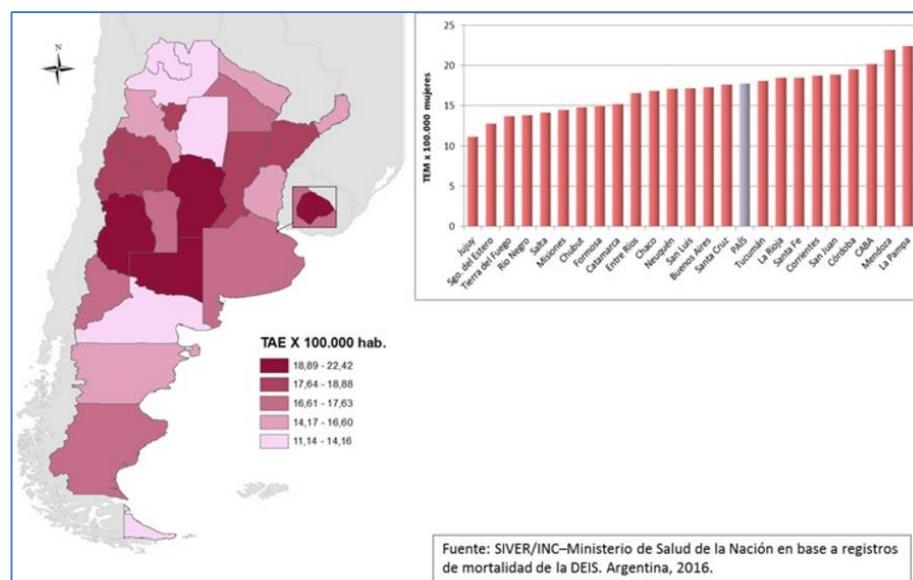
Según informes del año 2012 de la Organización Mundial de la Salud (OMS) y de la Organización Panamericana de las Salud (OPS) se registran más de 408.200 nuevos casos y más de 92.000 muertes por año. El CM es el primer cáncer en las Américas en cuanto a casos nuevos y el segundo en cuanto a muertes por cáncer en mujeres según informes del Banco Mundial. En cuanto a estimaciones de las tasas de incidencia en las Américas se posiciona en primer lugar las Bahamas, seguido de Uruguay y Estados Unidos y no muy lejos de allí se sitúa la Argentina. En cuanto a mortalidad las mayores tasas se presentan en Bahamas, Trinidad y Tobago y Uruguay.

En particular en la Argentina los datos de incidencia y de mortalidad del CM son fragmentados, ya que no existe un registro único. Existen diversos registros poblacionales, los que la mayoría de las veces dependen de voluntades individuales, ya que el apoyo gubernamental no es continuo en cuanto a la unificación de registros. Por este motivo, la información recabada tiene baja calidad y/o es Viniegra *et al.*, (2010).

A pesar de las falencias en los sistemas de registros, el CM es un problema preocupante, en Argentina, cada año, mueren, por esta causa, aproximadamente 5.400 mujeres, según datos del Ministerio de Salud de la Nación (MSAL, 2009) y se estima que se diagnostican alrededor de 17.000 casos nuevos por año (Ferlay *et al.*, 2010).

A nivel Provincial existen diferencias entre las provincias. Para traducir el verdadero riesgo de morir por CM, debe realizarse un ajuste por la estructura poblacional de cada jurisdicción, ya que hay jurisdicciones con mayor población añosa, que es la que presenta un mayor riesgo. Al realizar este ajuste se observa la situación comparativa de las distintas provincias y del promedio nacional como muestra la figura 2.

FIGURA 2. Estimación tasas estandarizadas por edad según población mundial por 100.000 mujeres de CM en la Argentina año 2014



Fuente: <http://www.msal.gov.ar/inc/index.php/acerca-del-cancer/estadisticas>

La figura 2 muestra los cálculos estimados de Mortalidad en cada una de las jurisdicciones de Argentina comparados con la tasa promedio País. En particular, se puede observar que la Provincia de Neuquén está próxima, pero por debajo de la tasa de mortalidad promedio del país. A pesar que la tasa estandarizada de mortalidad posiciona a la provincia de Neuquén con una mortalidad levemente inferior a la del resto del país, es el cáncer la causa de muerte más frecuente en la provincia, y en particular el CM la principal causa de muerte en mujeres, con una tasa de mortalidad de 20,5 por 10 mil habitantes, y ha mantenido una tendencia en ascenso desde el año 1986 como se indica en el Informe de CM.¹

El año 2014, la cantidad de mujeres en la provincia de Neuquén es de 297.764 (Proyección del censo, 2010). El subsector público de salud de Neuquén brinda cobertura a toda la población, entre ellas a aquellas pacientes que no cuentan con obra social o prepaga. Aproximadamente un 35% de la población sólo cuenta con cobertura por Salud Pública en Neuquén, lo que representa 104.218 mujeres.² Sólo en el subsector público de salud de Neuquén se diagnostican aproximadamente 100 nuevos casos de CM anualmente y mueren por CM 63

¹Una mirada en la provincia del Neuquén desde la Epidemiología que se llevó a cabo en un encuentro Interjurisdiccional para el logro de acuerdos para un Programa de abordaje del Cáncer de Mama". Sala de Situación, Ministerio de Salud de Neuquén. Septiembre del 2011.

²Comité Provincial de Biotecnologías Salud Pública, Provincia de Neuquén. Febrero 2014.

mujeres neuquinas por año.³ Hay que aclarar que en el hombre el CM es infrecuente en la provincia.

El tratamiento oportuno y las posibilidades de curación de esta enfermedad dependen, entre otras causas, del volumen del tumor al momento del diagnóstico. Al igual que para otras enfermedades, la detección precoz es fundamental, ya que los tumores de menos de 1 centímetro tienen hasta el 90% de probabilidad de curación. Además, el diagnóstico de tumores más pequeños permite aplicar tratamientos menos agresivos, reduciendo la probabilidad de tener que realizar mastectomías y quimioterapias Viniegra *et al.*, (2010).

Según Durand *et al.*, (2013), el estudio mamográfico es considerado el primero y único test para la detección temprana, el mismo se realiza en pacientes sin síntomas a modo de “screening”. En medicina se entiende por screening la estrategia aplicada sobre una población para detectar una enfermedad en individuos sin síntomas, la cual se puede llevar a cabo de dos maneras:

- i. La primera se refiere al tamizaje oportunista, es decir a la realización de estudios mamográficos indicados en el contexto de exámenes ginecológicos programados, prescriptos por el médico o por decisión propia de la paciente.
- ii. La segunda es a través de programas poblacionales, generalmente organizados por los servicios de salud, en los que, a través de actividades programadas de búsqueda activa, se invita a mujeres en rangos de edad específicos a la realización de mamografías periódicas de tamizaje.

La organización sanitaria en la provincia de Neuquén, se encuentra dividida en 7 regiones a nivel macro y en 28 áreas programáticas a nivel micro. Es meritorio destacar que, en el marco de la lucha contra el CM la provincia cuenta desde el año 2009 con un programa exclusivo denominado “**Red Provincial de Patología Mamaria**” cuyo objetivo es la prevención y control de la enfermedad (Novoa, 2009).

Objetivos de la investigación

Identificar patrones espaciales de indicadores de cáncer de mama en la provincia del Neuquén.

Los objetivos específicos propuestos son:

³Informe Registro Provincial de Tumores de la Provincia de Neuquén. 2011

- A. Describir patrones espaciales de mortalidad de CM de las áreas programáticas de la provincia de Neuquén.
- B. Proponer estructuras de conexión espacial.
- C. Comparar distintas redes de conexión espacial.
- D. Representar la distribución espacial del CM.
- E. Estimar tasas de mortalidad por CM según clase etaria en riesgo determinado según lo establecido por la OMS.
- F. Estimar las tasas de mortalidad por CM restringida a la distribución espacial.

Pertinencia de la investigación desde la óptica de la Salud Pública.

El conocimiento de la distribución geográfica de los casos de CM dentro de la provincia es de ayuda para planificar y programar acciones de control y vigilancia epidemiológica. Es por ello que en este trabajo se propone calcular tasas de mortalidad según las áreas programáticas en la que se encuentra dividida sanitariamente la provincia del Neuquén teniendo en cuenta su distribución geográfica (Pina, 2010).

Dentro de la salud pública se considera que el CM es un problema que debería ser analizado con herramientas estadísticas-matemáticas a partir de datos empíricos locales Viniegra *et al.*, (2010). Debido a esto, la propuesta es enmarcar este trabajo de tesis en los siguientes aspectos.

4.1. Primer Aspecto: Seguimiento y Control

A pesar de los grandes avances en clínica y en tecnologías de la información que se produjeron en los últimos años, el seguimiento y control de las enfermedades en la población sigue dependiendo de las fuentes de información tradicionales, como son las estadísticas de incidencias y/o mortalidad (Novoa, 2009). En la provincia de Neuquén no se han detectado trabajos que muestren tasas mortalidad ajustadas por edad de CM discriminadas por área programa.

Existen indicios sobre investigaciones globales “epidemiológicas”, algunas de ellas se mencionan a continuación por considerar que son significativos los aportes encontrados para el desarrollo de este trabajo de tesis.

En países de Latinoamérica estudios realizados por autores como Uscanga Sánchez *et al.*, (2014) y Nigenda *et al.*, (2011) son algunos indicios que muestran como plantean un conjunto de indicadores para poder dar seguimiento a

políticas/programas de CM sobre la base de lo que es considerado como relevante y factible en términos de la disponibilidad de información.

En nuestro país hay estudios realizados por la Organización panamericana y mundial de la salud, Viniegra *et al.*, (2010), donde se diagnosticó la situación de programas Nacionales y/o Provinciales. En este informe se detalla el estado actual del sistema de salud frente a la enfermedad del CM. Dentro del informe no se muestran resultados referidos a la provincia de Neuquén como sí para otras provincias.

En Neuquén a nivel ministerial en los años 2010 y 2014 se han realizado informes sobre el estado de situación de la enfermedad en términos generales de la provincia como tasas de mortalidad en el periodo 1989 al 2010(Comité Provincial de Biotecnología de Neuquén, 2014; Ministerio de Salud de Neuquén, 2011; Novoa, 2009).

4.2. Segundo Aspecto: Prevención.

En las últimas décadas se ha considerado prioritario que los sistemas de salud garanticen la prestación de los servicios de salud preventivos, ya que se reconoce su impacto en términos no sólo de las mejoras en la calidad de vida de los pacientes con tratamiento oportuno, sino también por la reducción de los costos de atención, curación y tratamientos a largo plazo. Debido a esto ha aumentado el interés en el uso de herramientas computacionales basadas en modelos matemáticos para soporte del proceso de toma de decisiones (Durand *et al.*, 2013). Con esta investigación se pretende estudiar el CM analizar su distribución espacial y que sea un aporte para la toma de decisiones para los gestores en salud en término de prevención.

Algunas investigaciones recientes referidas al tema que han utilizado modelos matemáticos y que fueron inspiradores para el trabajo que se presenta, son los siguientes: Griffin *et al.*, en el año 2007 presentó una investigación que tiene por objetivo determinar, en una red geográfica, el número de nuevos centros de salud y la mejor ubicación de ellos, así como los servicios que debería ofrecer cada uno, presenta un modelo de optimización que utiliza costos y realiza ciertas consideraciones geográficas. Mahar *et al.*, (2011) presenta una investigación donde lo central es el análisis de las redes de hospitales con múltiples ubicaciones, referidas a cómo se pueden aprovechar los beneficios de la puesta

en común al decidir dónde colocar los servicios especializados, en este modelo se realizan consideraciones de carácter financiero y niveles de servicio del paciente Wang (2012) hace una revisión de la literatura acerca de los diversos índices de accesibilidad para su comparación y evaluación, con énfasis en el desarrollo metodológico y las implicancias para las políticas públicas. Wang *et al.*, (2013) abordan el problema de la planificación en términos de cómo deberían ser distribuidos los recursos para lograr la mayor igualdad posible en el acceso de los pacientes a los proveedores de servicios. Y por último Huairén *et al.*, (2014), abordan la disparidad espacial de la salud en términos de indicadores accesibilidad basada en la red y según Información Geográfica Sistema (SIG).

4.3. Tercer Aspecto: Registro y Estadísticas.

Los eventos de salud, tradicionalmente, se han registrado de forma agregada para grandes áreas geográficas determinadas administrativamente, como son las provincias, departamentos, zonas y, con restricciones, los municipios. En los últimos años se ha incrementado la disponibilidad de obtención o construcción de bases de datos espaciales en formato digital que permiten vincular direcciones individuales con su localización espacial, como el sistema integrado de información Sanitaria Argentino (SIISA)⁴.

Por último, las técnicas estadísticas tradicionales solían analizar datos socioeconómicos, ambientales y de salud que, aun cuando son importantes y útiles, no consideran el espacio geográfico donde ocurren los eventos; en otras palabras, ellos no miden cómo los eventos de salud están asociados en un contexto de espacio. Debido a que el proceso de salud-enfermedad no es el resultado directo de los factores de riesgo y las condiciones biológicas/genéticas, los determinantes sociales y ambientales que influyen la salud y el bienestar de las poblaciones necesitan ser integrados dentro de los estudios epidemiológicos modernos. La característica especial acerca del análisis espacial es que “el lugar hace la diferencia”; por lo tanto, la ubicación de los eventos necesita ser integrada en el análisis. El análisis de los datos se torna en una tarea dinámica e integrativa, donde la información es mucho más que simples números o características estáticas, Piña (2010).

⁴<https://www.sisa.msal.gov.ar>

Para llevar a cabo esta investigación se trabajó con los registros de defunciones causados por CM y el tamaño poblacional estimado por la Dirección de Estadística de la Subsecretaría de Salud del Ministerio de Salud y Desarrollo Social de la Provincia de Neuquén entre los años 2001 y 2012. A partir de esta información, se calcularon las tasas crudas y las tasas ajustadas por edad de mortalidad (por cada 100.000 habitantes) según el método indirecto (población Argentina de referencia) en las 28 áreas programáticas de la provincia para una serie temporal del 2001 al 2012.

Estructura del Trabajo de Tesis

En este trabajo de tesis se aborda el marco conceptual en dos capítulos, en el primero se contextualiza la enfermedad del CM en la población mundial, nacional y provincial, como también su pertinencia en salud pública, en el segundo capítulo se presentan las herramientas estadísticas usuales para detectar la presencia de autocorrelación espacial y comparar redes de conexión entre observaciones. En el capítulo 3, se desarrollan los distintos métodos de cálculo de tasas descriptivas y las tasas asociadas a la probabilidad del CM. En el capítulo 4, se describen los datos a trabajar con sus distintas fuentes y la construcción de índices que fueron implementados en el estudio de la distribución espacial del CM. Los resultados y discusiones de los análisis estadísticos ilustrados sobre la enfermedad de CM en Neuquén con las interpretaciones, se muestran en el capítulo 5 y por último las conclusiones en el capítulo 6.

CAPÍTULO 2

ESTADÍSTICA EPIDEMIOLOGÍA ESPACIAL

La importancia del espacio dentro del estudio estadístico de variables de la salud es incuestionable. Los eventos relacionados con la salud, como enfermedad, muertes, nacimientos, exposición a riesgos o cualquier otro evento, pueden relacionarse con el sitio donde ocurren, por un par de coordenadas, una dirección o un área. El estudio de la distribución geográfica de los eventos de salud se volvió de gran interés entre los epidemiólogos durante las últimas tres décadas, a causa de un retorno a un concepto de salud más amplio y holístico, en el cual los individuos son vistos en su contexto socio-cultural-ambiental (Pina, 2010; Alazraqui et al., 2007).

El estudio de la distribución geográfica y la asociación espacial de los eventos en salud se denomina epidemiología espacial. Este concepto ha sido empleado para describir estudios sobre las causas y prevención de las enfermedades, empleando diferentes perspectivas de análisis en las que la localización de los eventos es un componente fundamental. Es una herramienta de investigación que facilita analizar los factores externos que pueden fomentar el desarrollo de una enfermedad y conocer en qué condiciones impacta más a la población.

John Snow fundó la epidemiología entre los años 1848 a 1854, cuando mostró la distribución geográfica de la mortalidad por cólera en la ciudad de Londres y permitió la resolución del brote mediante la identificación del foco. Anteriormente Seamon y Pascolis en 1789 ubicaron casos de fiebre amarilla en Nueva York como primeros indicios de la localización de eventos de la Salud. El inicio del relevamiento del dato “lugar de residencia del paciente” comienza a registrarse desde el siglo XVII. La cooperación entre geógrafos y médicos se inicia en el siglo XIX, pero este paradigma comienza con el declive, como consecuencia de los avances de la microbiología y los descubrimientos de Louis Pasteur “paradigma de los gérmenes” a fines del siglo XIX y comienzos del siglo XX. Durante el siglo XX continúa el declive de la importancia del espacio geográfico debido a las enfermedades crónicas y factores de riesgo como consecuencia directa del estilo de vida. A partir de allí la epidemiología comienza a recorrer un camino de transición.

El crecimiento del interés en los aspectos espaciales está estrechamente relacionado con la mayor atención que se presta a los problemas ambientales, los desarrollos tecnológicos derivados de la confluencia entre la epidemiología, la estadística, las herramientas informáticas y otras disciplinas científicas.

La incorporación de la perspectiva espacial a los estudios en salud lleva importantes contribuciones al entendimiento de los procesos de salud-enfermedad y puede conducir a diferentes resultados de aquellos obtenidos en los estudios que no consideran el espacio geográfico. Junto con la descripción de las características de las poblaciones es necesario precisar dónde ocurren los casos, qué clases de servicios busca la población, el lugar de potencial riesgo ambiental y las áreas socialmente vulnerables.

El patrón espacial de los problemas de salud es fundamental para comprender la exposición y desarrollar estrategias para prevenir eventos futuros, más allá de la naturaleza del proceso que subyace (contagioso, influenciado por el ambiente o relacionado con la variabilidad genotípica). La característica distintiva del análisis estadístico de datos espaciales es que el patrón de las localizaciones (objetos espaciales), la asociación entre los valores observados en diferentes localizaciones (dependencia espacial) y la variación sistemática del fenómeno en las distintas localizaciones (heterogeneidad espacial) se convierte en el mayor foco de investigación (Anselin, 1988).

La estadística presenta la mayor parte de sus resultados bajo el supuesto de independencia, cuando se viola esta condición, los resultados obtenidos ya no son válidos. En el caso de la estadística espacial el problema que se plantea es la pérdida de la condición de independencia de las observaciones tomadas en un área determinada. Así, en este tipo de modelos se supone que si encontramos una determinada observación en un punto de esta área es más fácil (o inversamente más difícil) encontrar observaciones semejantes en puntos próximos a éste que en puntos alejados. Otro caso típico de violación de la condición de independencia se presenta en el estudio de las series temporales. Los modelos de dependencia temporal han sido objeto de un profundo estudio y se conocen estructuras que expresan el comportamiento de las variables con respecto al tiempo. No ocurre lo mismo con aquellos modelos que intentan captar la dependencia de una variable en función de su "entorno". Esto se debe a la diferencia que existe en la relación en el espacio, es decir en el caso de datos longitudinales su relación es unidireccional, en cambio en los datos

espaciales su relación en el espacio es multidireccional. No es más importante la dimensión temporal del problema que la espacial, pero en el desarrollo científico no han recibido igual atención estos dos problemas, en buena parte debido a la complejidad que involucra las estructuras de dependencia espacial (López Hernández et al., 2000).

A continuación se detallan conceptos que serán útiles para abordar el tratamiento de datos espaciales.

2.1. Datos Espaciales

Los datos espaciales son medidas u observaciones que tienen asociada una localización específica y, de forma general, pueden definirse como datos georeferenciados (López Hernández et al., 2000). Estas observaciones pueden ser del tipo discreto o continuo. La localización que tiene asociada cada observación puede ser un punto cualquiera del espacio, o bien estar asociada a un área, dentro de una superficie, sobre la que se ha realizado una partición. Esta partición puede ser regular o irregular, y habitualmente se conoce como lattice o retícula. En función de toda esta casuística se presentan diferentes modelos que llevan asociados diferentes tratamientos estadísticos.

Según Cressie (1993), cualquier proceso espacial puede caracterizarse mediante un proceso estocástico. Sea $s \in \mathcal{R}^d$ una posición genérica localizada en el espacio Euclideo de d -dimensiones y sea $Z(s)$ una variable aleatoria localizado en s . Consideremos que s varía sobre el conjunto $D \subset \mathcal{R}^d$ generando el campo aleatorio multivariado o proceso aleatorio $\{Z(s): s \in D\}$. Si $d = 2$ estamos en el caso espacial y si $d = 3$ en el espacio-temporal. Dependiendo de las características de D se diferencian tres tipos de datos: geoestadísticos, agregados o lattice y mapas de puntos o procesos puntuales.

Según López et al., (2000), los datos asociados a referencias espaciales se suelen descomponer en:

$$\text{DATO} = \text{Variación a gran escala} + \text{Variación a pequeña escala}.$$

Esta primera componente, variación a gran escala, se suele corresponder con la variación no estocástica del proceso y resulta ser el resultado de una tendencia global que puede ser modelada atendiendo o no a criterios espaciales. La variación a pequeña escala es el residuo resultante de eliminar la tendencia. Una vez eliminada ésta podemos encontrar patrones de comportamiento que obedezcan al entorno en el que se encuentran las observaciones. Esta variación

espacial tiene una estructura de dependencia estocástica y puede también modelarse atendiendo a criterios espaciales.

En las siguientes secciones se describen los tipos de datos espaciales.

2.1.1 Datos Geoestadísticos

Son mediciones de un conjunto de variables aleatorias, continuas o discretas, tomadas en puntos fijos de un área geográfica concreta. Des un subconjunto fijo de \mathcal{R}^d que contiene un rectángulo de d –dimensiones con volumen positivo. Sus aplicaciones clásicas son en salud ambiental para medir contaminación ambiental, en climatología para predecir índices bioclimáticos, minería y geología para predecir yacimientos, y demás. El objetivo suele ser predecir el valor de la variable en toda la región continua de estudio, y/o una predicción para un instante temporal no medido (Cressie, 1993).

2.1.2 Mapas de Puntos o Proceso Puntual

Sea D una región geográfica plana y finita objeto en estudio. Un proceso puntual se puede definir como el proceso estocástico que genera un conjunto contable de eventos en una región geográfica plana D (Diggle, 2003). De forma general, la región D es finita de \mathcal{R}^d , es decir, $D \subset \mathcal{R}^d$.

Los campos de aplicación de los procesos puntuales son muy variados y diversos. Se han aplicado en ámbito forestal, en logística y urbanismo, en salud pública, en microbiología y demás. En general se pueden aplicar a cualquier ámbito en el que se quiera estudiar cómo se distribuyen un conjunto de puntos (eventos) en una región finita y plana del espacio (Quesada Rico, 2015).

2.1.3 Datos Agregados o Lattice.

Cuando el dominio de las observaciones de D es fijo y discreto numerable estamos en el caso de datos agregados, también llamados datos en áreas pequeñas, diseños ecológicos, datos en rejilla o datos reticulares. A veces también se denominan procesos reticulares. La forma de presentar estos datos está asociadas a datos regionales, como secciones censales, municipios, comarcas, etc.

En este tipo de diseño los datos son exhaustivos, es decir, los datos proporcionan información sobre todas las localizaciones en el dominio de observación, es decir sobre el conjunto D , además D está asociado a una estructura de vecindades. El objetivo prioritario es la modelización, suavizado de las observaciones e identificación de regiones con similares valores, no se realiza predicción (Quesada Rico, 2015).

En el contexto de epidemiología, la medida en cada una de las áreas o particiones suele ser una variable discreta, por ejemplo, casos de muerte o enfermedad por una determinada causa. Desde el punto de vista geográfico y ecológico, si se busca examinar la relación entre la variable respuesta en salud como tasas de muerte / enfermedad y exposiciones como pueden ser factores medioambientales, socioeconómicos, estilos de vida, etc. (González, 2015). teniendo en cuenta la estructura poblacional. En este trabajo se profundizó en el análisis para datos agregados o lattice, debido a que las variables aleatorias medidas están agregadas para las 28 áreas programáticas de la provincia.

2.2. Notación Elementales de Grafos

Muchas situaciones de la vida real pueden ser modelizadas, aunque sea en una primera aproximación, mediante diagramas, llamados grafos, construidos por puntos y líneas, tal que estas últimas conectan pares de esos puntos o uno consigo mismo. Los grafos aparecen en diversas disciplinas bajo distintos nombres, por ejemplo: redes en ingeniería, sociogramas en psicología, organigramas en economía y planificación, diagramas de flujo en programación, diagramas de estado en informática, estructuras moleculares en química, etc. Según indica Wilson (1979) el primero en designar a estos diagramas "grafos" fue Sylvester en 1878 al publicar sus resultados sobre Teoría de Invariantes en Química (Braicovich *et al*, 2009).

A continuación se presentan algunos conceptos de grafos, aquellos que serán utilizados en este trabajo.

Un grafo es una terna $G = (V, U, \Phi)$ que consiste en dos conjuntos no vacíos y disjuntos, V y U , de elementos llamados vértices y aristas respectivamente, y de una función Φ , frecuentemente llamada relación de adyacencia, que asocia a cada arista de U un par no ordenado de vértices (no necesariamente distintos)

de G . Se puede extender la definición de grafo para cuando $U = \emptyset$, en este caso la terna asociada es $(V, \emptyset, \emptyset)$ y se tiene el grafo discreto.

Si u es arista del grafo G y a y b son vértices tales que $\Phi(u) = (a, b)$, se dice que la arista u tiene extremos en los vértices a y b . Una arista en la que coinciden ambos extremos es llamada bucle o lazo. Los vértices que son extremos de una misma arista, que no es bucle, se llaman vértices adyacentes o vecinos. Un grafo G es de orden n si $n = |V|^5$ y de tamaño m si $m = |U|$, es decir que el orden y el tamaño coincide con la cantidad de vértices y aristas del grafo G respectivamente.

Se aceptan aristas diferentes pero con los mismos extremos, las mismas son llamadas aristas múltiples o paralelas. Cabe aclarar que en alguna bibliografía a las estructuras que nos ocupan se las llama grafos en caso que la función Φ sea inyectiva, es decir cuando no existen aristas múltiples y se llaman multigrafos independientemente si la función es o no inyectiva. En el desarrollo de este trabajo los llamaremos grafos en general y diremos que un grafo es simple cuando no tiene aristas múltiples ni bucles.

Sea $G = (V, U, \Phi)$ un grafo, diremos que G es:

- Finito si los conjuntos V y U son conjuntos finitos. En esta investigación solo se consideran grafos finitos.
- Cuando las aristas del grafo poseen un valor numérico asociado – denominado peso- se dice que el grafo es valuado. Si no existen valores numéricos asociados a las aristas significan que todas las aristas tienen el mismo peso.
- Trivial si es discreto y de orden 1.

El grado de un vértice v del grafo G , se nota $gr(v)$ y es el número de aristas con extremos en v , contando doble cada bucle. Un vértice se dice aislado cuando ninguna arista lo tiene por extremo, por lo tanto el grado de dicho vértice es igual a cero. Un vértice se dice pendiente cuando su grado es igual a uno. Un vértice v de un grafo G es llamado vértice par o vértice impar si el grado del mismo es par o impar, respectivamente.

⁵ Dado un conjunto A , la anotación $|A|$ indica el cardinal del conjunto

Proposición 1: Si G es un grafo con m aristas y n vértices v_1, v_2, \dots, v_n , se tiene que la suma de los grados de los vértices del grafo G es igual al doble del

número de aristas, es decir:
$$\sum_{i=1}^n gr(v_i) = 2 \cdot m$$

Es directa la demostración de esta igualdad, ya que al sumar los grados de los vértices de un grafo G , cada arista es contada dos veces, una vez por cada uno de los vértices en los cuáles incide la misma. Cabe aclarar que en el caso particular que dicha arista sea un bucle, la misma es contada dos veces en el mismo vértice, como ya se ha hecho mención con anterioridad (Braicovichef *al.*, 2009).

Se denomina vecindad o conjunto de vecinos del vértice N_v , al conjunto de vértices adyacentes de v . En grafos simples el cardinal de la vecindad N coincide con el grado del vértice v .

Un grafo $G = (V, U)$ es regular cuando todos sus vértices tienen igual grado. Si todos los vértices tienen grado k , se dice que G es regular de grado k o simplemente, k -regular. Si el grafo G es k -regular y de orden n , por la **proposición 1**, se tiene que el número de aristas del mismo es igual al valor $\frac{k \cdot n}{2}$

Un grafo simple es completo si cada vértice es adyacente a todos los restantes, notaremos K_n al grafo completo de orden n . Todo grafo completo de orden n es $(n-1)$ regular y tiene, por lo tanto $\frac{n \cdot (n-1)}{2}$ aristas.

Un grafo $G = (V, U)$ es bipartito si el conjunto V puede ser particionado en dos subconjuntos, V_1 y V_2 , tal que cada arista de G tiene un extremo en el conjunto V_1 y otro extremo en el conjunto V_2 . En particular, si G es un grafo tal que cada vértice del conjunto V_1 es adyacente a cada vértice del conjunto V_2 , entonces el grafo G es bipartito completo. Notaremos con $K_{r,s}$ al grafo bipartito completo, donde $|V_1| = r$ y $|V_2| = s$.

El grafo complementario de un grafo simple sin bucles $G = (V, U)$ es el grafo $\bar{G} = (V, W)$, en el cuál dos vértices serán adyacentes en \bar{G} sí, y sólo si, no lo eran en G . Si el grafo G es de orden n , es directo ver que la suma de los

cardinales de los conjuntos de aristas U y W es igual al valor $\frac{n \cdot (n-1)}{2}$, número que ya hemos visto que indica la cantidad de aristas del grafo completo de orden n .

Subgrafo: Dado un grafo $G = (V, U)$ se dice que el grafo $G' = (V', U')$ es un subgrafo de G si se tiene que: $V' \subseteq V$ y $U' \subseteq U$. Dado un grafo G , podemos decir que G' es subgrafo:

- Cubriente si $V' = V$.
- Inducido por U' si está constituido por las aristas de U' y los vértices de G sobre los cuales inciden estas aristas.
- inducido por V' si está constituido por los vértices de V' y las aristas de G cuyos extremos pertenecen a V' .

Sea $G = (V, U)$ un grafo, $V' \subseteq V$ y $U' \subseteq U$. Notaremos con $G - V'$ al subgrafo de G inducido por el conjunto de vértice $(V - V')$ y con $G - U'$ al subgrafo obtenido a partir de G eliminando las aristas pertenecientes a U' . Particularizando este caso, tomando conjuntos V' y U' de cardinalidad igual a 1, tenemos:

- El subgrafo restante respecto de un vértice v , es el subgrafo obtenido a partir del grafo original omitiendo el vértice v y todas las aristas incidentes en él, simbolizaremos a este subgrafo como \tilde{G}_v .
- De manera similar al caso anterior, se dice que el grafo que se obtiene al quitar una arista a al grafo G es llamado subgrafo restante respecto de una arista a y se nota como \tilde{G}_a .

Dado un grafo $G = (V, U)$ diremos que una cadena C de longitud L , $L \geq 1$ es una sucesión de vértices y aristas $C : v_1, a_1, v_2, a_2, v_3, \dots, v_L, a_L, v_{L+1}$, tal que la arista a_i tiene por extremos los vértices v_i, v_{i+1} , $1 \leq i \leq L$; no necesariamente $a_i \neq a_j$ y eventualmente $v_i = v_{i+1}$.

Cualquier subsucesión de la cadena C determina una subcadena de C . Admitimos que cada vértice define una cadena nula, de longitud cero.

Sea la cadena $C : v_1, a_1, v_2, a_2, v_3, \dots, v_L, a_L, v_{L+1}$, la misma es:

- simple si $a_i \neq a_j$ cualquiera sean $i \neq j$, $i, j \in \{1, \dots, L\}$. Es equivalente decir que una cadena es simple si no se repiten aristas en la sucesión.
- elemental si $v_i \neq v_j$ cualquiera sean $i \neq j$, $i, j \in \{1, \dots, L+1\}$ excepto que v_1 y v_{L+1} correspondan a un mismo vértice. De manera equivalente se puede decir que una cadena es elemental si no se repiten vértices, excepto v_1 y v_{L+1} .

Convendremos que toda cadena nula es elemental y simple, ya que en la misma no se repiten aristas ni vértices.

Dada una cadena $C : a_1, a_2, \dots, a_L$, $L \geq 1$, si la arista a_1 incide en el vértice x y la arista a_2 no incide en dicho vértice y además si la arista a_L incide en el vértice y y la arista a_{L-1} no lo hace, se dice que la cadena C lleva desde x hasta y , la notaremos $C : x \rightarrow y$. En el caso que se tenga que $x = y$, se dice que la cadena es cerrada y es llamada ciclo, en caso contrario se dice que la cadena es abierta.

Observación:

Habitualmente, cuando no haya lugar a confusión, al referirnos a cadenas omitiremos explicitar sus vértices y su longitud. Es obvio que sólo si el grafo G es simple la cadena queda bien definida dando sólo la sucesión de los vértices correspondientes.

La distancia entre los vértices u y v , de un grafo G se denota $d_{u,v}(G)$ y es la menor de las longitudes de las cadenas que conectan u con v . En el caso que los vértices no estén conectados entre sí, se considera que la distancia entre ellos es $+\infty$.

Un grafo G es conexo si es trivial o, equivalentemente, para todo par de vértices de G existe al menos una cadena que los une. Caso contrario, G es disconexo o no conexo.

Un grafo desconexo consiste en dos o más subgrafos conexos, cada uno de estos subgrafos es una componente conexa. Si G_1, G_2, \dots, G_t son las t componentes conexas de G , se tiene que: $G = \bigcup_{i=1}^t G_i$.

Es obvio que si el grafo es conexo, se dice que tiene una componente conexa.

Diremos árbol a todo grafo conexo, sin ciclos. En particular, todo vértice aislado es un árbol trivial y toda cadena elemental y abierta es un árbol. El tamaño de un árbol de orden n es siempre igual a $n - 1$. Un grafo sin ciclos es llamado grafo acíclico.

Un Grafo cordal es aquel que no posee ciclos inducidos de 4 vértices, equivalentemente un grafo es cordal si todo ciclo de longitud mayor o igual a 4 tiene una arista entre vértices no consecutivos del ciclo.

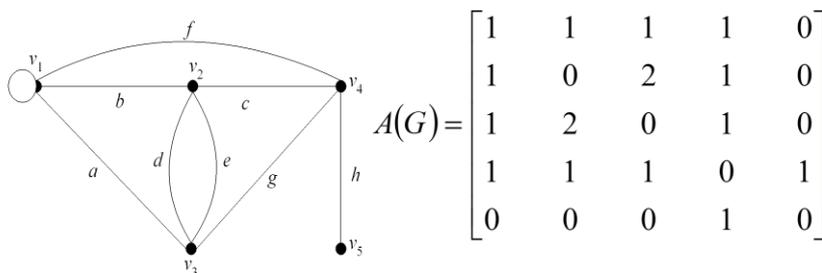
Existen distintas formas de representar un grafo, estos pueden ser representados mediante conjuntos, pudiendo ser la representación geométrica o topológica, también la misma puede ser realizada mediante matrices o mediante listados de aristas, siendo esta la adecuada para trabajar en software.

Las matrices que se presentan son aquellas ligadas a nociones asociadas a pares de vértices, las que son sumamente útiles para trabajar conexidad y transitabilidad.

Sea G un grafo de orden n , $n \geq 1$, la matriz adyacencia de G se define como $A(G) = (a_{ij})$ donde a_{ij} es el número de aristas de la forma (i, j) , eventualmente $i = j$. También se las suele denominar matriz de conexión.

A continuación se presenta modo de ejemplo en la figura 3 un grafo de 5 vértices, por lo que la matriz adyacencia será de orden 5.

FIGURA 3. Grafo G no dirigido de 5 vértices y Matriz de Adyacencia



De la definición de matriz adyacencia, resulta que:

- $A(G)$ es simétrica, ya que las aristas se indican mediante pares no ordenados.
- La matriz adyacencia del grafo completo K_n es de orden n y son iguales a 0 los elementos de la diagonal principal e iguales a 1 los restantes, ya que no tiene bucles y cada vértice está relacionado con todos los demás mediante una única arista.

Un grafo G es planar si admite una representación en el plano tal que curvas correspondientes a aristas distintas no se cortan salvo, tal vez, en sus puntos extremos. A tal representación se la denomina representación plana de G o inmersión en el plano de G .

Un grafo planar simple G es un grafo planar maximal si el agregado de una arista entre dos cualesquiera de sus vértices no adyacentes el grafo deja de ser planar.

2.3. Grafos y Redes

Como se mencionó en el apartado 2.2 los grafos bajo ciertas disciplinas son llamados redes. En este trabajo será indistinto llamar de una u otra forma a esta disciplina matemática.

Existen distintos índices para analizar redes, en los mismos juegan un papel determinantes el orden (n), y el tamaño (m) (Kansky *et al*; 1989)

2.3.1 Medidas de Conexión o Cohesión

El interés por esta clase de medidas, obedece a que el incremento de las conexiones a través del tiempo significa una mayor cohesión espacial para el territorio donde se emplaza la red. Las miden el grado de conectividad recíproca entre los nodos de la red. Los índices de este tipo son múltiples, los que se utilizaré en este trabajo se detallan a continuación:

2.3.1.1 Índice β o de Conexión Máxima: es el más simple, es el cociente entre el tamaño y orden del grafo. Cabe aclarar que estos índices se calculan sobre grafos simples. Es obvio que un aumento en el número de aristas en la red, representa una mayor conectividad entre los nodos.

$$\beta = \frac{m}{n} \quad \text{con} \quad 0 \leq \beta \leq \frac{n-1}{2}$$

En particular si $\beta < 1$, nos indicará una red inconexa.

Es interesante comparar éste índice con el valor máximo posible de aristas del grafo, que sería el tamaño del grafo completo, que se obtiene de acuerdo a la siguiente fórmula: Máximo de $\beta = n \cdot (n - 1) / 2$. Este valor indica la máxima conexión o coherencia que podría tener la red, y permite calcular el porcentaje de conectividad. El resultado debe interpretarse con cierta precaución, porque un valor de β igual a 100 % es totalmente ideal, y por lo tanto alejado de la realidad.

2.3.1.2 Índice μ o Número Ciclomático: expresa el número de ciclos que posee el grafo. Se calcula haciendo $\mu = m - (n - p)$ donde m es el número de aristas, n el número de vértices y p es el número de componentes conexas.

Dicho resultado adquiere mayor sentido al ser comparado con el número máximo posible de ciclos del grafo completo del mismo orden.

2.3.1.3 Índice α : es un indicador de la complejidad de la red o grafo y se obtiene al hacer cociente entre el Número Ciclomático y el máximo número posible de ciclos del grafo completo del mismo orden.

Si el grafo es no planar, $\alpha = \frac{\mu}{\frac{n(n-1)}{2} - (n-1)}$, donde el número máximo de ciclos se calcula haciendo la diferencia entre el número total de aristas del grafo completo de orden n y número de arista de un árbol del mismo orden.

Si el grafo es planar $\alpha = \frac{\mu}{2n-5}$, esto se obtiene a partir del siguiente: Teorema "Sea G un grafo planar maximal con n vértices ($n \geq 3$) y m aristas, entonces $m = 3.n - 6$ ", si a este valor le restamos el número de arista de un árbol de orden n se obtiene que el número máximo de ciclos igual a $(2n - 5)$.

2.3.1.4 Índice γ : consiste en hacer el cociente entre el tamaño m del grafo de orden n y el tamaño del grafo completo del mismo orden. Obteniéndose

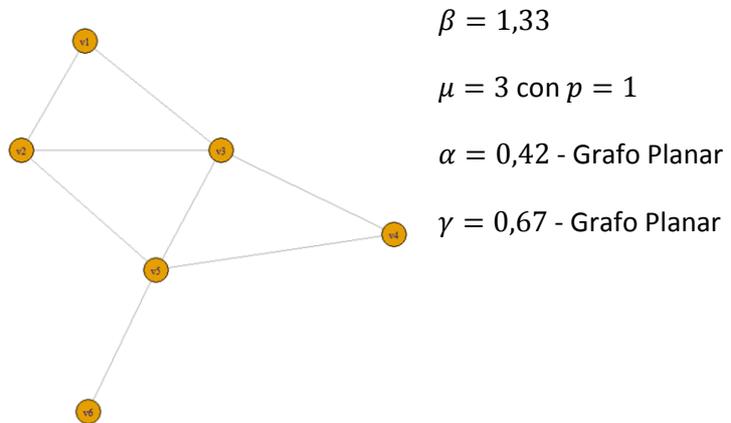
Si el grafo es no planar: $\gamma = \frac{m}{\frac{n(n-1)}{2}}$,

Si el grafo es planar: $\gamma = \frac{m}{3n-6}$

Al igual que el índice β , la cercanía al 1, muestra una red idealmente mejor conectada. Este índice puede expresarse en porcentajes y su interpretación nos indica que cantidad de aristas deberían introducirse en cada vértice para obtener una red más integrada.

El siguiente ejercicio, dado el siguiente grafo de la figura 4, se calculó los índices de conexión detallados anteriormente.

FIGURA 4. Grafo G no dirigido de 6 vértices y medidas de conexión asociadas.



2.3.2 Medidas de Accesibilidad y Centralidad.

En estas medidas los nodos adquieren una jerarquía en función de la facilidad de acceso desde cada uno a los demás, pues se analiza la accesibilidad topológica de cada uno de ellos en la red teniendo en cuenta la cantidad de aristas que es necesario atravesar para llegar a cada vértice de referencia desde cualquier otro vértice.

2.3.2.1 Medidas de centralidad

Las medidas de centralidad cuantifican la relevancia de los vértices y/o aristas de un grafo. En general esta relevancia se relaciona con la cercanía del vértice y/o aristas a los demás vértices y/o aristas. En particular el grado de un vértice es una medida de centralidad y posiblemente sea la más sencilla. Existen

numerosas medidas de centralidad, pero sólo se presentaran aquellas que sean de utilidad en este trabajo.

El diámetro, es una medida de centralidad que nos permite calcular el máximo de las distancias entre todos sus pares de vértice de un grafo G y se denota

$$diam = \text{máx } d_{u,v}(G)$$

El diámetro de un grafo es infinito si y solo si el grafo es no conexo. El $diam$ del grafo de la figura 4, es 3.

Para el cálculo de las demás medidas se requiera la construcción de la matriz de Adyacencia o de Conexión (A_{ij}). Sumando el número de conexiones en sentido horizontal, se detectan los nodos mejor y peor conectados.

Grado de Centralidad: El Grado de Centralidad es igual al número de vínculos que un nodo tiene con otros los nodos, sería el grado de cada uno de los nodos.

$$C_D(i) = \sum_{j=1}^n a_{ij} = \sum_{j=1}^n a_{ji}$$

Siendo los a_{ij} elementos de la matriz de adyacencia A y n el orden del grafo.

Grado de Centralidad Normalizado

$$C_D(i)^* = \frac{1}{n-1} \cdot C_D(i) \text{ siendo } n \text{ el orden del grafo.}$$

Nodos con mayor grado o más conexiones son más centrales a la estructura y tienden a tener una mayor capacidad de influir sobre los demás.

Centralidad de Intermediación: se basa en el número de caminos de menor longitud o geodésico entre cada par de vértice del grafo que pasan a través de un nodo. Los nodos con una alta intermediación juegan el rol de conectar a los diferentes grupos. En la siguiente fórmula, g_{jik} son todos los geodésicos ligando el nodo j y el nodo k , que pasan a través del nodo i ; g_{jk} es la distancia geodésica entre los nodos j y k .

$$C_B(i) = \sum_{j \neq i \neq k} \frac{g_{jik}}{g_{jk}}$$

Centralidad de Intermediación Normalizada

$$C_B(i)^* = \frac{2}{(n-1) \cdot (n-2)} \cdot C_B(i)$$

Los nodos con alta intermediación en una red, son los agentes y conectores que mantienen juntos a los demás. Los nodos con mayor intermediación al ser retirados incrementan la distancia entre los demás nodos.

Centralidad de Cercanía: la medida de centralidad más sofisticada es la cercanía que hace hincapié en la distancia de un nodo con respecto a todos los demás, centrándose en la distancia geodésica de cada nodo a todos los demás. La cercanía puede ser considerada como una medida de cuánto tiempo tomará para que la información se propague de un determinado nodo a los otros nodos de la red. En la siguiente ecuación, $C_c(i)$ es la centralidad de cercanía, y $d(i,j)$ es la distancia entre dos nodos de la red.

$$C_c(i) = \frac{1}{\sum_j d(i,j)}$$

Centralidad de Cercanía Normalizada

$$C_c(i)^* = \frac{n-1}{\sum_j d(i,j)}$$

Centralidad del Autovector: La centralidad del autovector es una medida de la influencia que tiene cada nodo en una red. Asigna puntuaciones relativas a todos los nodos de la red basadas en el principio bien conocido de que las conexiones a los nodos de alta puntuación contribuyen más a la puntuación del nodo en la cuestión que conexiones iguales a los nodos de baja puntuación (Umadevi, 2013; Noori, 2011). En general, las conexiones con nodos que son en sí mismos influyentes van a influenciar más a otro nodo que las conexiones a los nodos menos influyentes. Si denotamos la centralidad del nodo i por x_i , entonces podemos determinar este efecto haciendo x_i proporcional a la media de las centralidades de los vecinos de i :

$$x_i = \frac{1}{\lambda} \sum_{j=1}^n A_{ij} x_j$$

donde λ es un autovalor de la matriz A .atendiendo al concepto de autovalores y de autovectores de una matriz y definiendo el vector de las centralidades $x = (x_1 x_2 \dots)$, se puede reescribir esta ecuación de la siguiente manera.

$$\lambda x = Ax$$

Es obvio que x es el autovector asociado al autovalor λ de la matriz A .

Para que las centralidades sean no negativas el λ debe ser el mayor autovalor de la matriz de A . La centralidad del autovector definido de esta manera otorga a cada nodo una centralidad que depende tanto de la cantidad como de la calidad de sus conexiones: tener un gran número de conexiones es relevante, sin embargo un nodo con un menor número de contactos pero de alta calidad puede exceder a un nodo con un mayor número de contactos mediocres. La centralidad del autovector resulta ser en muchas situaciones una medida reveladora (Umadevi, 2013; Newman, 2008).

En la siguiente tabla se muestran a modo de ejemplo las medidas de centralidad para cada vértice del grafo de la figura 4.

TABLA 1. Medidas de Centralidad para Grafo G no dirigido de 6 vértices

Vértices	Grado de Centralidad	Grado de Centralidad Normalizado	Intermediación	Intermediación Normalizada	Cercanía	Cercanía Normalizada	Autovector
v1	2	0.4	0	0	0.111	0.556	0.610
v2	3	0.6	1	0.1	0.142	0.714	0.840
v3	4	0.8	2.5	0.25	0.167	0.834	1
v4	2	0.4	0	0	0.125	0.625	0.638
v5	4	0.8	4.5	0.45	0.167	0.834	0.924
v6	1	0.2	0	0	0.100	0.500	0.306

2.3.2.2 Medidas de Accesibilidad

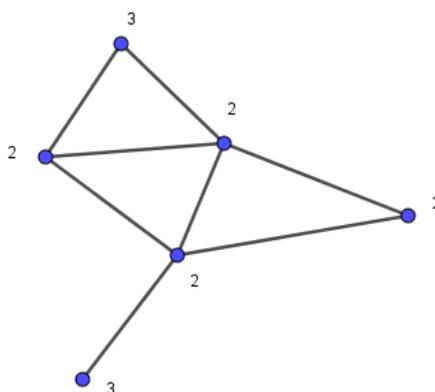
Para estas medidas es fundamental el uso de la Matriz de Accesibilidad Topológica (A_T), la misma se construye a partir de la matriz de adyacencia A , manteniendo los 1 en sus posiciones y reemplazando los 0 por las distancias entre los vértices de la red, expresada por el número de aristas que deben atravesarse para llegar de un vértice a otro, siguiendo el camino de menor longitud.

A partir de esta matriz se pueden calcular las siguientes medidas: el Número Asociado, y el índice de Shimbel, con el objetivo de reconocer y analizar la accesibilidad en los vértices del grafo. También de estos se pueden hallar otras medidas como grado de dispersión del grafo, su longitud media y la media de los recorridos.

Es importante aclarar que dicha accesibilidad también puede medirse considerando un grafo valuado, es decir un grafo en las que las aristas tengan un valor asociado, este valor puede ser cantidad de kilómetros, costos de traslados o bien tiempo de recorrido.

Número Asociado (NS) también llamado número de Köning mide la distancia topológica, la misma queda expresada por la longitud del camino de mayor longitud para alcanzar el nodo más distante por el camino de menor longitud. Representa la accesibilidad de este vértice al más alejado de la red, y significa que cuando más bajo es dicho número, más alto es el grado de accesibilidad. En la matriz de Accesibilidad Topológica se lo identifica porque es el número mayor de cada fila. En el ejemplo de la siguiente figura 5 se puede observar el NS en cada vértice del grafo.

FIGURA 5. Número Asociado de cada vértice un grafo G de 6 vértices



Índice o Número de Shimbel⁶(SHI): se obtiene sumando los valores de cada fila en la Matriz de Accesibilidad topológica y representa el número de aristas que

⁶ Alfonso Shimbel; "Structural Parameters of Communication Network"; Bulletin of Mathematical Biophysics; XV; 1953; 501-507

es necesario atravesar desde un nodo a los demás por el camino de menor longitud, siendo más accesible el nodo que represente el índice más bajo

$$SHI = \sum d_{u,v}(G)$$

En el caso que los vértices tengan igual NS, será más accesible el de menor índice de Shimmel.

Índice de accesibilidad topológica Ω .

A cada vértice se le puede calcular la accesibilidad topológica relativa (también llamado índice Ω) de la siguiente manera

$$\Omega_i = \frac{SHI_i - SHI_1}{SHI^1 - SHI_1} \cdot 100$$

Donde, Ω_i es el índice omega para el vértice i ; SHI_i es el Shimmel del vértice i ; SHI_1 es el Shimmel más bajo y SHI^1 es el Shimmel más alto. Los valores de Ω_i más bajos corresponden a vértices más accesibles, mientras que los más altos a los vértices menos accesibles.

Para el grafo de la figura 4, se muestra a modo de ejemplo la matriz de accesibilidad topológica, como también los índices SHI y Ω_i de cada vértice.

FIGURA 6. Grafo G no dirigido de 6 vértices y Matriz de Accesibilidad Topológica asociada.

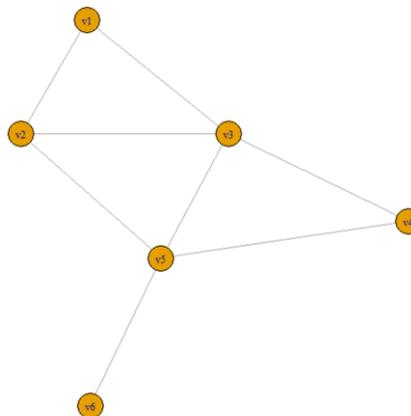


TABLA 2. Medidas de Accesibilidad para Grafo G no dirigido de 6 vértices

Vértices	SHI	Ω_i
v1	9	750
v2	7	550
v3	6	450
v4	8	650
v5	6	450
v6	10	850

Medias de los recorridos o longitud media, consiste en hacer el cociente entre la distancia topológica de cada nodo (NS) y el número total de nodos del grafo. Esta medida permite determinar el número medio de aristas que es preciso recorrer para llegar desde un nodo a cualquier otro de la red. A su vez permite determinar los nodos más accesibles de la red.

A partir de estas últimas medidas se derivan otros índices para evaluar la accesibilidad global de la red: índice G de Dispersión e Índice de Accesibilidad Media (IAM).

Índice G de Dispersión(IGD): mide el nivel de accesibilidad conjunta del grafo y se obtiene sumando todos los índices de Shimmel de todos los vértices del grafo.

$$IGD = \sum_i SHI_i$$

Este índice permite comparar la red con otras de similares características.

Índice de Accesibilidad Media (IAM): este índice determina un valor promedio de la accesibilidad en el grafo a partir del cociente entre, el índice G de dispersión y el número n de vértices.

$$IAM = \frac{IGD}{n}$$

Este promedio permite comparar diferentes grafos y también ver la accesibilidad de cada vértice (SHI) respecto a la media (IGD); los que superan este promedio son los menos accesibles. En el ejemplo del grafo de la figura 4, IGD es igual a 46 y IAM es 7,67.

2.4. Matriz de Vecinos W

El problema de la multidireccionalidad en el contexto espacial determina lo que se denomina dependencia espacial y puede resolverse mediante la definición de la matriz de vecinos o comúnmente llamada matriz de contactos o de pesos espaciales. Esta matriz W permite introducir la variable *dependencia entre las unidades* y, frecuentemente, es causa de disputas en relación a qué y cómo debería ser especificada (Herrera *et al.*, 2011), es decir W define el conjunto de vecinos para cada localización permitiendo “de esta manera”, incorporar el espacio dentro del análisis. Para cuantificar la estructura de dependencia espacial en un conjunto de datos, hay que definir para el conjunto de puntos o áreas, la relación espacial existente entre ellos Haining (1990). El trabajo de Ord (1975) realiza la conversión de estos términos dentro de un arreglo matricial y posiciona este elemento como clave para especificar modelos espaciales. Desde entonces, la matriz de pesos espaciales ha recibido considerable atención (Anselin, 2002), pero a pesar de ello no existen respuestas definitivas sobre cómo se deberían especificar las relaciones de dependencia espacial en ella.

Suponiendo que la superficie a estudiar está dividida en n regiones o áreas, la matriz W será de orden $n \times n$ y se representará de la siguiente forma:

$$W = \begin{bmatrix} 0 & w_{1,2} & \dots & w_{1,j} & \dots & w_{1,n} \\ w_{2,1} & 0 & \dots & w_{2,j} & \dots & w_{2,n} \\ \dots & \dots & \ddots & \dots & \dots & \dots \\ w_{i,1} & \dots & \dots & 0 & \dots & w_{i,n} \\ \dots & \dots & \dots & \dots & \ddots & \dots \\ w_{n,1} & w_{n,2} & \dots & \dots & \dots & 0 \end{bmatrix}$$

Donde los elementos w_{ij} ($i, j = 1, 2, 3 \dots n$) son los pesos o ponderaciones que representan la relación entre dos localizaciones i y j . Los elementos de la diagonal principal de esta matriz son iguales a cero, pues ninguna localización puede ser vecina consigo misma. En determinadas situaciones, suele añadirse un suprándice a la matriz para indicar el orden de contigüidad al que se refiere (en general, W^j representa al j -ésimo orden de contigüidad, siendo $j \geq 1$). Por ejemplo, $W^{(2)}$ identificará al conjunto de vecinos de los vecinos de cada localización o vecinos de segundo orden.

El criterio más común utilizado para determinar los pesos $w_{i,j}$ es el de cercanía o distancia geográfica, pero esto no es necesariamente válido para todas las aplicaciones como se pone de manifiesto mediante el conocido principio de

alotopía: «A menudo lo que sucede en una región se encuentra relacionado con otro fenómeno localizado en otras partes diferentes y remotas del espacio»(Ancot et al.1982) este principio destaca que la utilización de este tipo de criterio puede ser una posición muy optimista ya que existe una gran incertidumbre que caracteriza a su elección (Herrera et al., 2011).

El problema introducir la dependencia espacial es identificar cuáles son las regiones u observaciones vinculadas y cómo introducirlas al análisis, debe tenerse en cuenta que el espacio es irregular, heterogéneo y las influencias pueden ser de cualquier tipo. Por lo tanto la consideración de la cercanía geográfica como criterio de orden, tal como es mencionada por Tobler (1979), es solo una de las posibilidades.

Anselin (1988, p.16) designa como «la necesidad de determinar cuál de las otras unidades en el sistema espacial tiene influencia particular sobre lo unidad bajo consideración (...) expresada en nociones topológicas de vecindario o vecinos más cercanos», esto es reafirmado por Haining (1990). Este primer paso es absolutamente crucial, pero puede no ser tan simple y directo como escribir una matriz binaria o una matriz de pesos estandarizada por filas. En algunos casos, se puede tener información completa sobre la especificación de la matriz de ponderación. En otros casos, esta matriz será una mera hipótesis, como ocurre habitualmente entre los trabajos aplicados.

Desde una perspectiva univariante, la matriz de pesos proviene de un problema de subidentificación que afecta, en general, a la mayoría de los modelos espaciales. La dependencia espacial puede ser representada mediante un modelo espacial autoregresivo⁷ no restringido tal como:

$$\begin{aligned}
 y_1 &= y_2 w_{12} + y_3 w_{13} + \dots + y_j w_{1j} + \dots + y_n w_{1n} + \varepsilon_1 \\
 y_2 &= y_1 w_{21} + y_3 w_{23} + \dots + y_j w_{2j} + \dots + y_n w_{2n} + \varepsilon_2 \\
 &\vdots \\
 y_n &= y_1 w_{n1} + y_2 w_{n2} + \dots + y_j w_{nj} + \dots + \dots + y_{n-1} w_{1(n-1)} + \varepsilon_n \\
 \varepsilon_1, \dots, \varepsilon_j, \dots, \varepsilon_n &\sim N(0, \sigma^2)
 \end{aligned}$$

$$Y = WY + \varepsilon$$

⁷ Es importante destacar que la dependencia espacial puede estar presente en la parte sistemática de las ecuaciones, en el término de error o en ambos términos simultáneamente. Por simplicidad, nos enfocaremos en la especificación espacial autoregresiva como modelo básico, siendo extensibles los comentarios para las demás alternativas.

“donde es fácil ver su escasa utilidad práctica dado que esto resultará en un sistema con muchos más parámetros que observaciones. La solución al problema de la sobreparametrización, que surge cuando permitimos que cada relación de dependencia posea su propio parámetro, es imponer estructura sobre las relaciones de dependencia espacial” (Herrera *et al.*, 2011). Esta es la razón de por qué necesitamos una matriz de pesos espaciales.

La problemática de la especificación de la matriz parece más compleja, aunque la práctica usual ha favorecido al tipo de soluciones simples. El enfoque dominante implica un tratamiento exógeno del problema. Las unidades espaciales cercanas o vecinas son identificadas mediante una variable binaria usando, por ejemplo, el tradicional criterio de adyacencia o los k vecinos más cercanos, por mencionar los más comunes. Después de esto, la matriz binaria puede ser estandarizada en alguna forma habitual por filas, columnas o de acuerdo a la suma total. Otras matrices son construidas usando alguna variante de distancia entre los centroides de las unidades espaciales y con la posterior estandarización de la matriz. En algunas aplicaciones, la geografía puede ser sustituida por otro dominio, como la similitud en las estructuras socio-económicas en orden de obtener las medidas de distancia.

Profundizando el enfoque exógeno del problema hay varias alternativas para establecer el conjunto de vecinos de un área o polígono (unidades espaciales). Los criterios mencionados en la literatura, a semejanza de la localización de las piezas de un tablero de ajedrez y relacionadas con sus movimientos son, considerar vecinos a aquellas unidades espaciales que posean un borde común (arista). Otra posibilidad es considerar como vecinos a aquellas unidades que poseen un vértice común. Las elecciones de estos conjuntos de vecinos son denominadas, respectivamente, “criterio tipo torre” y “criterio tipo alfil” también pueden seleccionarse vecinos mediante una combinación de ambos criterios, dando lugar al “criterio tipo reina” considerando vecinos a que los comparten borde y vértice (Anselin, 1988).

Los polígonos o las áreas de un mapa pueden transformarse en puntos, y viceversa. La elección de un punto representativo de un polígono es típicamente resuelta por criterios geométricos, en particular mediante el punto central o centroide del polígono. Dado el centroide de cada unidad espacial, es posible definir una red de conexión de puntos o nodos que es sencillamente un grafo.

Un problema con la elección de vecinos por medio de la distancia (y contigüidad) es la existencia de puntos aislados que pueden no contener vecino alguno para un radio determinado. Esto sucede habitualmente cuando la densidad de los puntos sobre el plano no es regular o cuando se encuentran algunos nodos distribuidos por agrupamientos (*clusters*) y otros aislados. Para salvar este problema se suele determinar un radio d de amplitud mayor tal que asegure que cada observación tiene al menos un vecino.

Un criterio alternativo es el los k -vecinos más cercanos. En este caso, considerando la distancia geométrica entre las regiones, se seleccionan los k vecinos (kV) más cercanos de cada uno de los nodos que representan a las regiones.

Adicionalmente a los métodos anteriormente descritos este proyecto busca incorporar otra alternativa para la selección de vecinos, la Triangulación de Delaunay (también llamada grafo de Delaunay) Restringida (TDR).

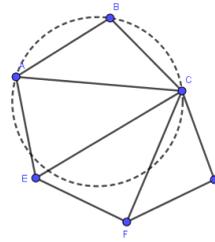
Esta triangulación es una de las triangulaciones más interesantes por ser aplicable a la resolución de una gran diversidad de problemas aparentemente sin relación entre sí, esto se debe a sus propiedades geométricas y al hecho de contar con algoritmos bastante eficientes para su cálculo (Delaunay, 1934).

Dado un conjunto de puntos, al hacer la Triangulación de Delaunay (TD) se obtiene un grafo conexo, planar y cordal. Este conjunto de triángulos cumple la siguiente condición (llamada condición de Delaunay): la circunferencia circunscripta⁸ de cada triángulo de la red no contiene ningún otro vértice del conjunto. Dado un conjunto de puntos es sencillo triangular, pero lo importante es determinar aquella triangulación que cumpla con la condición de Delaunay.

Para determinar los nodos o vértices vecinos, utilizó esta TDR pero se condicionó con radio de distancia d máxima para eliminar triangulaciones absurdas, es por eso que recibe el nombre de Triangulación de Delaunay Restringida (TDR). Se presenta a continuación, a modo de ejemplo una TD y se indica solo una de las cuatro circunferencias que cumplen con la condición

⁸La circunferencia circunscripta a un triángulo es aquella que contiene a los tres vértices del mismo, que por ser los tres vértices no alineados es única. El centro de esta circunferencia es el punto de intersección de las rectas mediatrices de los tres lados.

FIGURA 7. Triangulación de Delaunay en un grafo de 6 nodos.



Una vez establecido el grafo de vecindad o equivalentemente la lista de conjuntos de vecinos en el área de estudio, se debe proceder a asignar ponderaciones espaciales a cada relación. Existen diferentes criterios para la determinación de los pesos de la matriz W , siguiendo distintas hipótesis de interacción. Cada hipótesis puede conducir a matrices de ponderaciones diferentes. Cuando se desconoce sobre el proceso espacial asumido, es conveniente mantener una elección binaria, con $w_{ij} = 1$ cuando i y j son vecinos, y $w_{ij} = 0$ cuando no lo son (Bavaud, 1998). No obstante, también es posible construir una matriz de pesos no binaria, donde los elementos w_{ij} expresan el grado de interacción espacial potencial entre cada par de localizaciones. En efecto, como alternativa a los pesos binarios, pueden considerarse funciones que combinan la distancia, el perímetro y otras características geográficas de las unidades espaciales (González, 2015).

Es decir, la intensidad de la relación puede ser constante e igual, como es en el caso de la relación binaria (matriz de adyacencia), o no constante y no necesariamente iguales para cada de las unidades espaciales.

A continuación se presentan algunas estandarizaciones que se utilizan para ponderar la relación entre vecinos:

- Dos unidades espaciales son vecinas si se encuentran a menos de una determinada distancia fijada, siendo la intensidad de la relación inversamente proporcional a la distancia euclídea entre cada par de nodos que representan a cada unidades. La expresión de los elementos de la matriz de conexiones es la siguiente

$$w_{ij} \begin{cases} 0 & \text{si } d_{ij} > d \\ \frac{1}{d_{ij}} & \text{si } d_{ij} \leq d \end{cases}$$

Donde (i, j) denota las unidades espaciales vecinas (nodos o vértices), d indica la distancia máxima euclídea entre i y j para que dos unidades sean consideradas vecinas.

- Dos unidades espaciales son vecinas si se encuentran a menos de una distancia d fijada y la relación existente entre dos unidades o poblaciones será inversamente proporcional a la distancia entre ellas y directamente proporcional al cociente entre los tamaños de las mismas. La expresión de los elementos de la matriz de conexiones es la siguiente

$$w_{ij} \begin{cases} 0 & \text{si } d_{ij} > d \\ \frac{1}{d_{ij}} \cdot \frac{p_j}{p_i} & \text{si } d_{ij} \leq d \end{cases}$$

Donde p_i y p_j hacen referencia a la cantidad de habitantes de las unidades espaciales i y j respectivamente.

Es directo observar que la matriz W no será simétrica y que las unidades espaciales más pobladas ejercen más influencia que las menos pobladas dentro de una misma vecindad.

Existen otros criterios alternativos algunos de los cuales se presentan a continuación, en términos de los elementos de la matriz W

- $w_{ij} = \frac{1}{|x_i - x_j|}$ donde x indicaría alguna variable socioeconómica, por ejemplo el PBI per cápita (Case *et al.*, 1993).

$w_{ij} = [d_{ij}]^{-a}$ donde d_{ij} es la distancia entre las regiones i y j y a coeficiente de fricción positivo que se determina de acuerdo a la situación (Clif *et al.*, 1981).

- $w_{ij} = [d_{ij}]^{-2}$ donde d_{ij} es la distancia entre dos puntos o regiones i y j (Anselin, 1980).
- $w_{ij} = \gamma_{ij} \alpha_i \beta_{i(j)}$ donde γ_{ij} es un factor de contigüidad binario (1,0), α_i es la proporción del área de la unidad i respecto al total de área de todas las unidades del sistema y $\beta_{i(j)}$ es la proporción del perímetro de la unidad i en contacto con la unidad j (Dacey, 1969).

Sin embargo, es posible que alguna o algunas de estas especificaciones plantean problemas de endogeneidad que deberán ser considerados en el momento de la estimación del modelo. Este tipo de problema puede surgir a partir de propuestas que utilizan variables socio-económicas, como el nivel de empleo o el producto bruto *per cápita*, para la elección de los pesos (Gonzalez, 2015).

Como pudo observarse, es lógico la fuerza de las relaciones entre vecinos se atenúa con la distancia, tal como consideran Clif *et al.*, 1981, de manera que los pesos son proporcionales a la inversa de la distancia entre localizaciones. Asimismo, si se conoce que existen flujos entre áreas vecinas, que describen la estructura de los pesos espaciales, debe considerarse su utilización como pesos generales.

Una vez elegidos los pesos espaciales, lo habitual es trabajar con alguna transformación de la matriz en el modelo espacial, ya que mejora las propiedades estadísticas de los estimadores y sus estadísticos. En particular, Tiefelsdorf *et al.*,(2007), han considerado los siguientes sistemas de codificación para la matriz W :

- Normalización por fila: los elementos de la matriz W^* (obtenida a partir de la matriz W) se obtienen haciendo

$$w_{ij}^* = \frac{w_{ij}}{\sum_j w_{ij}}$$

De tal forma que la suma de cada fila de la matriz de W^* será igual la unidad.

- Estandarización global, que calculan los nuevos elementos de la matriz C como:

$$c_{ij} = \frac{nw_{ij}}{\sum_i \sum_j w_{ij}}$$

- Normalización de los pesos espaciales, donde los nuevos pesos se obtienen haciendo :

$$u_{ij} = \frac{w_{ij}}{\sum_i \sum_j w_{ij}}$$

De tal forma que la suma sea igual a la unidad: $\sum_i \sum_j u_{ij} = 1$

- Varianza de estandarización de los pesos espaciales, donde los nuevos valores se obtienen:

$$s_{ij} = \frac{n \cdot s_{ij}^*}{\sum_i \sum_j s_{ij}^*}$$

Donde $s_{ij}^* = \frac{w_{ij}}{\sqrt{\sum_j w_{ij}^2}}$

Es obvio que la elección de los criterios para definir vecinos y el esquema de codificación elegido para los pesos influye en los resultados y conclusiones obtenidas.

2.5. Autocorrelación y Heterogenidad Espacial

La variabilidad espacial está presente en cualquier investigación geográfica en particular y en cualquier ciencia vinculada al espacio en general como la epidemiología, geología, agricultura, entre otras (TISNÉS, 2012).

Fundamentalmente, los datos que tienen una localización espacial/geográfica asociada, pueden tener propiedades que se refieren a su localización individual y también a los datos que los rodean. Estas características son:

-Naturaleza georreferenciada: posición absoluta o relativa sobre el espacio con información valiosa a considerar.

-Multidireccional : relaciones que se establecen sobre el espacio.

-Multidimensional: no se distingue sobre el pasado, presente y futuro. Todo es presente, todo es pasado y todo es futuro.

Estas peculiaridades en el espacio dan lugar a los “Efectos Espaciales” como la Dependencia Espacial, la Autocorrelación Espacial y la Heterogeneidad Espacial.

La Autocorrelación Espacial (AE) es la concentración o dispersión de los valores de una variable en un mapa, es decir, refleja el grado en que objetos o actividades en una unidad geográfica son similares a otros objetos o actividades en unidades geográficas próximas (Goodchild, 1987). Este tipo de autocorrelación está asociada a la primera ley geográfica de Tobler (1970) que afirma: *todo está relacionado con todo lo demás, pero las cosas cercanas están más relacionadas que las distantes*. La Dependencia Espacial (DE) se produce cuando el valor de la variable dependiente en una unidad espacial es parcialmente función del valor de la misma variable en unidades vecinas como también de factores externos, de otras variables (Coro C, 2003), es decir, cuando la autocorrelación es sustantiva y no existe un factor de aleatoriedad. En el análisis de datos agregados geográficamente es frecuente encontrar que los valores de las variables estén autocorrelacionados espacialmente. La diferencia entre AE y DE está, fundamentalmente, en el uso de las palabras y estriba en que el primer caso se refiere simultáneamente a un fenómeno y técnica estadística, y el segundo, a una explicación teórica (Vilalta *et al.*,2005). Los

conceptos de DE y AE no son sinónimos, pero en este trabajo serán utilizados indistintamente.

Otro efecto espacial que hay que considerar es el concepto estadístico de la HE que indica la presencia de diferencias sistemáticas en la ocurrencia de un fenómeno en distintas regiones geográficas, de tal forma que este podría tener diferentes distribuciones (media, varianza u otros parámetros) en un subgrupo espacial de los datos o bien, simplemente cambiar con la ubicación de las unidades (O'Loughlin y Anselin 1992). La Heterogeneidad Espacial (HE) puede presentarse debido a:

- Simplemente un problema estadístico como consecuencia de la heteroscedasticidad en un modelo de regresión, en el término del error y viola el supuesto de la no correlación entre los errores. Si ignoramos este tipo de autocorrelación los estimadores de nuestra regresión serán ineficientes, nuestras estadísticas de t y F estarán sesgadas y la bondad de nuestro ajuste (R^2) será engañoso (O'Loughlin y Anselin 1992). Se considera como *ruido* porque no afectará los coeficientes del modelo y porque sólo nos interesa para mejorar nuestras estimaciones, o bien,
- A la existencia de una variación espacial sustantiva de la variable en cuestión, al igual a la DE, en este caso también estará presente el problema de la heteroscedasticidad, que se presenta cuando el valor que toma la variable dependiente en cada unidad geográfica está realmente determinado por el valor que ésta asume en las unidades vecinas. Es decir, la autocorrelación estará todavía presente después de otras variables explicativas. Si ignoramos este tipo de dependencia espacial los coeficientes estimados estarán sesgados. También puede suceder que se dé la combinación de DE y de HE y que no se pueda distinguir una de la otra para el tratamiento de los datos.

En el estudio de variables epidemiológicas es crucial la importancia del espacio, por lo tanto, el problema de la DE o AE, ha sido objeto de una gran cantidad de estudios en el área de las ciencias de la salud como Haining, 1990; Cressie, 1993; Elliot, 2000; Lawson, 2001; Cliff y Ord, 1981; Ripley, 1981 y Tiefelsdorf 2000 entre otros.

2.6. Índice de Autocorrelación Espacial: Coeficiente de Moran

Los estadísticos de autocorrelación espacial como el Coeficiente de Moran (1948) y el Cociente de Geary (1954) han permitido medir y analizar el grado de correlación espacial entre observaciones (Cliff y Ord, 1981). Cuando se han definido los pesos espaciales la autocorrelación espacial puede ser estimada. La significancia de los valores del Coeficiente de Moran obtenido puede ser probada a través del procedimiento de simulación de Monte-Carlo. En esta simulación las ubicaciones son permutadas para obtener una distribución del Coeficiente de Moran bajo la hipótesis nula de que hay una distribución aleatoria es decir, que no existe una correlación espacial entre las observaciones. Si el valor observado del Coeficiente de Moran es mayor que el esperado bajo simulación, entonces existe un agrupamiento o una estructura de los datos mediada por la distancia a la cual se encuentran valores más cercanos entre sí y son más parecidos (autocorrelación espacial positiva). Por el contrario, un valor significativo y negativo del Coeficiente de Moran indica que valores cercanos espacialmente son más diferentes entre sí que los esperados por azar (autocorrelación espacial negativa) (Bruno *et al.*, 2014).

El estadístico I del Coeficiente de Moran se calcula como:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n (y_i - \bar{y}) w_{ij} (y_j - \bar{y})}{S_o \sum_{i=1}^n (y_i - \bar{y})^2}$$

Dónde:

- w_{ij} denota los elementos de la matriz de pesos espaciales W correspondientes a las localidades i y j .
- y_i los valores que toma la variable Y de interés en la localización i .
- \bar{y} la media de la variable Y
- $S_o = \sum_{i=1}^n \sum_{j=1}^n w_{ij}$.

El valor del índice I oscila entre -1 y 1 y según Goodchild, (2008a) el I de Moran es esencialmente, el coeficiente de correlación de Pearson con una matriz de pesos definida por usuario, con grandes valores positivos indicando similitud entre vecinos y valores cercanos a cero indicando ausencia de autocorrelación espacial.

Su distribución asintótica es normal:

$$I \sim N(E(I); V(I))$$

Los momentos del estadístico de Moran, bajo la hipótesis nula de aleatoriedad, son:

$$E(I) = \frac{-1}{n-1}$$

$$V(I) = E(I^2) - E(I)^2$$

$$\text{Donde } E(I^2) = \frac{A-B}{C}$$

$$A = n[(n^2 - 3n + 3)S_1 - nS_2 + 3S_0^2]$$

$$B = D[(n^2 - n)S_1 - 2nS_2 + 6S_0^2]$$

$$C = (n-1) \cdot (n-2) \cdot (n-3)S_0^2$$

$$D = \frac{\sum_{i=1}^n (y_i - \bar{y})^4}{[\sum_{i=1}^n (y_i - \bar{y})^2]^2}$$

$$S_1 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (w_{ij} + w_{ji})^2$$

$$S_2 = \sum_{i=1}^n \left(\sum_{j=1}^n w_{ij} + \sum_{j=1}^n w_{ji} \right)^2$$

Para el cómputo del índice de autocorrelación espacial se puede utilizar indistintamente las definiciones de la matriz de pesos W , siendo habitual proceder previamente a la estandarización de la misma. Sin embargo, los resultados obtenidos puede variar, a veces significativamente, en función de la matriz W especificada (González, 2015).

2.7. Análisis de Componentes Principales espacial (ACPe)

El Análisis de Componentes Principales espacial (ACPe) permite estudiar las relaciones entre las variables medidas (análisis covariabilidad) y al mismo tiempo la estructura espacial (autocorrelación). El ACPe es una modificación del Análisis de Componentes Principales (ACP) en la que se construyen ejes (variables sintéticas) que no sólo optimizan la explicación de la varianza sino también la autocorrelación espacial permitiendo investigar la porción del ordenamiento de las observaciones relacionada a la Autocorrelación Espacial. El ACPe incorpora la restricción dada por los datos espaciales mediante el cálculo del índice de Moran para medir la dependencia o correlación espacial entre las observaciones en un sitio y el promedio de las observaciones en el vecindario de ese sitio. Para

delimitar los vecindarios utiliza una matriz de pesos espaciales W determinando cuáles y cuántas observaciones cercanas a cada sitio conforman el vecindario (Bruno *et al.*, 2014).

En resumen, el ACPe surge como un análisis que pretende integrar el análisis multivariado con datos georreferenciados, este análisis recibe el nombre Multispati-PCA. Esta integración se realiza a través de dos análisis, un primer análisis que resume la información multivariada como el ACP clásico y un segundo paso basado en una técnica univariada de estadística espacial que es aplicada sobre las variables sintéticas obtenidas a partir del ACP. Recordemos que el ACP es una técnica de reducción de dimensión que maximiza simultáneamente la varianza de la proyección de las observaciones sobre la componente principal y la suma de cuadrados de la correlación entre la componente principal y las variables (Bruno *et al.*, 2014).

CAPÍTULO 3

CÁLCULO DE TASAS Y ESTIMADORES DEL RIESGO DE LA POBLACIÓN.

Considerar el total de casos de una enfermedad o el número de muertes es útil para determinar la magnitud de un problema en salud pública, pero no resulta aplicable al problema de comparar grupos de población, como tampoco es útil al momento de comparar tendencias. Si se parte del mismo riesgo de desarrollar un evento, un grupo de población más numeroso desarrollará más eventos que uno pequeño, solo por su tamaño. Por esta razón, para comparar diferencias relativas entre grupos a lo largo del tiempo y, suponiendo que el riesgo es constante en todo el período, se debe hablar del riesgo de la población. Para cumplir con los objetivos de este trabajo, se calcularon una serie de indicadores de riesgo.

3.1. Tasas crudas mortalidad y tasas de mortalidad estandarizadas

El indicador más básico de mortalidad es *la tasa bruta o tasa cruda de mortalidad (TCM)*, que indica el número de muertes por 1.000 habitantes en una determinada población durante un período determinado de tiempo, por lo general un año. Sin embargo, cualquier interpretación de la estructura geográfica de los casos está limitada por la falta de información sobre la distribución espacial de la población que podría estar “en riesgo” y, consecuentemente, ha generado esos casos observados. Por tanto, se prefiere la utilización de tasas que permitan incorporar el efecto de la población en riesgo, en vez de considerar los casos brutos.

La *TCM* es una medida global que no tiene explícitamente en cuenta la composición de la población. Su valor será el promedio de los valores para los subgrupos individuales, ponderados por sus tamaños relativos. Cuanto más grande es el subgrupo, mayor influencia tendrá en la medida cruda. Por lo tanto, la tasa de mortalidad de una población es el promedio ponderado de las tasas de mortalidad para sus subgrupos componentes.

Supongamos, por ejemplo, una población de tamaño N formada por k grupos de edad o estratos y donde cada estrato de edad i ($1 \leq i \leq k$) con un número específico N_i de personas. En el siguiente año, cada estrato de edad i experimentará D_i muertes, siendo el tamaño poblacional total $N = \sum N_i$ y es el número total de muertes $D = \sum D_i$, entonces la tasa de mortalidad cruda será D/N , que también puede ser escrita como el promedio ponderado según las tasas de mortalidad específicas por edad, que hallan haciendo el cociente entre d_i y n_i , como sigue:

$$\frac{D}{N} = \frac{\sum d_i}{N} = \frac{\sum n_i \cdot \left(\frac{d_i}{n_i}\right)}{N} = \sum \frac{n_i}{N} \cdot \frac{d_i}{n_i} = \sum w_i \cdot \frac{d_i}{n_i}$$

Donde w_i es el peso del grupo etario i cabe señalar que $\sum w_i = \sum \frac{n_i}{N} = \frac{\sum n_i}{\sum n_i} = 1$.

La *TCM* es el primer indicador o la forma de resumen más sencilla y directa de la experiencia poblacional. Pero la mortalidad está fuertemente relacionada con la edad, de manera que las tasas de mortalidad específicas por edad serán muy distintas una de otra. El resumen que produce la *TCM* oculta la heterogeneidad de las tasas de mortalidad específicas por estrato, y no permite la comparación entre distintas zonas geográficas, ya que las diferencias observadas entre ellas pueden deberse a factores de riesgo que no hayan sido tenidos en cuenta, el sexo, el nivel socioeconómico u otros factores que determinan la magnitud de las tasas brutas y distorsionan su interpretación.

En Epidemiología, la mayoría de las tasas (incidencia, prevalencia y mortalidad) son fuertemente dependientes del grupo etario, ya se, en algunos casos, el riesgo de un evento (enfermar o morir) aumenta con la edad y, en otras situaciones, disminuye. La comparación de tasas crudas, a través del tiempo o entre poblaciones, puede ser engañosa si la composición por edades subyacente difiere entre las poblaciones que se comparan. Para superar este inconveniente es posible aplicar un proceso de normalización o ajuste, comúnmente denominado estandarización.

La mayor parte de los textos de epidemiología presentan el tema de estandarización de tasas en relación al ajuste por edad. Esta tendencia no es coincidencia, dado que virtualmente todos los eventos mórbidos o mortales ocurren con distinta frecuencia en los diferentes grupos de edad.

Los términos *ajuste* y *estandarización* se refieren a los procedimientos para facilitar la comparación de las medidas de resumen entre grupos. Dichas

comparaciones a menudo se complican por las diferencias entre los grupos en cuanto a factores que influyen en las medidas de interés pero que no son el foco de atención. El ajuste intenta quitar los efectos de factores *extraños* que pudiesen impedir una comparación justa. El término *Ajuste* desde el punto de vista más general, engloba tanto la estandarización como otros procedimientos para quitar los efectos de los factores que distorsionan una comparación. La estandarización se refiere a los métodos de ajuste basados en promedios ponderados, siendo seleccionados los pesos de manera de presentar una base *apropiada* para la comparación, es decir se define un estándar, pudiendo ser por ejemplo, el número de personas en varios estratos de una de las poblaciones en la comparación, una suma de las poblaciones, o alguna otra población externa relevante.

Una de las principales razones para estandarizar las tasas, se debe a que es más fácil comparar medidas de resumen de dos o más poblaciones que múltiples estratos de tasas específicas. Esto es sobretodo importante cuando se comparan tasas de varias poblaciones o cuando cada población tiene un número importante de estratos. Otra razón para estandarizar las tasas, es que números pequeños en algunos estratos pueden producir tasas específicas inestables. Cuando las poblaciones muestrales son tan pequeñas que sus estratos contienen fundamentalmente tasas inestables y ceros, el procedimiento de estandarización directa puede no ser apropiado.

A continuación se detallan los dos métodos de estandarización de tasas, pero en este trabajo se utilizará el método de estandarización indirecta porque hay pequeños números (de la variable observada) de las áreas analizadas.

3.2. Estandarización de tasas por el método directo (TME_D)

En la estandarización directa, las tasas específicas por estrato de la población de estudio se aplican a distribución por edad de una población estándar, es decir la estandarización directa aplica el mismo conjunto de pesos a las tasas específicas por edad. Este método no es más que un promedio ponderado de las tasas específicas por edad, para cada una de las poblaciones que se deseen comparar, en relación a las tasas de una población estándar.

Se cuenta con una región dividida en m áreas y se quiere hallar la tasa de mortalidad estandarizada por el método directo, para cada j –ésima área ($1 \leq j \leq m$). Para el cálculo de esta tasa es necesario determinar k estratos en

cada una de las j – ésima áreas. Cada i – ésimo estrato ($1 \leq i \leq k$) tendrá una tasa cruda de mortalidad, la que será denominada r_i .

La tasa de mortalidad estandarizada por el método directo para la j – ésima área será notada TME_{Dj} y está dada por:

$$TME_{Dj} = \frac{\sum(\text{tasas específicas de los estratos} \times \text{pesos})}{\sum(\text{pesos estándares})}$$

Siendo N_i : número de personas en el i – ésimo estrato de la población estándar

$$TME_{Dj} = \frac{(r_1N_1 + r_2N_2 + r_3N_3 + \dots + r_iN_i \dots + r_kN_k)}{(N_1 + N_2 + N_3 + \dots + N_k)}$$

Considerado $N = \sum_{i=1}^k N_i$

$$TME_{Dj} = \sum \left(r_i \times \frac{N_i}{N} \right)$$

Siendo $\frac{N_i}{N}$ el peso del estrato i -ésimo de la población estándar, denominado W_i

$$TME_{Dj} = \sum (r_i \times W_i)$$

Cuando los conteos específicos por estrato son pocos, como suele suceder en el caso de estratos etarios, las estimaciones de las tasas específicas por estrato a través del método directo pueden estar fuertemente influenciadas por variaciones aleatorias. En su lugar, se recomienda aplicar el procedimiento de estandarización indirecta que permite obtener tasas específicas por estrato, a partir de una población estándar de tamaño y relevancia suficiente.

3.3. Estandarización de tasas por el método indirecto (TME_I)

Estas tasas se promedian utilizando como ponderaciones los tamaños de cada estrato en población estudiada. Simbólicamente, la tasa de mortalidad estandarizada por el método indirecto para el área programática j – ésima (TME_I), está dada por la siguiente ecuación.

En la estandarización indirecta, las tasas se promedian utilizando como ponderaciones los tamaños de cada estrato en población estudiada. Este método no es más que un promedio ponderado de las tasas específicas por

edad, para cada una de las poblaciones que se deseen comparar, en relación a las tasas de una población estándar.

Se cuenta con una región dividida en m áreas y se quiere hallar la tasa de mortalidad estandarizada por el método indirecto, para cada j –ésima área ($1 \leq j \leq m$). Para el cálculo de esta tasa es necesario determinar k estratos en cada una de las j –ésima áreas. Cada i –ésimo estrato de la población estándar ($1 \leq i \leq k$) tendrá una tasa cruda de mortalidad, la que será denominada R_i .

La tasa de mortalidad estandarizada por el método indirecto para la j –ésima área será notada TME_{Ij} y está dada por:

$$TME_{Ij} = \frac{\sum \left(\begin{array}{l} \text{tasas específicas de los estratos de la pobla} \\ \text{ción estándar} \times \text{pesos de la población en estudio} \end{array} \right)}{\sum (\text{pesos de la población en estudio})}$$

Siendo n_i : número de personas en el i –ésimo estrato de la población de estudio.

$$TME_{Ij} = \frac{(R_1 n_1 + R_2 n_2 + R_3 n_3 + \dots + R_i n_i + \dots + R_k n_k)}{(n_1 + n_2 + n_3 + \dots + n_k)}$$

Considerado $n = \sum_{i=1}^k n_i$

$$TME_{Ij} = \sum \left(R_i \times \frac{n_i}{n} \right)$$

Siendo $\frac{n_i}{n}$ el peso del estrato i -ésimo de la población en estudio, denominado w .

$$TME_{Ij} = \sum (R_i \times w_i)$$

La estandarización indirecta evita el problema de las estimaciones imprecisas de las tasas específicas por estrato en una población de estudio al tomar las tasas específicas por estrato de una población estándar de suficiente tamaño y relevancia. Estas tasas son promediadas utilizando como pesos los tamaños de los estratos de la población de estudio. Así, el procedimiento es la imagen en espejo de la estandarización directa. En la estandarización directa, la población de estudio provee las tasas y la población estándar los pesos y en la estandarización indirecta, la población

estándar provee las tasas y la población de estudio los pesos, lo que se esquematiza en la siguiente tabla.

TABLA 3. Comparación de tasas según tipo de estandarización.

	Población de estudio	Población estándar
Tasa directamente estandarizada	Tasas	Pesos
Tasa indirectamente estandarizada	Pesos	Tasas

3.4. Razón de Mortalidad Estandarizada (RME) o Razón Estandarizada de Mortalidad (REM).

Una medida de la mortalidad que tiene en cuenta la estructura de edad es la Razón de Mortalidad Estandarizada (RME) o Razón Estandarizada de Mortalidad (REM). Esta medida se define como la tasa entre las muertes observadas en un área y las muertes esperadas si los riesgos para cada grupo de edad fueran los mismos que en cierta población de referencia. El uso de esta tasa se debe a la diferencia que existe entre las mortalidades en los diferentes grupos de etarios y a las diferentes estructuras de edad en las áreas geográficas.

En definitiva, la REM no es más que un estimador del riesgo relativo de mortalidad ajustado por edad en una determinada área, en particular REM de la j – ésima que se define como:

$$REM_j = \frac{o_j}{e_j}, \text{ con } 1 \leq j \leq m \text{ siendo } m \text{ la cantidad de áreas}$$

Donde o_j es el número de casos observados de la enfermedad en el área j – ésima, y e_j es el número de casos esperados que se calculan aplicando el método de estandarización indirecta, el que toma tasas específicas por estrato de edad de la población estándar.

Bajo el supuesto de que el número de muertes en los diferentes estratos de edad no están correlacionados y mientras los errores de muestreo asociados con la población estándar son insignificantes, se puede calcular el error estándar de j – ésima área a partir de la expresión:

$$D(REM_j) = \frac{\sqrt{o_j}}{e_j}$$

Es de interés determinar una gama de posibles valores para el verdadero REM, razonablemente consistentes con los datos observados. Los límites de confianza exactos para el REM se obtienen determinando, en primer lugar, los límites inferior y superior para la media $\mu = E(o_j)$ de la distribución de Poisson, correspondiente a los casos observados, que llamaremos L y U respectivamente.

Se asume en general que el modelo estadístico adecuado para los casos observados (o_j) para la j -ésima área es la distribución de Poisson ($o_j \sim \text{Poisson}(\mu_j)$) donde $\mu_j = \theta_j e_j$.

Luego, el límites inferior (LI) y el límite superior (LS) para el REM se calculan como L/e_j y U/e_j respectivamente (Haenszel, 1962).

3.5. RIESGO RELATIVO Y ODDS RATIO

Medir la frecuencia de enfermedades u otros estados de salud es solo el comienzo del proceso epidemiológico. La comparación de las frecuencias de enfermedades puede utilizarse para calcular el riesgo de que determinada exposición provoque cierto sobre la salud. Pueden establecerse comparaciones absolutas o relativas; las medidas resultantes describen la fuerza con la que se asocia una exposición a una determinada evolución.

3.5.1 Comparación absoluta: Diferencia de riesgos o proporciones

La diferencia de riesgos o proporciones, también llamada exceso de riesgo, es la diferencia entre la incidencia o prevalencia entre dos grupos, puede ser un grupo de expuestos y otro grupo de no expuestos. Es una medida útil para dimensionar la magnitud del problema en la salud pública que causa la exposición. Cuando se comparan dos o más grupos es importante que los mismos sean similares, excepto en aquello que se busca compara.

Muchos estudios comparan dos grupos (un grupo G1 con exposición y un grupo G2 sin exposición) en una respuesta binaria, Y (éxito y fracaso). Esto se muestran en una tabla de contingencia de 2×2 , ubicando los dos grupos en las filas y en las columnas los niveles de respuesta de Y . Se denotaremos π_1 y $1 - \pi_1$ la probabilidades de éxito y fracaso respectivamente del grupo G1,

análogamente, π_2 y $1 - \pi_2$ las probabilidades de éxito y fracaso respectivamente del grupo G2.

La diferencia de proporciones ($\pi_1 - \pi_2$) compara las probabilidades de éxito entre filas, toma valores entre -1 y 1, y es 0 cuando $\pi_1 = \pi_2$

Se ρ_1 y ρ_2 las proporciones de la muestra de cada una de las fila, la diferencia muestral $\rho_1 - \rho_2$ estima a $\pi_1 - \pi_2$. Cuando los valores de las filas corresponden a variables aleatorias con distribución binomial, se supone una probabilidad constante e independiente, la estimación del desvío estándar para $\rho_1 - \rho_2$ es

$$SE = \sqrt{\frac{\rho_1(1 - \rho_1)}{N_1} + \frac{\rho_2(1 - \rho_2)}{N_2}}$$

Donde N_1 y N_2 corresponden a los tamaños de los grupos.

El intervalo de confianza $100(1 - \alpha)\%$ para $\pi_1 - \pi_2$ esta dado por

$$(\rho_1 - \rho_2) \pm z_{\frac{\alpha}{2}} \cdot SE$$

3.5.2 Comparación Relativa: Riesgo Relativo

Es una medida de asociación, se usa principalmente en el análisis de trabajos prospectivos en los cuales dos grupos son seguidos para determinar la ocurrencia de algún evento. En cada grupo, el riesgo de ocurrencia de dicho evento se obtiene haciendo el cociente entre el número de casos y el resultado de interés por el número total de casos. La razón entre las proporciones en cada grupo es una medida del riesgo de un grupo comparado con el otro y se denomina **riesgo relativo (RR)**. Es decir el RR representa la probabilidad de que ocurra un evento, es la proporción entre los sujetos que presentan un evento y el total de sujetos susceptibles de presentarlo. Para poder calcular el RR hay que calcular previamente los riesgos en los dos grupos, y para eso hay que conocer la prevalencia o la incidencia acumuladas de la enfermedad por lo que esta medida suele utilizarse en los estudios de cohortes y en ensayos clínicos (Agresti, 1996), de la siguiente manera:

$$RR = \frac{\pi_1}{\pi_2}$$

Siendo π_1 y π_2 las probabilidades de éxitos de cada grupo.

El RR es mejor indicador de la intensidad de una asociación que la diferencia de riesgo, ya que se expresa en relación con un nivel basal de frecuencia, se relaciona con la magnitud de la tasa de incidencia o prevalencia basal, cosa que no ocurre en la diferencia de riesgos. En poblaciones en las que las diferencias

de riesgo son similares, los riesgos relativos pueden ser muy distintos, esto depende de la magnitud de las tasas basales.

Por otra parte, el concepto de *odds* diferente respecto al RR, la *odds* nos indica cuánto más probable es que se produzca un suceso a que no se produzca y se define como:

$$odds = \frac{\pi}{1 - \pi}$$

En una tabla de contingencia de 2x2, en la fila 1 (asociada al grupo 1) la probabilidad de éxito es $odds_1 = \frac{\pi_1}{1 - \pi_1}$, y en la fila 2 (asociado al grupo 2) corresponde $odds_2 = \frac{\pi_2}{1 - \pi_2}$ como probabilidad de éxito.

Se define *odds ratio* (OR) y se nota θ para dos filas o grupo como:

$$\theta = \frac{odds_1}{odds_2} = \frac{\pi_1 / (1 - \pi_1)}{\pi_2 / (1 - \pi_2)}$$

Observaciones:

- El valor de la OR es un número real positivo, es decir $\theta \in \mathcal{R}^+$.
- Cuando $\theta = 1$ es un valor de independencia y de referencia para la comparación, pues cuando los grupos X e Y son independientes, $\pi_1 = \pi_2$ entonces $odds_1 = odds_2$ resulta que $\theta = odds_1 / odds_2 = 1$.
- Si $\theta > 1$ las probabilidades de éxito de la fila 1 son mayores que las de la fila 2, es decir cuando $\pi_1 > \pi_2$.
- Cuando $\theta < 1$ es menos probable el éxito de la fila 1 que el de la fila 2, $\pi_1 < \pi_2$.

La OR no cambia el valor cuando en la tabla de contingencia se invierte las filas por columnas, tampoco cambia cuando se trata indistintamente las columnas como la variable de respuesta y las filas como variables explicativas, o viceversa. Por lo tanto, no es necesario identificar una clasificación como variable de respuesta para estimar θ . Cuando ambas variables son variables de respuesta, la OR se puede definir usando probabilidades conjuntas como:

$$\theta = \frac{\pi_{11} / \pi_{12}}{\pi_{21} / \pi_{22}} = \frac{\pi_{11} \pi_{22}}{\pi_{12} \pi_{21}}$$

También la OR recibe el nombre de cociente de producto cruzado. El estimador del θ es el OR muestral $\hat{\theta}$, que es igual a la razón de las probabilidades muestrales de las dos filas de la tabla de contingencia y se halla como sigue:

$$\hat{\theta} = \frac{\rho_1/(1-\rho_1)}{\rho_2/(1-\rho_2)} = \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \frac{n_{11}n_{22}}{n_{12}n_{21}}$$

Los OR son de uso frecuente en estudios retrospectivos, en estimaciones de la relación entre dos variables binarias, en investigaciones del efecto de otras variables sobre esa relación a través de la regresión logística y, en el análisis el resultado de estudios de casos y controles donde no es factible usar el riesgo relativo.

La distribución muestral de la razón de probabilidades es muy sesgada, al menos que el tamaño de la muestra sea extremadamente grande debido a esta asimetría, la inferencia estadística para la OR usa una medida equivalente alternativa, en particular el $\log(\theta)$.

El $\log(\hat{\theta})$ de muestra, tiene una distribución de muestreo menos sesgada y de forma de campana. Su distribución normal aproximada tiene una media de $\log(\theta)$ y un error estándar (SE) de:

$$SE = \sqrt{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}}$$

Como la distribución de muestreo está más cerca de la normalidad para $\log\theta$ que para θ , es mejor construir intervalos de confianza (IC) para $\log\theta$ y transformar hacia atrás utilizando la función inversa del logaritmo, que es la función exponencial para encontrar un intervalo de confianza para θ . El intervalo de confianza para $\log\theta$ esta dado por la siguiente formula:

$$IC = \log \theta \pm Z_{\alpha/2}(SE)$$

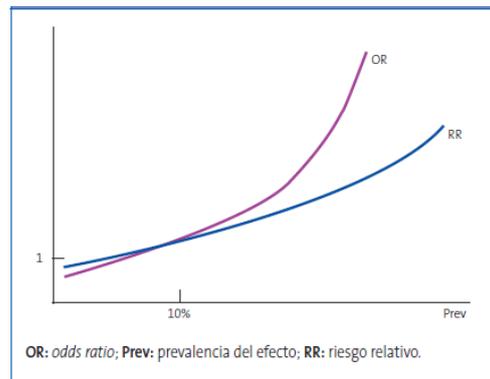
Pueden parecer similares el RR y la OR pero son conceptos muy diferentes, la utilización de cada uno de ellos depende, en gran medida, del tipo de diseño que seleccionado. El RR y la OR se comportan de forma diferente en función de la prevalencia del efecto que se esté observando. Los valores de la OR y el RR son muy similares cuando la prevalencia del efecto a estudiar es baja, esto se puede deducir de la fórmula de OR:

$$\theta = \frac{\rho_1/(1-\rho_1)}{\rho_2/(1-\rho_2)} = RR \times \left(\frac{1-\rho_2}{1-\rho_1} \right)$$

Se puede observar que cuando ρ_1 y ρ_2 tienden a cero se tiene que $\theta \cong RR$.

En bibliografía se encuentra que en eventos con tasa de prevalencia de alrededor de un 10 %, pueden considerarse similares la OR y el RR, esto aunque la OR siempre es un poco más baja que el RR para valores menores que 1 y un poco más alta para valores mayores que 1. En la figura 8 se muestra la relación entre la OR y el RR, se puede observar, a medida que la frecuencia del evento aumenta, la OR crece mucho más rápido que el RR.

FIGURA 8. Relación entre el Riesgo Relativo y Odds Ratio.



CAPITULO 4

DATOS Y VARIABLES

4.1. Datos y Fuentes de Información

Se trabajó con los datos de defunciones por Cáncer de Mama (CM) y estimaciones poblacionales obtenidas en la Dirección de Estadística de la Subsecretaria de Salud del Ministerio de Salud y Desarrollo Social de la Provincia de Neuquén. A partir de esta información, se calcularon, según el método indirecto, las tasas crudas y las tasas ajustadas por edad de mortalidad (por 100.000 habitantes), en las 28 áreas programáticas de la provincia para una serie temporal del 2001 al 2012.

Posteriormente se obtuvieron las Razones Estandarizadas de Mortalidad (REM) en las 28 áreas programáticas, estas razones son estimadores de los riesgos relativos, las mismas representan la relación entre el número observado y el número esperado de defunciones en cada área. El número esperado de defunciones se calculó por la estandarización por el método indirecto, donde las tasas específicas por edad de Argentina se aplicaron a la estructura por edad en las áreas programáticas de la provincia de Neuquén.

Se consultó el registro de pacientes con CM en la Provincia de Neuquén, registro que es elaborado por el servicio de Anatomía Patológica del Hospital Regional Castro Rendón, por ser el único hospital que cuenta con este servicio en el sector público y es el que analiza los resultados de toda la Provincia. El servicio de Anatomía Patológica cuenta con una base de datos, en salud pública, desde el año 1980 a la fecha, de los casos de CM de residentes. Para este trabajo se realizó un recorte de los casos de CM entre los años 2009 y 2015 ambos incluidos, con estos datos se calcularon las tasas de pacientes con CM, tasas de incidencia de nuevos casos en salud pública (Tasas NC) y se trabajó sobre la relación entre las variables asociadas a cada caso, grupo etario y localización.

Actualmente el estudio mamográfico es considerado el primer y único test para la detección temprana del CM, el mismo se realiza en pacientes sin síntomas a modo de "screening" (Novoa, 2009), por ello se incorporaron al trabajo, a partir

de la información proporcionada por el **Programa de Prevención de Tumores de la Subsecretaría de Salud** las siguientes variables:

- Cantidad de mamógrafos públicos y su localización.
- Tipo de mamógrafos.
- Capacidad del mamógrafo que se mide en cantidad de mamografías realizadas por hora.
- Población blanco, mujeres entre 50 y 70 años, es decir mujeres que deben realizarse la mamografía a modo “screening”, para el programa anual de prevención de tumores.
- Población de mujeres entre 50 y 70 años sin cobertura explícita de Salud.
- Cantidad de Establecimientos de Salud, también llamados efectores donde se solicitan las mamografías.
- Nivel de complejidad de los distintos tipos de efectores.

Estas variables se consideraron para poder realizar el cálculo de la accesibilidad.

Por otro lado es importante aclarar también que, en esta investigación se consideró un indicador socioeconómico por área programática (Lamfre *et al*, 2016), construido para el análisis de los determinantes sociales de la salud calculado a partir de la información brindada por la Dirección de Epidemiología de la Subsecretaría de Salud de Neuquén. El cálculo del indicador socioeconómico se hizo mediante el Análisis de Componentes Principales (ACP), y a partir de las variables:

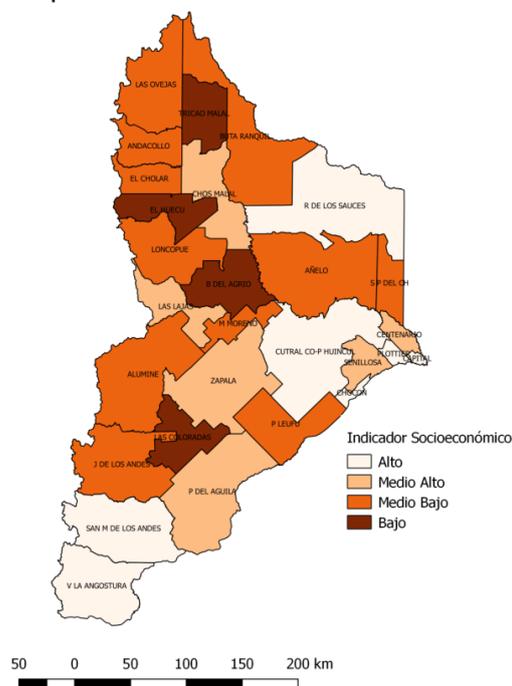
- Hogares cuyo Jefe de Hogar tiene primaria incompleta
- Hogares sin cobertura explícita en salud
- Hogares sin Red Pública de Agua
- Hogares con más de 3 personas por cuarto
- Hogares con CALMAT 4⁹
- Hogares sin gas de red para cocinar

⁹ La división por calidad de materiales de la vivienda (CALMAT) evalúa los materiales predominantes de los componentes constitutivos de la vivienda (pisos, paredes y techos) con relación a su solidez, resistencia y capacidad de aislamiento térmico, hidrófugo y sonoro. Se incluye asimismo la presencia de determinados detalles de terminación: cielorraso, revoque exterior y cubierta del piso.

- Hogares con Necesidades Básicas Insatisfechas (NBI)
- Hogares sin Baño.

A continuación se presenta el mapa de la provincia de Neuquén dividido en las áreas programáticas y se muestra el indicador socioeconómico (ISE) donde las categorías van desde “Alto”, “Medio alto”, “Medio bajo” y “Bajo”, “Alto” indica áreas con baja vulnerabilidad en condiciones socioeconómicas en contraposición de la categoría “Bajo” que corresponde a áreas con condiciones de más vulnerabilidad desde el punto de vista económico.

FIGURA 9. Indicador socioeconómico según área programática. Provincia de Neuquén. 2016



Fuente: Lamfre L, Hasdeu S. Construcción de un indicador de privación como marco de referencia para guiar intervenciones sanitarias en la provincia de Neuquén. *Rev Argent Salud Pública.* 2016; Sep;7(28):7-11

La información georeferenciada de Salud de la provincia de Neuquén se obtiene según el servicio de catálogo SIG¹⁰ (Sistema de Información Georeferenciada).

¹⁰<http://catalogo.neuquen.gov.ar/>

4.2. Metodología

En la metodología se describen cinco secciones en el cual se llevó a cabo este trabajo. La primera explica el método de cálculo utilizado para las Tasas de nuevos casos de incidencia de CM (Tasas CN) y Razón Estandarizada de Mortalidad (REM) para área programática de la provincia de Neuquén; en la segunda se detalla los métodos de contigüidad propuestos y pesos espaciales asignado a la estructura espacial utilizada en el cálculo de la Autocorrelación Espacial. En la tercera sección se determinan los grafos asignados a cada estructura espacial propuesta en la sección 2. Se muestra en la cuarta sección las variables que intervienen en el cálculo para la accesibilidad y forma de calcularla. Por último se realiza un análisis de componentes principales espacial integrando los resultados obtenidos en las secciones anteriores.

4.2.1 Cálculo de tasas y REMs

Para cada una de las 28 áreas programáticas de la provincia de Neuquén se calcularon, las tasas crudas y las tasas ajustadas por edad *TMEs*, según población proporcionada por la subsecretaría de Salud. Las tasas se expresaron en cantidad de defunciones por 100.000 personas en riesgo por año. Para el cálculo de las tasas de mortalidad se utilizó la planilla de cálculo de Excel con el método indirecto, utilizando las tasas específicas por edad de la población de Argentina y de la Provincia de Neuquén, según el atlas de mortalidad por CM de Argentina en periodo 2007-2011 (Arriaga et al.2011). También se utilizó un método de estandarización indirecto en el cálculo de las tasas esperadas para cada una de las áreas programáticas para luego calcular las REMs, que representan la relación entre el número observado y el número esperado de defunciones. Las tasas específicas por grupo etario para la Argentina y para la provincia de Neuquén se aplicaron a la estructura por edad de las áreas programáticas de la provincia. Las tasas calculadas se indican en la tabla 3.

TABLA 4. Tasas de mortalidad específicas por grupo de edad para Argentina y Provincia de Neuquén de Cáncer de Mama (por cada 100.000). Periodo 2007-2011.

Grupo Etario	Tasas por año de Prov. de Neuquén	Tasa por año de Argentina
0-4	0	0
5 a 9	0	0
10 a 14	0	0
15 a 19	0	0,1
20 a 24	0,8	0,2
25 a 29	0,9	0,9
30 a 34	0,9	3,8
35 a 39	8,9	8,4
40 a 44	14,4	16,2
45 a 49	27,9	27,1
50 a 54	43,1	40,1
55 a 59	74,6	56,2
60 a 64	74	72,3
65 a 69	93,2	82,7
70 a 74	102,5	93,8
75 a 79	110,1	124,1
80 y más	291,5	210,5
Total	21,7	26,3

Fuente: Atlas de mortalidad por Cáncer de Mama en Argentina 2007-2011.

Para estimar los intervalos de confianza (IC) se realizó una simulación de 1000 veces sobre los datos de pacientes con CM para el periodo 2001 al 2015. Para esta simulación se utilizó la varianza de los casos en el periodo especificado para cada área programática, considerando como la suma del total de casos sobre el número de registros menos 1, más un factor de corrección por finitud de 0,01 multiplicado por un número aleatorio entre 0 y 1.

4.2.2 Análisis de la Autorrelación Espacial

Una de las primeras cuestiones a resolver es determinar las estructuras de vecindades o equivalentemente la construcción de una matriz que represente la

relación de vecinos entre áreas. Como no existe una única teoría que guíe la construcción de dicha matriz de contacto o de conexión para las áreas programáticas de Neuquén, se consideraron diferentes criterios entre los habitualmente utilizados y también se propone un nuevo criterio, este último es una adaptación de un criterio que se suele utilizar.

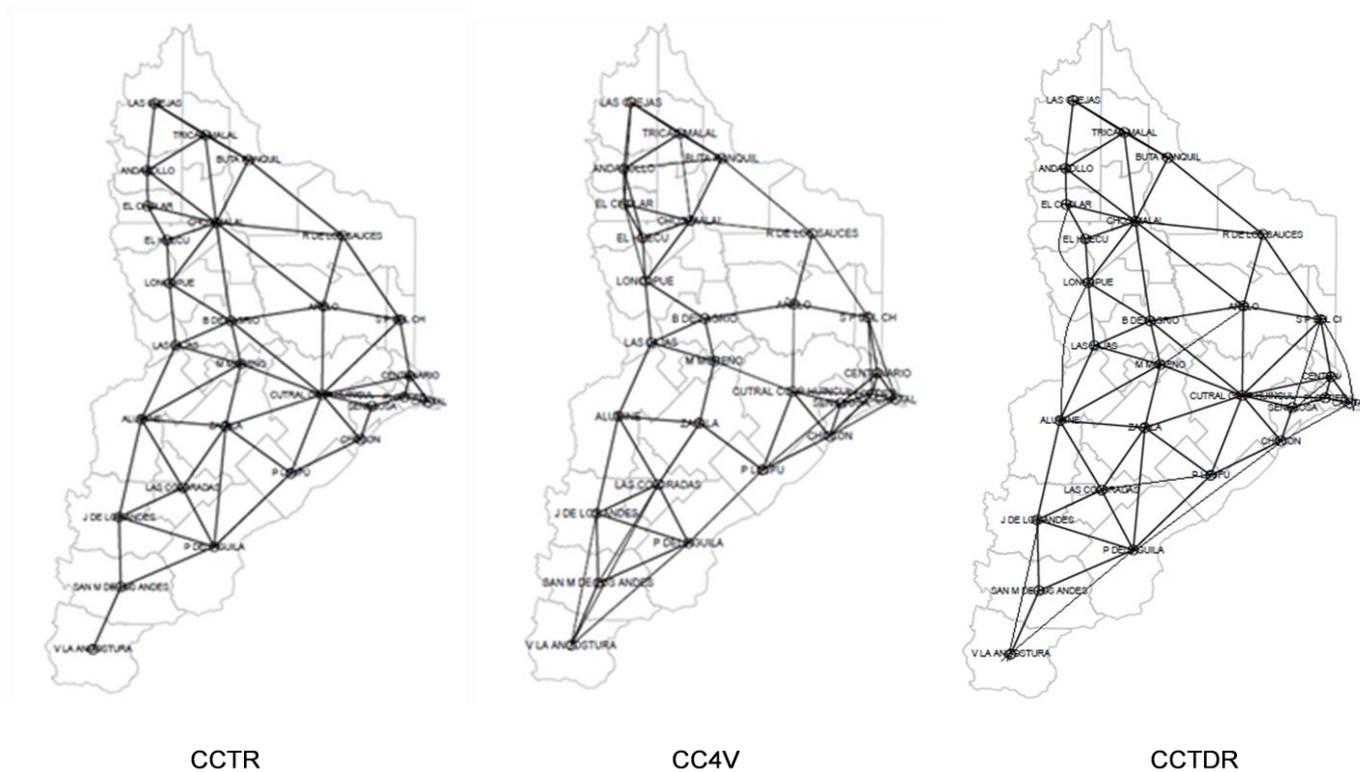
Se utilizó el software R por ser un software libre y de constante actualización. En particular el paquete **spdep** del mismo, para determinar las listas de vecinos basadas en distintos criterios de contigüidad, es importante aclarar que se trabajó la estructura de vecinos a partir de los centroides de cada área programática. A continuación se mencionan los criterios agregando la función del paquete **spdep** involucrada en cada caso:

- Criterio de Contigüidad Tipo Reina (CCTR), donde se utilizó la función **poly2nb**.
- Criterio de Contigüidad de k vecinos más cercanos (CCKV), donde se utilizó la función **knearneigh**, en particular se trabajó con k igual a 4.
- Criterio de Contigüidad de Triangulación de Delaunay Restringida (CCTDR), donde se utilizó la función **tri2nb**. Para este criterio se partió del criterio de contigüidad de la Triangulación de Delaunay (CCTD), restringiendo aquellas triangulaciones que desde lo geográfico no tienen sentido. Para determinar cuáles de las triangulaciones deben ser eliminadas se fijó una distancia máxima, que es igual a la mediana de las distancias medias de cada uno de los centroides de cada área programática con respecto a los centroides de las 27 restantes. La estructura de vecinos correspondiente CCTD también pueden ser obtenida con la función **tri2nb**.

Luego de determinar estos criterios de contigüidad se realiza un análisis multivariado de las variables analizadas teniendo en cuenta su correlación espacial que proporcionan estos criterios o redes.

A continuación, en la figura 10, se presentan las tres redes de conexión, mapas de contactos o grafos de vecinos obtenidos a partir de cada uno de los tres criterios mencionados CCTR, CCKV, CCTDR.

FIGURA 10. Redes de Conexión de la Provincia de Neuquén según criterio de contigüidad



Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

Se detalla en la Tabla 5 el número de vecinos de cada vértice o áreas en cada una de las redes o grafos.

TABLA 5. Número de vecinos de cada área programa según criterio de contigüidad.

Provincia de Neuquén

Área Programa	CCTDR	CCTR	CC4V
TRICAO MALAL	4	4	4
CHOS MALAL	9	9	6
BAJADA DEL AGRIO	5	6	4
CUTRAL CO-P HUINCUL	7	10	6
AÑELO	6	5	4
MARIANO MORENO	6	5	4
ZAPALA	6	6	4
ANDACOLLO	4	4	5
EL HUECU	3	3	4
LONCOPIE	6	4	5
LAS LAJAS	5	4	4
ALUMINE	6	5	4
VILLA LA ANGOSTURA	3	1	4
SAN MARTIN DE LOS ANDES	3	3	4
PIEDRA DEL AGUILA	6	5	5
LAS COLORADAS	5	4	6
JUNIN DE LOS ANDES	5	4	5
P LEUFU	5	4	5
CHOCON	6	3	5
SENILLOSA	5	3	6
PLOTTIER	4	4	6
CAPITAL	4	2	5
CENTENARIO	4	4	4
SAN PATRICIO DEL CHAÑAR	6	4	5
BUTA RANQUIL	3	4	5
LAS OVEJAS	2	3	4
EL CHOLAR	4	3	5
R DE LOS SAUCES	4	4	4
Suma	136	120	132
aristas	68	60	66
Nodos	28	28	28

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

Luego de seleccionar las redes de conexión se procede a determinar las ponderaciones espaciales w_{ij} , con la función ***nb2listw*** del software R, que toma un objeto de la lista de vecinos y lo convierte en un peso. Existen varios estilos de conversión como se indicó anteriormente, en el software se pueden obtener agregando a la función ***nb2listw***, *style* = "tipo de ponderación". En particular *style* = " W^* ", cuyo peso se obtiene estandarizando por fila, para que su suma se igual a uno. Para este caso, las ponderaciones en áreas con pocos vecinos serán más grandes que las que se originan en las áreas con muchos vecinos.

Si incorporamos el argumento *style* = "B", se genera una matriz de pesos binaria, asignando un 1 cuando dos observaciones son vecinas y 0 cuando no lo son. En este caso, las sumas de los pesos para las áreas difieren según el número de vecinos que tengan. También se utilizaron otros dos argumentos que asignan el mismo peso para todos los enlaces o vecinos: *style* = "C" que proporciona una matriz estandarizada globalmente y *style* = "U" que se estandariza los pesos de manera que la suma total sea igual a 1.

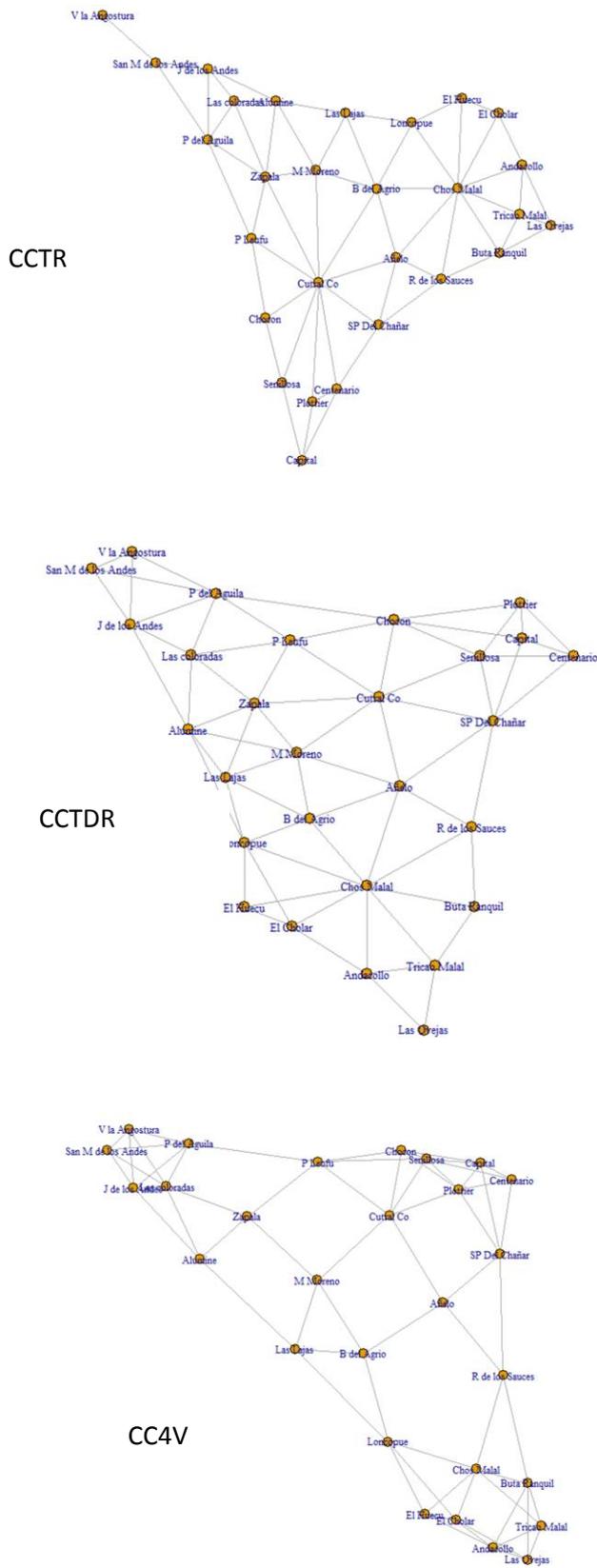
Por último se aplicó el argumento de estabilización de varianza propuesta por Tiefelsdorf *et al.*(2000) cuyo *style* = "S", este esquema comparado con W^* los pesos varían menos pero las sumas por filas de los pesos varían más (en W^* son siempre iguales a 1). Por otro lado registra una menor variabilidad en las sumas por filas con respecto a los estilos B, C y U.

4.2.3 Propiedades topológicas de los Grafos

Una vez definidas las estructuras de vecindades que serán utilizadas en la autocorrelación espacial se construyeron los grafos correspondientes a cada una de ellas, se utilizó el paquete del software R, ***igraph*** y para el cálculo de indicadores propuesto por Kansky¹¹ se trabajó con funciones: ***kansky_un, kansky_di, kansky_df*** (Beauguitte, 2014). Para el análisis de centralidad se utilizaron funciones: ***degree, closeness, betweenness, eigenvector*** de R.

¹¹Geógrafo estadounidense que, en su tesis impresa en 1963, propuso una serie de indicadores para caracterizar redes, indicadores derivados directamente de la teoría de grafos y han sido de gran utilidad en la geografía del transporte y todavía se usan ampliamente en la actualidad

FIGURA 11. Grafos según diferentes Redes de Conexión Provincia de Neuquén.

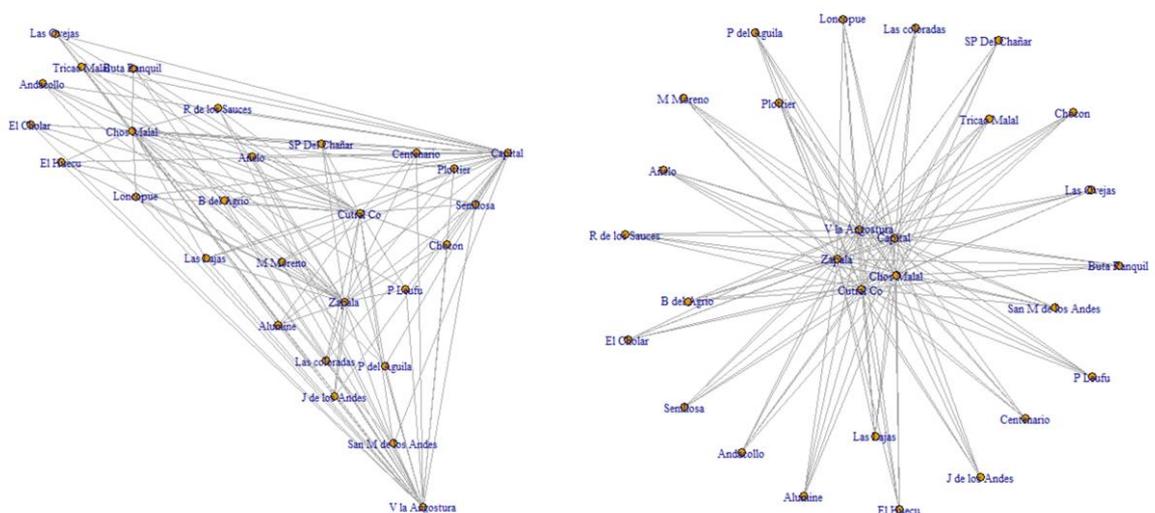


Fuente: Elaboración propia en base a información de la Subsecretaría de Salud
De la Provincia de Neuquén

4.2.4 Análisis de la Accesibilidad

Para analizar la funcionalidad de la red se consideró un grafo bipartito completo que modela la accesibilidad, los grafos bipartitos se caracterizan por tener dos conjuntos de vértices, en este caso en particular los grupos de vértices se conformaron en función si tienen o no las áreas programas mamógrafos públicos. En la figura 12 se observa este grafo desde dos perspectivas, en el primero se considera la ubicación geográfica de las áreas programáticas y el segundo no se tienen en cuenta la ubicación geográfica de las mismas.

FIGURA 12. Grafos bipartitos de accesibilidad según ubicación geográfica de las áreas programáticas. Provincia de Neuquén.



Fuente: Elaboración propia en base a información de la Subsecretaría de Salud De la provincia de Neuquén

Las variables que se consideraron para el cálculo de la accesibilidad son: capacidad teórica del mamógrafo para realizar mamografía de “screening” que se obtuvo realizando el cálculo pacientes por hora (pac./hs.), por horas de cada día (hs./día) por días de cada año (días/año). Cabe aclarar que también se tuvo en cuenta el tipo de mamógrafo de cada hospital, si es digital o analógico en el tiempo que lleva a realizar una mamografía. Otra variable que se consideró es la población blanco que considera el programa de prevención de tumores de la subsecretaría de Salud de la provincia, que son mujeres de 50 a 70 años que deben realizarse una mamografía cada dos años. También para calcular la accesibilidad se incluyó la distancia real (ruta y caminos reales) entre los hospitales cabecera de cada área programática.

La accesibilidad se calculó con la siguiente formula (Wang *et al.*,2013; Durand *et al.*,2013)

$$A_i = \sum_{j=1}^n \left[\frac{cap_j d_{i,j}^{-\beta}}{\sum_{k=1}^m p_k d_{k,j}^{-\beta}} \right]$$

Siendo A_i la accesibilidad de la ubicación i , cap_j capacidad teórica disponible del mamografo j , $d_{i,j}$ distancia real de viaje de i y j , p_k población blanco de la localidad k y β coeficiente de fricción, que en este estudio se lo consideró $\beta = 1$ (Anselin, 2002).

4.2.5 Análisis de Componentes Principales espacial (ACPe).

Para realizar un ACPe, se parte de una matriz $\tilde{X} = WX$ está compuesta por los promedios ponderados de los valores de los vecinos de acuerdo a la matriz de conexión espacial (W), como se indicó en el párrafo anterior, esta matriz es llamada matriz lagged. Las matrices X y \tilde{X} tienen las mismas columnas (variables) y las mismas filas (observaciones) y el análisis Multispati-PCA consiste en el análisis de este par de matrices (X y \tilde{X}) mediante un análisis de coinercia (Dra *et al.*, 2008). En particular en este trabajo la matriz X está conformada por 28 observaciones que corresponden a las áreas programa de la provincia y 4 variables analizadas: REMs, tasas de nuevos casos de CM, accesibilidad al mamografo e indicador socioeconómico. Para la matriz W se utilizó las tres redes de conexión propuestas: CCRT, CC4V y CCTRD con las distintas ponderaciones. Se empleó los paquetes **maptools**, **spdep**, **ade4**, **Guerry**, **spacemakeR** del software R y en particular la función **multispati** para el ACPe.

Se ajustó pruebas de autorrelación espacial, es decir se calculó el coeficiente de Moran y su significancia por simulación de Montecarlo para cada variable analizada con la función **moran.mc** del software R.

CAPÍTULO 5

RESULTADO Y DISCUSIÓN

Tras la implementación de la metodología antes descrita, se presentan los resultados junto a las interpretaciones y discusión de esta tesis, organizados en los siguientes ejes:

5.1. Distribución espacial de las Tasas de Mortalidad Estandarizadas por edad (por cada 100.000) de Cáncer de mama (CM). Provincia de Neuquén. Periodo 2001 a 2012

Se presentan a continuación, por área programa, las tasas de mortalidad estandarizadas por edad y su respectivos intervalos del 95% de confianza por cada 100.000 habitantes, para CM en la Provincia de Neuquén, periodo 2001 al 2012 (Tablas 6).

TABLA 6. Tasas de mortalidad estandarizadas por edad (TME), desvío estándar (DE (TME)) e intervalos de confianza (IC) para las tasas estandarizadas (por 100.000) de Cáncer de Mama. Provincia de Neuquén. Período 2001-2012.

AREA PROGRAMA	TME	DE(TME)	IC(95%) para TME	
			LI	LS
ALUMINE	21,2	0,01	21,1	21,2
ANDACOLLO	14,5	0,31	13,8	15,1
BUTA RANQUIL	7,4	0,01	7,4	7,4
CAPITAL	34,5	0,31	33,9	35,1
CENTENARIO	30,8	0,21	30,4	31,3
CHOS MALAL	27,0	0,24	26,5	27,4
CUTRAL CO-PLAZA HUINCUL	35,0	0,59	33,9	36,2
EL CHOCON	18,4	0,02	18,4	18,5
EL CHOLAR	65,9	0,45	65,0	66,8
JUNIN DE LOS ANDES	21,2	0,14	20,9	21,4
LAS LAJAS	14,9	0,01	14,9	14,9
LONCOPIE	3,1	0,00	3,1	3,1
MARIANO MORENO	7,0	0,01	7,0	7,0
PICUN LEUFU	10,7	0,01	10,7	10,8
PIEDRA DEL AGUILA	26,5	0,01	26,5	26,6
PLOTTIER	40,0	0,69	38,6	41,4
RINCON DE LOS SAUCES	18,6	0,01	18,6	18,6
SAN MARTIN DE LOS ANDES	10,9	0,11	10,7	11,1
SAN PATRICIO DEL CHAÑAR	27,6	0,12	27,4	27,9
SENILLOSA	24,7	0,01	24,7	24,7
TRICAO MALAL	11,4	0,02	11,4	11,4
VILLA LA ANGOSTURA	8,5	0,08	8,4	8,7
ZAPALA	25,5	0,27	25,0	26,1

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

**5.2. Distribución espacial de las REM de Cáncer de Mama.
Provincia de Neuquén. Periodo 2001 a 2012.**

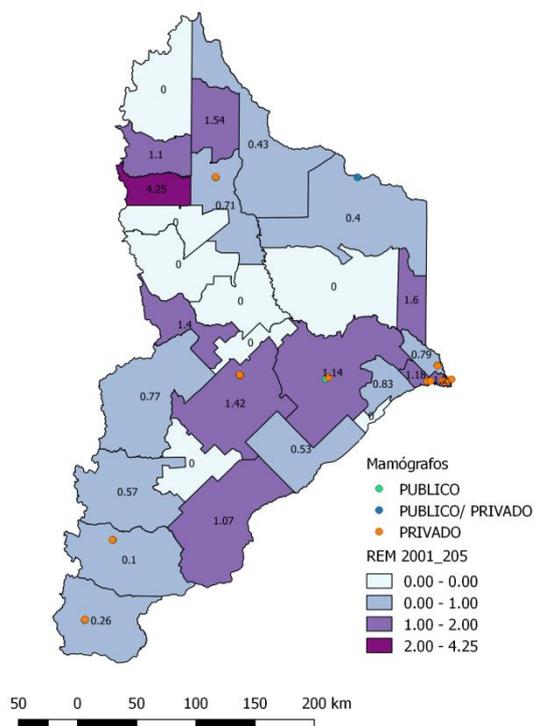
TABLA 7. Casos observados (O_i), casos esperados (e_i), Razón Estandarizada de Mortalidad (REM) e intervalo de confianza (IC) para los Razones Estandarizadas para el Cáncer de Mama. Provincia de Neuquén. Período 2001-2012.

AREA PROGRAMA	CASOS OBSERVADOS O_i	CASOS ESPERADOS e_i	REM	IC(95%) para REM	
				LI	LS
ALUMINE	6	3	1,86	0,61	3,10
ANDACOLLO	3	2	1,33	0,07	2,59
BUTA RANQUIL	2	3	0,65	-0,11	1,41
CAPITAL	326	108	3,03	2,75	3,30
CENTENARIO	41	15	2,70	2,01	3,40
CHOS MALAL	15	6	2,33	1,34	3,31
CUTRAL CO-PLAZA HUINCUL	67	21	3,12	2,49	3,75
EL CHOCON	1	1	1,61	-1,04	4,27
EL CHOLAR	5	1	5,70	1,51	9,90
JUNIN DE LOS ANDES	15	8	1,88	1,08	2,68
LAS LAJAS	3	2	1,31	0,07	2,55
LONCOPIE	1	4	0,27	-0,17	0,71
MARIANO MORENO	1	2	0,61	-0,40	1,62
PICUN LEUFU	2	2	0,94	-0,15	2,04
PIEDRA DEL AGUILA	5	2	2,33	0,62	4,04
PLOTTIER	43	12	3,52	2,64	4,41
RINCON DE LOS SAUCES	5	3	1,63	0,43	2,84
SAN MARTIN DE LOS ANDES	12	13	0,96	0,50	1,41
SAN PATRICIO DEL CHAÑAR	7	3	2,42	0,91	3,92
SENILLOSA	6	3	2,17	0,71	3,63
TRICAO MALAL	1	1	1,00	-0,65	2,65
VILLA LA ANGOSTURA	4	5	0,76	0,14	1,39
ZAPALA	32	14	2,26	1,60	2,91

Fuente: Elaboración propia en base a información de la Subsecretaria de Salud de la provincia de Neuquén

FIGURA 13. Mapeo de los Razón Estandarizada de Mortalidad estandarizados.

Provincia de Neuquén. Periodo 2001-2012



Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

5.3. Análisis descriptivo de los Nuevos Cáncer de Mama en Salud Pública periodo 2001-2015

TABLA 8. Nuevos Casos de Cáncer de Mama en Salud Pública según edad . Provincia de Neuquén. Período 2001-2015.

LI	LS	FA	FR	FAA	FRA
[18,00	28,86)	8	0,01	8	0,01
[28,86	39,71)	76	0,08	84	0,09
[39,71	50,57)	258	0,27	342	0,36
[50,57	61,43)	291	0,3	633	0,66
[61,43	72,29)	200	0,21	833	0,87
[72,29	83,14)	100	0,1	933	0,97
[83,14	94,00]	24	0,03	957	1

Nota: LI: Límite inferior de edad, LS: Límite superior de edad, FA: frecuencia absoluta simple, FR: frecuencia relativa simple, FAA: frecuencia absoluta acumulada, FRA: frecuencia relativa acumulada.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

El CM se trata de una enfermedad que afecta mayoritariamente a mujeres entre 45 y 70 años de edad, sobre la cual se desconocen cuáles son los causantes aunque se reconocen algunos factores que aumentan el riesgo de enfermar, con respecto a la edad, ser mayor de 50 años es un factor de riesgo (INC¹²).

En el caso de Neuquén en los casos registrados en salud pública el 64% de las mujeres con CM tienen más de 50 años como se observa en la tabla 8.

TABLA 9. Nuevos Casos de Cáncer de Mama en Salud Pública según factor de riesgo. Provincia de Neuquén. Período 2001-2015.

Factor de Riesgo	Condición frente al CM		Total
	CM	SCM	
Mayor o igual a 50 años	615	73008	73623
Menor a 50 años	342	143882	144224
Total	957	216890	217847

Nota: Se consideró la población del año 2009 como población central en los años considerados, sin tener en cuenta las edades inferiores a 15 años. CM: con cáncer de mama, SCM: sin cáncer de mama.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

Se detectó asociación significativa entre el factor de riesgo de la edad y condición frente al CM con $p < 0,0001$, es decir existen evidencias muestrales para rechazar la hipótesis nula de independencia entre el factor de riesgo y condición frente al CM. El odds ratio calculado de ser mayor de 50 años respecto a la condición “si tienen CM” es de 3,54 con un intervalo de confianza (IC) de (3,10; 4,05), el IC no contiene al 1 y el riesgo relativo es 3,52 con un IC de (3,09; 4,02), hay un efecto dañino. Es decir que las mujeres mayores de 50 años tienen 3 veces más chance de registrar CM, respecto de las mujeres menores de 50 años en la provincia de Neuquén.

En el periodo evaluado, hay áreas programas que registran 1 nuevo caso de CM, lo que dificulta calcular odds ratio por área programa, por ello se agruparon los datos por zonas sanitarias de la provincia. A continuación

¹²Instituto Nacional del Cáncer (INC) ,<http://www.msal.gov.ar/inc>

se presenta las tablas de contingencias 2 × 2 de las 6 zonas sanitarias en que se divide la provincia, incluyendo el hospital Castro Rendón en la Zona Metro.

TABLA 10. Nuevos Casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria I. Provincia de Neuquén. Período 2001-2015.

Factor de Riesgo	Condición frente al CM		Total
	CM	SCM	
Mayor o igual a 50 años	51	3640	3691
Menor a 50 años	19	13785	13804
Total	70	17425	17495

Nota: La zona Sanitaria I está compuesta por las áreas programas: Centenario, San Patricio del Chañar y Senillosa. CM: con cáncer de mama, SCM: sin cáncer de mama.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

TABLA 11. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria II. Provincia de Neuquén. Período 2001-2015.

Factor de Riesgo	Condición frente al CM		Total
	CM	SCM	
Mayor o igual a 50 años	47	4579	4626
Menor a 50 años	27	16342	16369
Total	74	20921	20995

Nota: La zona Sanitaria II está compuesta por las áreas programas: Aluminé, Bajada del Agrio, El Huecú, Las Lajas, Loncopue, Mariano Moreno, Zapala. CM: con cáncer de mama, SCM: sin cáncer de mama.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

TABLA 12. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria III. Provincia de Neuquén. Período 2001-2015.

Factor de Riesgo	Condición frente al CM		Total
	CM	SCM	
Mayor o igual a 50 años	33	2442	2475
Menor a 50 años	21	8418	8439
Total	54	10860	10914

Nota: La zona Sanitaria III está compuesta por las áreas programas: Chos Malal, Buta Ranquil, Tricao Malal, Andacollo. CM: con cáncer de mama, SCM: sin cáncer de mama.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

TABLA 13. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria IV. Provincia de Neuquén. Período 2001-2015.

Factor de Riesgo	Condición frente al CM		Total
	CM	SCM	
Mayor o igual a 50 años	41	4955	4996
Menor a 50 años	26	16229	16255
Total	67	21184	21251

Nota: La zona Sanitaria IV está compuesta por las áreas programas: Junín de los Andes, Las Coloradas, San Martín de los Andes, Villa la Angostura. CM: con cáncer de mama, SCM: sin cáncer de mama.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

TABLA 14. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Sanitaria V. Provincia de Neuquén. Período 2001-2015.

Factor de Riesgo	Condición frente al CM		Total
	CM	SCM	
Mayor o igual a 50 años	65	4837	4902
Menor a 50 años	33	20315	20348
Total	98	25152	25250

Nota: La zona Sanitaria V está compuesta por las áreas programas: Añelo, Cutral Có, Rincón de los Sauces, Picún Leufú, Piedra del Águila. CM: con cáncer de mama, SCM: sin cáncer de mama.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

TABLA 15. Nuevos casos de Cáncer de Mama en Salud Pública según factor de riesgo en Zona Metropolitana. Provincia de Neuquén. Período 2001-2015.

Factor de Riesgo	Condición frente al CM		Total
	CM	SCM	
Mayor o igual a 50 años	365	20979	21344
Menor a 50 años	178	67516	67694
Total	543	88495	89038

Nota: La zona Metropolitana está compuesta por las áreas programas: Hospital Castro Rendón, Neuquén Capital y Plottier. CM: con cáncer de mama, SCM: sin cáncer de mama.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

En las seis zonas Sanitarias se detectó asociación significativa entre el factor de riesgo de la edad y condición frente al CM con $p < 0,0001$, es decir existen evidencias muestrales para rechazar la hipótesis nula de independencia entre el factor de riesgo y condición frente al CM.

En la Zona Sanitaria I, el odds ratio calculado de ser mayor de 50 años respecto a la condición “si tienen CM” es de 10,17 con un intervalo de confianza (IC) de (6,03; 17,14), el IC no contiene al 1 y el riesgo relativo es 10,04 con un IC de (5,84; 16,07). Es decir que las mujeres mayores de 50 años tienen 10 veces más chance de registrar CM, respecto de las mujeres menores de 50 años.

En la Zona Sanitaria II, el odds ratio calculado de ser mayor de 50 años respecto a la condición “si tienen CM” es de 6,21 con un intervalo de confianza (IC) de (3,88; 9,95), el IC no contiene al 1 y el riesgo relativo es 6,16 con un IC de (3,81; 9,8). Es decir que las mujeres mayores de 50 años tienen 6 veces más chance de registrar CM, respecto de las mujeres menores de 50 años.

En la Zona Sanitaria III, el odds ratio calculado de ser mayor de 50 años respecto a la condición “si tienen CM” es de 5,42 con un intervalo de confianza (IC) de (3,15; 9,33), el IC no contiene al 1 y el riesgo relativo es 5,36 con un IC de (3,08; 9,16). Es decir que las mujeres mayores de 50 años tienen 5 veces más chance de registrar CM, respecto de las mujeres menores de 50 años.

En la Zona Sanitaria IV, el odds ratio calculado de ser mayor de 50 años respecto a la condición “si tienen CM” es de 5,16 con un intervalo de confianza (IC) de (3,17; 8,42), el IC no contiene al 1 y el riesgo relativo es 5,13 con un IC de (3,12; 8,32). Es decir que las mujeres mayores de 50 años tienen 5 veces más chance de registrar CM, respecto de las mujeres menores de 50 años.

En la Zona Sanitaria V, el odds ratio calculado de ser mayor de 50 años respecto a la condición “si tienen CM” es de 8,27 con un intervalo de confianza (IC) de (5,45; 12,56), el IC no contiene al 1 y el riesgo relativo es 8,18 con un IC de (5,34; 12,32). Es decir que las mujeres mayores de 50 años tienen 8 veces más chance de registrar CM, respecto de las mujeres menores de 50 años.

En la Zona Metropolitana, el odds ratio calculado de ser mayor de 50 años respecto a la condición “si tienen CM” es de 6,6 con un intervalo de confianza (IC) de (5,51; 7,9), el IC no contiene al 1 y el riesgo relativo es 6,5 con un IC de

(5,43; 7,76). Es decir que las mujeres mayores de 50 años tienen 6 veces más chance de registrar CM, respecto de las mujeres menores de 50 años.

TABLA 16. Nuevos casos de Cáncer de Mama en Salud Pública según año. Provincia de Neuquén. Período 2001-2015.

Año	FA	FR	FAA	FRA
2001	35	0,04	35	0,04
2002	59	0,06	94	0,1
2003	68	0,07	162	0,17
2004	65	0,07	227	0,24
2005	63	0,07	290	0,3
2006	62	0,06	352	0,37
2007	43	0,04	395	0,41
2008	38	0,04	433	0,45
2009	111	0,12	544	0,57
2010	58	0,06	602	0,63
2011	54	0,06	656	0,68
2012	58	0,06	714	0,74
2013	88	0,09	802	0,83
2014	86	0,09	888	0,92
2015	73	0,08	961	1

Nota: LI: Límite inferior de edad, LS: Límite superior de edad, FA: frecuencia absoluta simple, FR: frecuencia relativa simple, FAA: frecuencia absoluta acumulada, FRA: frecuencia relativa acumulada.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

En el año 2009 se crea en el ministerio de salud de la provincia un programa exclusivo denominado “**Red Provincial de Patología Mamaria**” (Novoa, 2009), en ese año se registraron 111 casos nuevos de CM en la provincia. Tomando el año 2009 como referencia, el 45 % de los casos nuevos de CM se registraron hasta el año 2008 y 55% de los casos nuevos de CM desde el 2009 en adelante.

5.4. Comparación de Grafos (redes de conexión)

Para describir los grafos de las estructuras de vecindades propuestas, se calculó índices globales descriptos en el marco conceptual.

TABLA 17. Índices globales según estructura de vecindades. Provincia de Neuquén

Índices Globales	CCTR	CCTDR	CC4V
Diámetro	8	6	6
N° Ciclomático(μ)	33	39	41
Alpha (α)	0,54	0,63	0,67
Beta(β)	2,14	2,35	2,42
Gamma(γ)	0,7692	0,8461	0,8717
Grado de Conectividad	6,3	5,72	5,55
Índice de G	2222	2246	2052
IAM	79,35	80,21	73,28

Fuente: Elaboración propia

En las redes analizadas se observa que CCTR es la que presenta mayor distancias entre todos sus pares de vértice, pues su diámetro es 8, en cambio CCTDR y CC4V su distancia máxima entre todos su pares de vértices es de 6, siendo redes más centrales.

En términos de complejidad las redes CC4V, CCTDR y CCTR se ordenan en forma decreciente respectivamente, en función del cálculo del indicador α .

En cuanto a cercanía la red CC4V ($\gamma = 0,8717$) es la mejor conectada, seguida por la CCTDR ($\gamma = 0,8461$) y por último CCTR ($\gamma = 0,7692$).

Ambas redes cuentan con un indicador $\beta \geq 1$ que nos dicen que son redes conexas, con mayor conexión entre nodos o vértices.

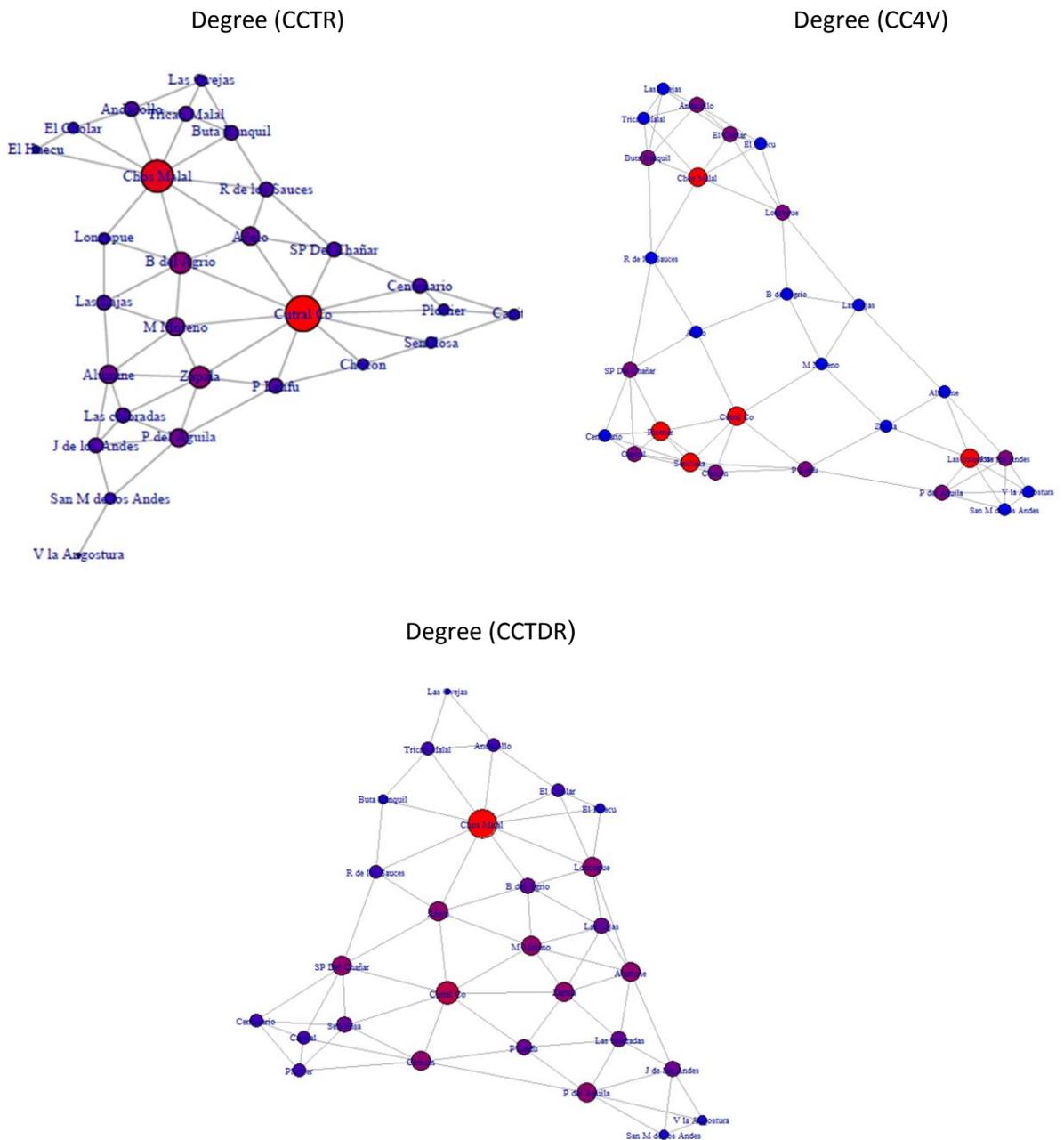
Con respecto al grado de centralidad, la red CCTR presenta nodos con 6 vínculos, en cambio en las redes CCTDR y CC4V poseen nodos con 5 vínculos. En ambas redes se aprecian nodos más centrales a la estructura y tienden a tener una mayor capacidad de influir sobre los demás.

Según accesibilidad de la red en términos topológicos las redes analizadas se ordenarían en forma decreciente en CCTDR, CCTR y CC4V a partir del indicador IAM.

A partir de estos índices podemos concluir que los tres grafos propuestos son redes complejas, ambas redes presenta índices similares desde los puntos de vistas de integridad, cercanía, centralidad y accesibilidad en términos generales.

Para profundizar el estudio se analizó 4 medidas de centralidad en función de los nodos de los grafos: Grado (Degree); Cercanía (Closeness); Intermediación (Betwenness) y Centralidad del EigenVector.

FIGURA 14. Grado de Centralidad (Degree) según los Grafos de las estructuras de vecindades. Provincia de Neuquén

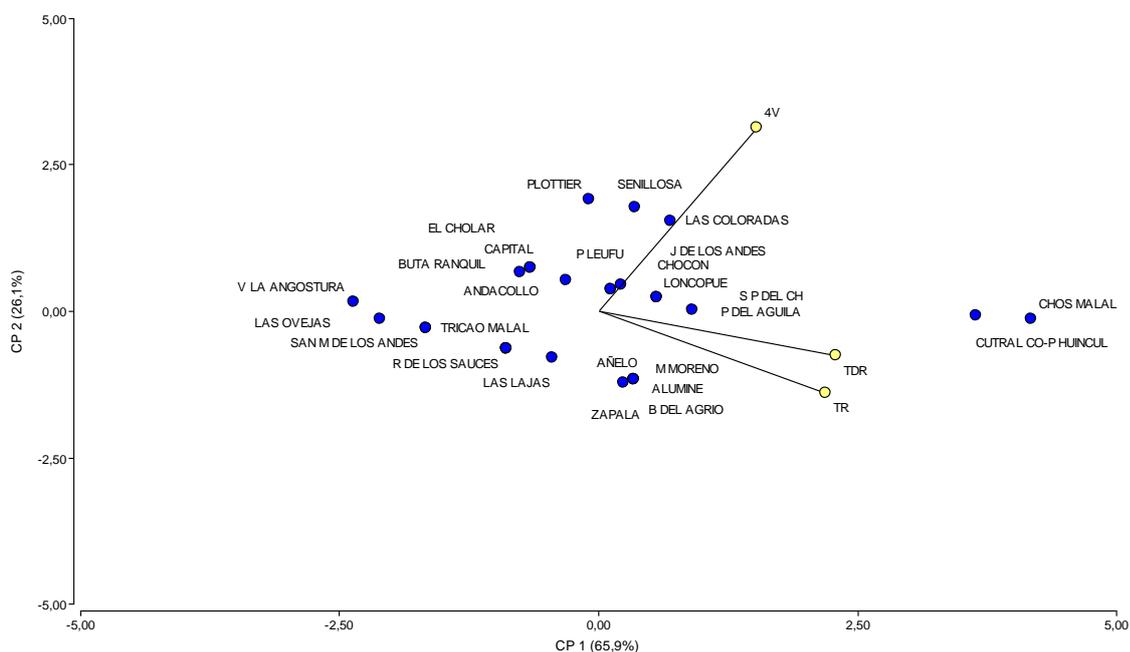


Fuente: Elaboración propia

Las redes de la figura 14 se indican con el tamaño y color del nodo, las áreas programas más centrales con respecto al grado de conexión (Degree) según el tipo de red utilizada en este trabajo. Los nodos presentan una asimetría en las redes CCTR y CCTDR, como era de esperarse por la definición de contigüidad de cada una de ellas destacándose las áreas programas de Chos Malal y Cutral Co-Plaza Huincul que presentan mayor centralidad.

En la red tipo Reina (CCTR) presenta una centralidad promedio de 0,16 con un desvío estandar de 0,07, para la red 4 vecinos (CC4V) es $0,19 \pm 0,07$ y $0,36 \pm 0,11$ para red Triangulación de Delaunay Restringida (CCTDR) en terminos generales, pero veamos un poco más en detalle como es la variabilidad de la áreas programáticas en termino de centralidad en un Análisis de componentes principales (ACP) integrando las 3 redes propuestas.

FIGURA 15. Biplot del grado de centralidad (Degree) según los Grafos de las estructuras de vecindades. Provincia de Neuquén



Nota: TR es la red tipo Reina (CCTR); TDR es la red Triangulación de Delaunay Restringida (CCTDR) y 4V es la red 4 vecinos más cercanos (CC4V).

Fuente: Elaboración propia

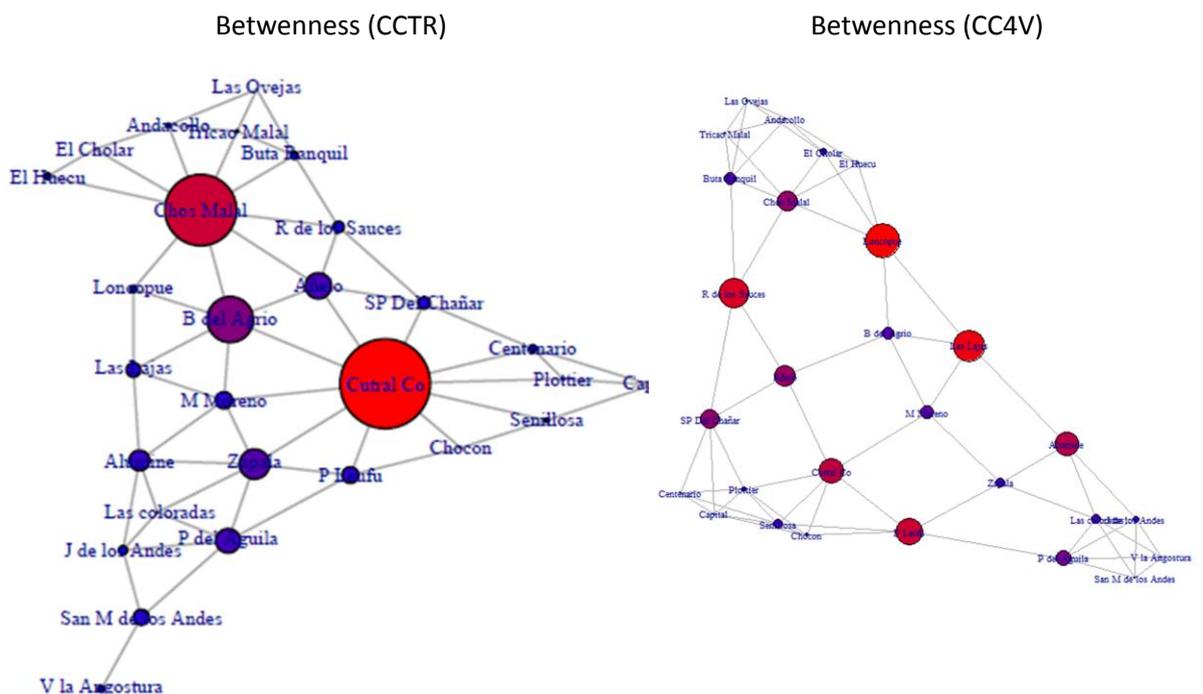
Luego de realizar un ACPse observó nodos más centrales y con mayor capacidad de influir en los demás en las redes CCTR y CCTDR, siendo las áreas programas de mayor centralidad Cutral Co-P Huincul, Chos Malal, Piedra del Aguila entre otras, esto se concluye a que la componente principal(CP1) explica el 65,9% de la variabilidad de los grados de centralidad observada en las redes, siendo las redes CCTDR y CCTR las que más contribuyen a la primer componente principal.

La red CC4V contribuye a la segunda componente principal (CP2) que explica el 26,1% de la variabilidad de los grados de centralidad observada en las redes. Las áreas programas más centrales en esta red son Plottier, Senillosa, Las Coloradas, Cutral Co y Chos Malal.

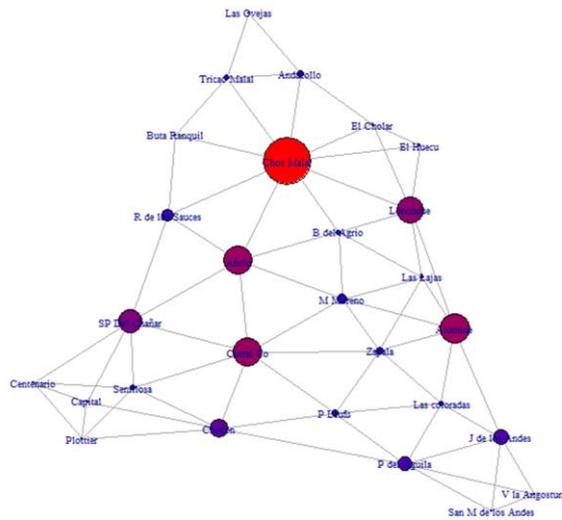
Tambien se puede apreciar que la red CC4V más cercanos no tiene correlación en los grados de los vertices de las redes CCTR y CCTDR, en cambio las redes CCTR y CCTDR estan correlacionadas.

Para continuar con este análisis exhaustivo de los indicadores locales de la red realizó un análisis similar que con el grado de centralidad pero ahora con el índice de intermediación para cada una de las redes analizadas.

FIGURA 16. Índice de Intermediación según los Grafos de las estructuras de vecindades. Provincia de Neuquén



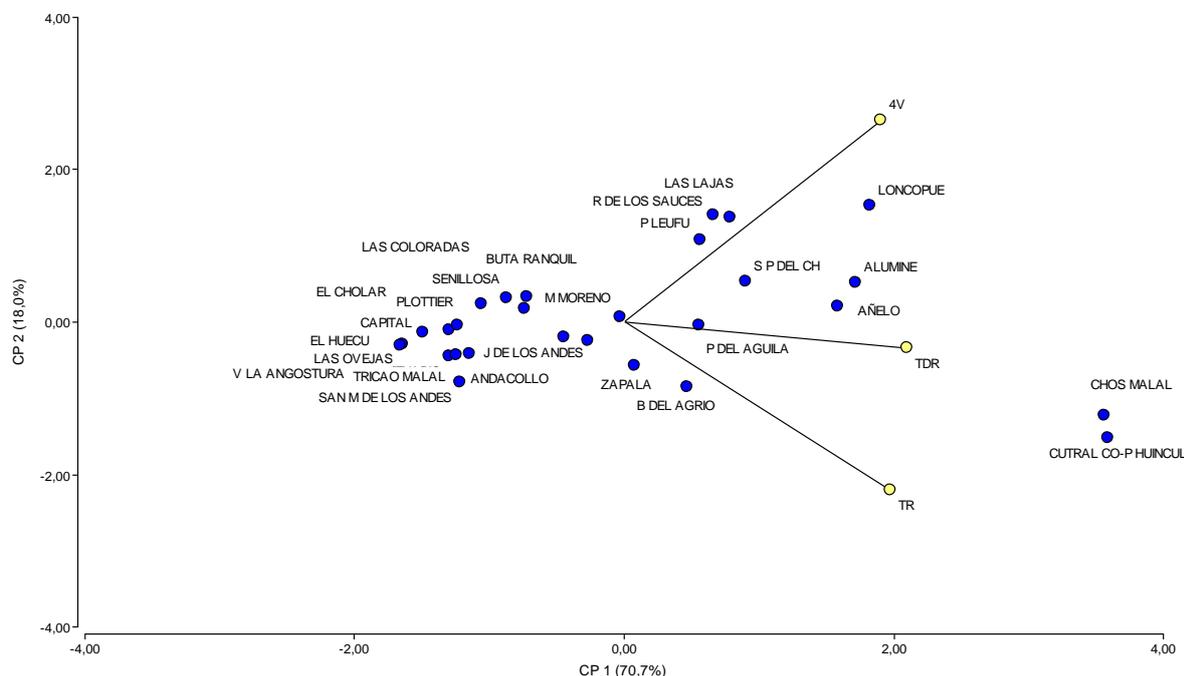
Betweenness (CCTDR)



Fuente: Elaboración propia

Las redes de la figura 16 se indican con el tamaño y color del nodo, las áreas programas con mayor intermediación (Betweenness) según el tipo de red utilizada. El promedio de intermediación y el desvío estándar de los tres tipos de redes de conexión analizada son las siguientes, 0.05 ± 0.42 en CCTR, 0.06 ± 0.21 en CC4V y 0.04 ± 0.28 en CCTDR. Observándose también asimetría en los nodos, en términos de intermediación, destacándose Cutral Co- Plaza Huincul, Chos Malal, Bajada del Agrio, Loncopue entre otras.

FIGURA 17. Biplot de los índices de intermediación según los Grafos de las estructuras de vecindades. Provincia de Neuquén



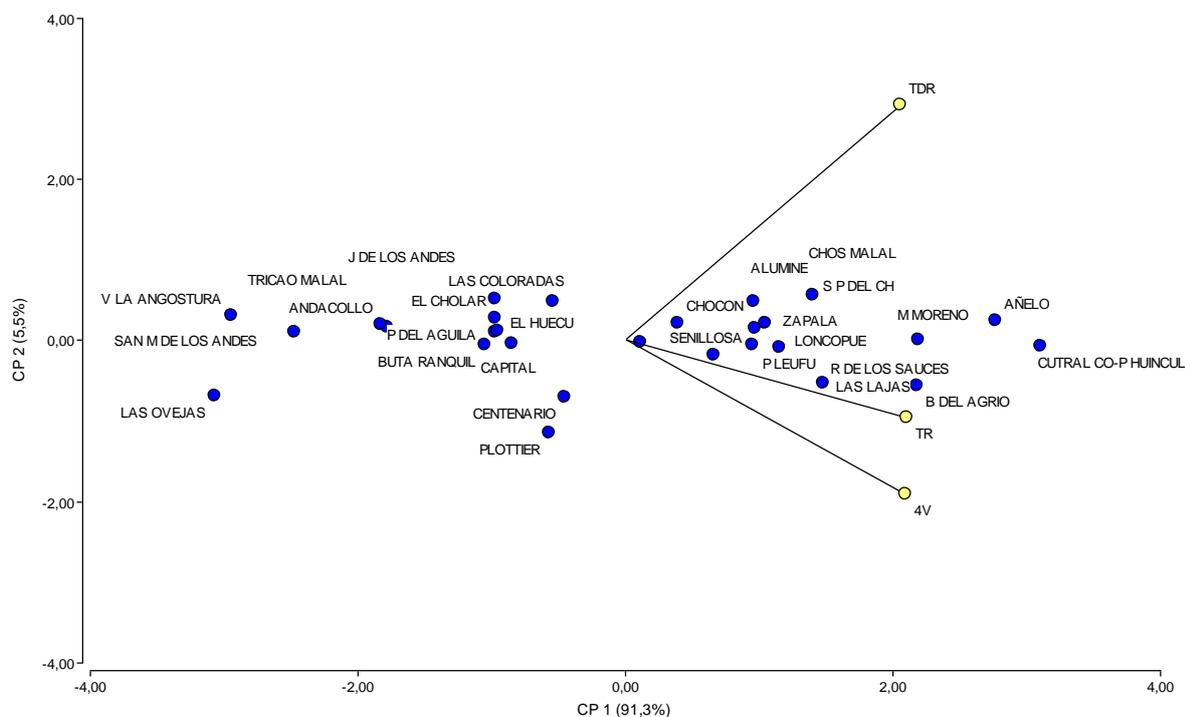
Nota: TR es la red tipo Reina (CCTR); TDR es la red Triangulación de Delaunay Restringida (CCTDR) y 4V es la red 4 vecinos más cercanos (CC4V)

Fuente: Elaboración propia

Del análisis de componentes principales (ACP) ambas redes contribuyen a la CP1 con un 70,7% de la variabilidad de los índices de intermediación observada en las mismas. Se concluye que la primer componentes es un gradiente con respecto a la intermediación las áreas programas con alto índice de intermediación son Chos Malal, Cutral Co-PHuincul, Añelo, Alumine, Loncopue entre otras y las áreas con menor índice de intermediación como Villa la Angostura, El Huecú, Las Ovejas y demás. Es decir que hay áreas bien representadas con respecto al grado de intermediación y otras no muy bien representadas. Con respecto a la correlación entre las redes, se observa que la red CC4V y CCTR tienen poca correlación entre las misma, pero si se puede observar que CCTR y CCTDR tienen una alta correlación.

El índice promedio de cercanía y su desvió estándar son: 0.35 ± 0.07 en CCTR, 0.34 ± 0.04 en CC4V y 0.38 ± 0.05 en CCTDR.

FIGURA 19. Biplot de los índices de Cercanía según los Grafos de las estructuras de vecindades. Provincia de Neuquén



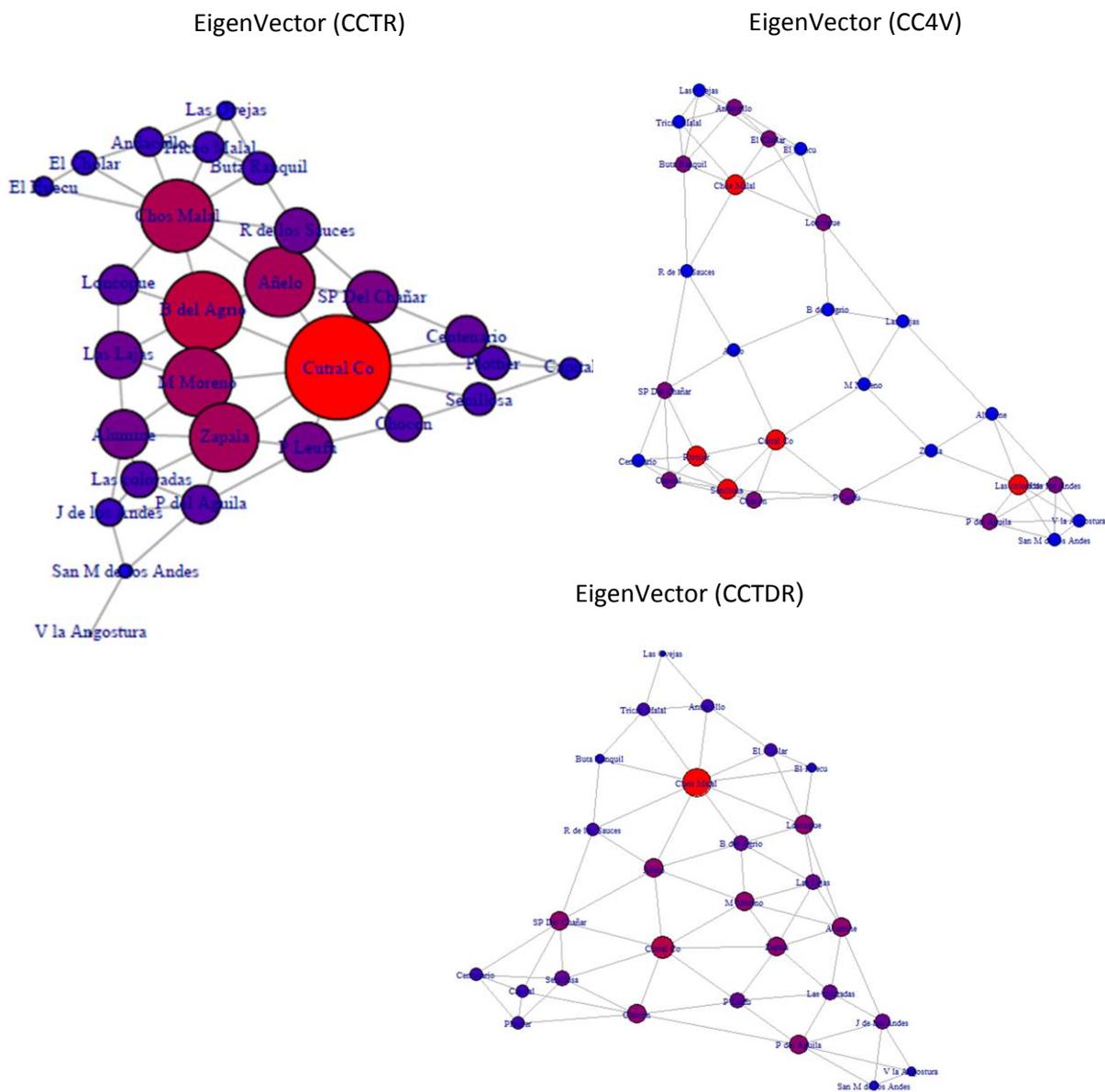
Nota: TR es la red tipo Reina (CCTR); TDR es la red Triangulación de Delaunay Restringida (CCTDR) y 4V es la red 4 vecinos más cercanos (CC4V).

Fuente: Elaboración propia

En cuanto a la cercanía luego de realizar un ACP, nuevamente puede considerarse la CP1 como un gradiente respecto a la cercanía, debido a que ambas redes contribuyen a la conformación de esta componente y capta un 91,3% de variabilidad observada.

Las áreas mejores presentadas en este índice son entre otras, Cutral Co-Phuincul, Añelo, Mariano Moreno, Bajada del Agrio y las menor representadas con Villa la Angostura, Las Ovejas, San Martín de los Andes. Existe una alta correlación entre la cercanía de los nodos en las redes CCTR y CCTDR, y sigue diferenciándose las redes de CC4V y CCTDR en términos de correlación, es decir no se observa, muy baja, correlación entre las mismas.

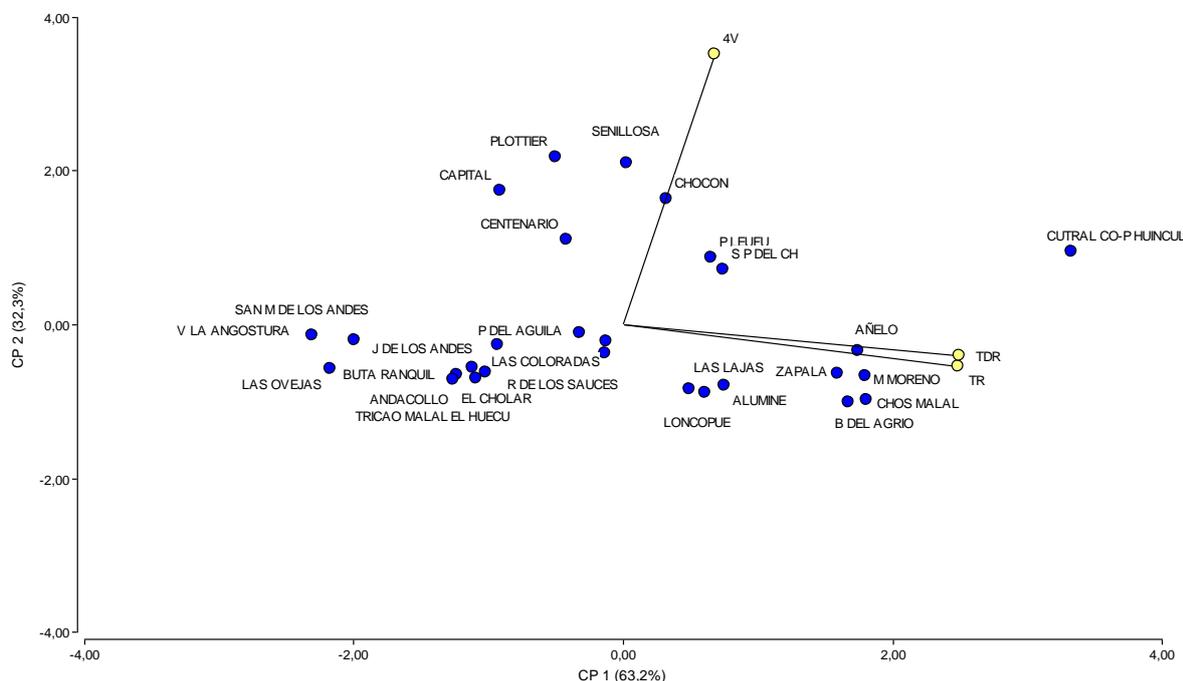
FIGURA 20. Índices de centralidad por EigenVector según los Grafos de las estructuras de vecindades. Provincia de Neuquén



Fuente: Elaboración propia

El tamaño y color de los nodos de las redes presentadas en la figura 20, corresponden a mayor índice de centralidad de las áreas programas analizadas. Los índices resumen para estas tres redes son: 0.42 ± 0.22 para red CCTR, 0.41 ± 0.28 para CC4V y para CCTDR 0.57 ± 0.25

FIGURA 21. Biplot de los índices de centralidad por EigenVector según los Grafos de las estructuras de vecindades. Provincia de Neuquén



Nota: TR es la red tipo Reina (CCTR); TDR es la red Triangulación de Delaunay Restringida (CCTDR) y 4V es la red 4 vecinos más cercanos (CC4V).

Fuente: Elaboración propia

En términos de nodos influyentes en la centralidad las redes CCTDR y CCTR son los que presentan mayores índices, captando una variabilidad explicada del 63% de las tres redes observadas, las áreas programa más influyentes son Cutral Co- PHuincul, Añelo, Zapala, Mariano Moreno y ChosMalal. La Red CC4V se diferencia de las redes CCTDR y CCTR en términos de correlación, pero capta áreas influyentes como Capital, Plottier, Chocón y Senillosa, además tiene mayor contribución el CP2 del ACP captando un 32,3% de variabilidad total observada en estas áreas. Luego del análisis de estas redes, se pueden observar áreas programas con índices altos de centralidad como Cutral Co- PHuincul, ChosMalal Zapala, entre otras, en contra posición a áreas muy desfavorecida como las áreas de Villa la Angostura, Las Ovejas, San Martín de los Andes. Las redes CCTR y CCTDR tienen una alta correlación en los 4 índices analizados y contribuyen a la conformación del CP1 que capta la mayor

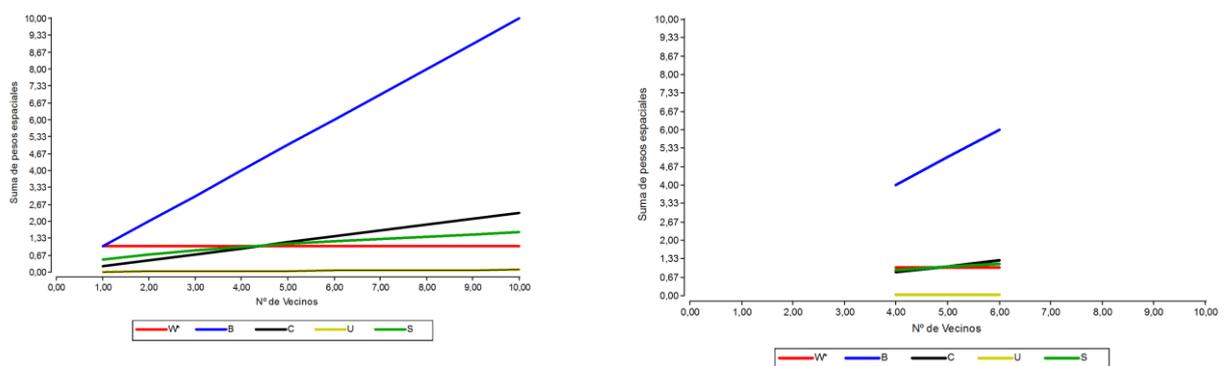
variabilidad observada de estos índices en el área programas a diferencia de la red CC4V.

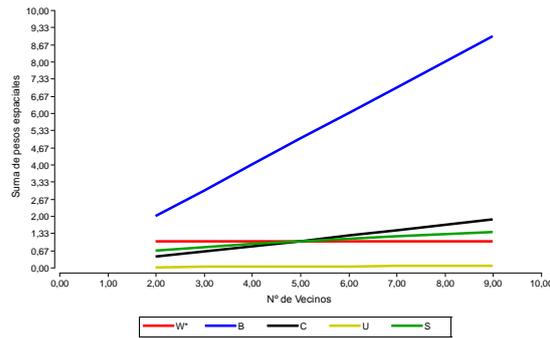
En términos generales las redes a CCTDR y CCTR son muy similares respecto a los índices locales de centralidad propuestos, diferenciándose con la red CC4V, pero en sí, este análisis permite caracterizar y evidenciar un gradiente respecto la centralidad que tienen las áreas programáticas de la provincia de Neuquén en función de su localización geográfica, diferenciándose polos de mucha centralidad en áreas como Cutral Co- PHuincul, Chos Malal Zapala respecto a áreas más desfavorables como Villa la Angostura, Las Ovejas, Junín de los Andes, San Martín de los Andes. En función de estrategias de prevención en cuanto a “descentralización de equipamientos” los actuales mamógrafos públicos se encuentran estratégicamente ubicados en áreas con mayor centralidad geográfica como Cutral Co- PHuincul, Chos Malal, Zapala como así también en áreas no centrales como Villa la Angostura y Neuquén Capital.

5.5. Análisis de la autocorrelación espacial.

Luego de analizar los tres criterios de contigüidad o redes de conexión CCTR, CC4V y CCTDR, se analizó los distintos tipos de conversión, es decir los pesos entre vecinos W^* , B , C , U y S en cada una de las redes de conexión presentadas. En las siguientes figuras se observa el peso que otorga a cada nodo según tipo de conversión en los tres criterios de contigüidad: CCTR, CC4V, CCTDR, se disponen en ese orden en la figura 22.

FIGURA 22. Estilos de conversión según criterios de contigüidad. Provincia de Neuquén





Fuente: Elaboración propia

En términos generales el comportamiento del sistema de estandarización es similar en los distintos criterios, diferenciándose muy levemente en algunos de ellos según el número de vecinos o enlace de cada estructura.

El estilo de conversión W^* es contante sin importar el criterio de contigüidad, donde el mínimo y máximo de las sumas de los pesos espaciales son iguales a 1, con una suma total de los pesos es igual a 28, que coincide con el número de área programa.

El estilo de conversión B , es el que se diferencia notoriamente a los demás estilos debido a relación directa y perfecta en función de números de vecinos, donde el mínimo de la suma de los pesos espaciales es igual al número mínimo de vecinos y el máximo de la suma de los pesos espaciales es igual al número máximo de vecinos según el criterio de contigüidad. La suma total de los pesos espaciales es igual a la suma del número de vecinos o enlaces, 120 en red CCTR, 132 en CC4V y 136 en CCTDR.

El estilo de conversión U , no es constante en función al número de vecinos, pero la pendiente de la recta es tan pequeña que parece casi constante, la diferencia entre el mínimo y máximo de la suma de pesos espaciales es muy pequeña. Este estilo estandariza con valores de sumas de pesos espaciales inferiores a estilo W^* y la variación es muy pequeña a medida que se aumenta el número de vecinos. La suma total de los pesos espaciales es igual a la unidad.

En los estilos de conversión C y S , las sumas de los pesos espaciales de S es superior al estilo C , pero su comportamiento se invierte a partir de un número determinado de vecinos, en estos casos a partir de 5 vecinos

aproximadamente. La suma total de los pesos espaciales de ambos estilos es de 28, igual al número de área programa.

A continuación se presentan unas tablas donde se pueden observar en detalles que complementan la información descripta de la figura 22.

TABLA 18. Pesos Espaciales según estilo de conversión, para el criterio de contigüidad de Tipo Reina (CCTR). Provincia de Neuquén

Estilo de conversión	Peso Mínimo	Peso Máximo	Mínimo de la suma de los pesos por fila	Máximo de la suma de los pesos por fila	Suma total de los pesos
W*	1,00	1,000	1,000	1,000	28
B	0	1,000	1,000	10,000	120
C	0,233	0,233	0,233	2,333	28
U	0,0083	0,0083	0,0083	0,083	1
S	0,155	0,492	0,492	1,558	28

Fuente: Elaboración propia

TABLA 19. Estadísticos descriptivos según estilo de conversión, para el criterio de contigüidad de 4 vecinos más cercanos (CC4V). Provincia de Neuquén

Estilo de conversión	Peso Mínimo	Peso Máximo	Mínimo de la suma de los pesos por fila	Máximo de la suma de los pesos por fila	Suma total de los pesos
W*	0,167	0,250	1,000	1,000	28
B	1,000	1,000	4,000	6,000	132
C	0,212	0,212	0,848	1,273	28
U	0,0080	0,0080	0,0300	0,045	1
S	0,189	0,231	0,924	1,132	28

Fuente: Elaboración propia

TABLA 20. Estadísticos descriptivos según estilo de conversión, para el criterio de contigüidad Delaunay Restringido (CCTDR). Provincia de Neuquén

Estilo de conversión	Peso Mínimo	Peso Máximo	Mínimo de la suma de los pesos por fila	Máximo de la suma de los pesos por fila	Suma total de los pesos
W*	0,111	0,500	1,000	1,000	28
B	1,000	1,000	2,000	9,000	136
C	0,205	0,205	0,411	1,852	28
U	0,0070	0,0070	0,0140	0,066	1
S	0,153	0,324	0,649	1,377	28

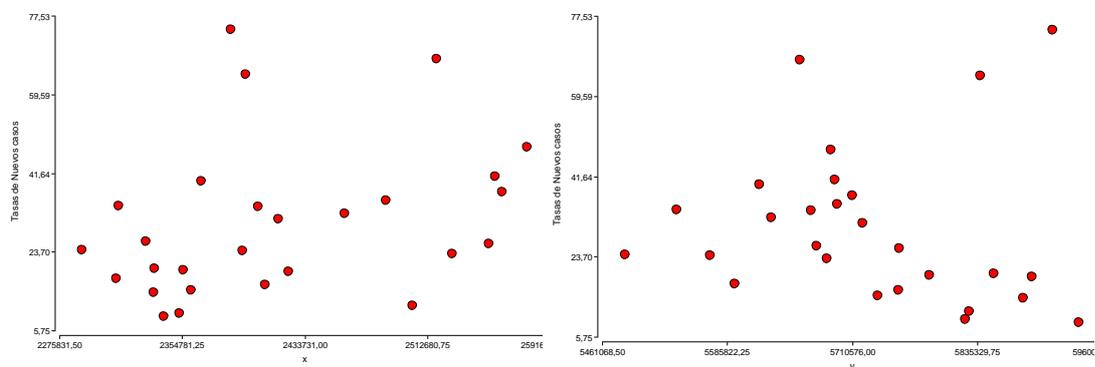
Fuente: Elaboración propia.

En ambos criterios, los pesos mínimo y máximo difieren solamente cuando se utiliza la matriz estandarizada W^* y S , para las demás matriz no se registran diferencias en estas medidas (González, 2015).

5.6. Detección de tendencias espaciales

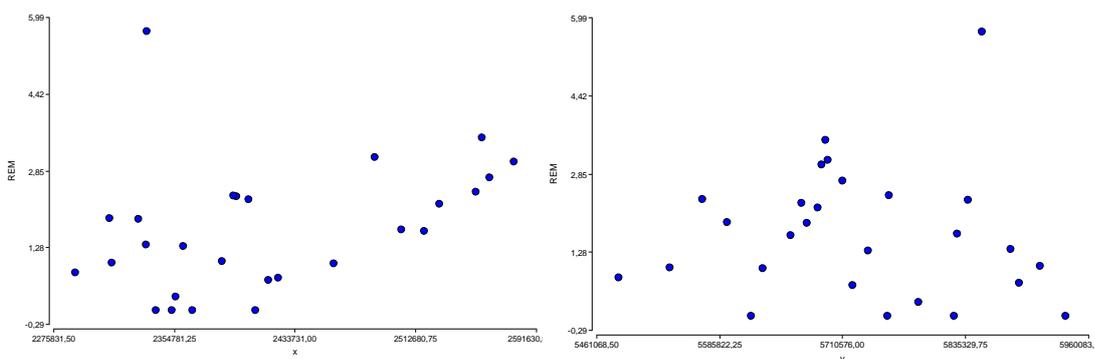
La estacionariedad es otro de los supuestos que se debiera cumplir para la realización del análisis geoestadístico. Este supuesto considera que el nivel promedio de la variable debe ser constante en todos los puntos del área de estudio, es decir, no deben existir tendencias. Su cumplimiento puede verificarse mediante la realización de gráficos de dispersión entre la variable en estudio y las coordenadas espaciales (Córdoba *et al.*, 2014). La figura 23, 24, 25 y 26 se muestra los diagramas de dispersión de las cuatro variables que se consideran en el estudio: Tasas de nuevos casos de cáncer de mama en Salud pública, Razón de Mortalidad Estandarizada (REM), Accesibilidad al Mamógrafo e Índice Socio Económico (ISE).

FIGURA 23. Tasas de nuevos casos de cáncer de mama en Salud Pública, de las áreas programáticas según coordenadas espaciales. Provincia de Neuquén. Período 2001-2015.



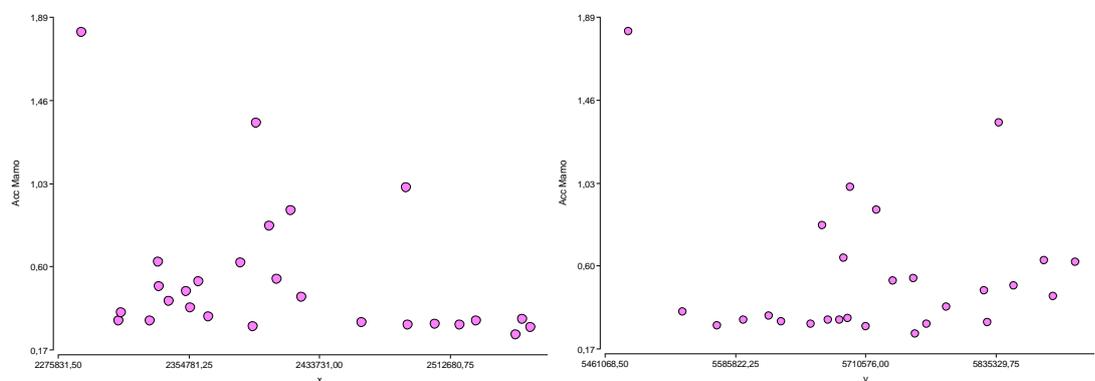
Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

FIGURA 24. Razón de Mortalidad Estandarizada de las áreas programáticas según coordenadas espaciales. Provincia de Neuquén. Período 2001-2012.



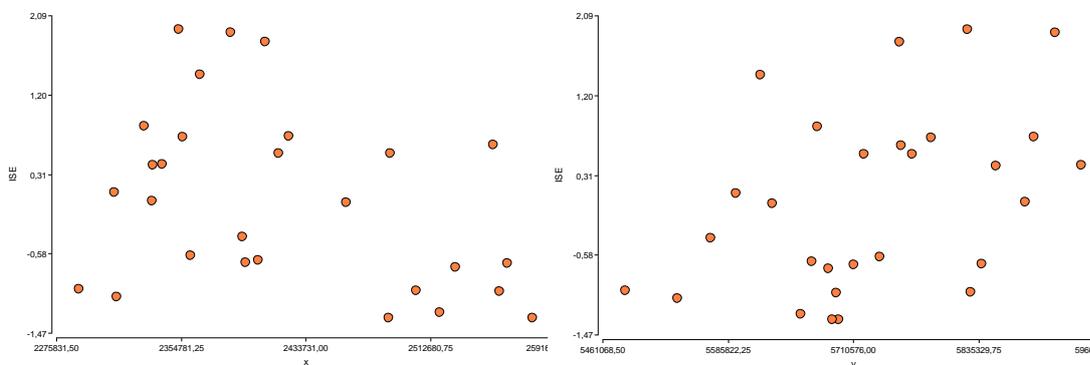
Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

FIGURA 25. Accesibilidad al Mamógrafo de las áreas programáticas según coordenadas espaciales. Provincia de Neuquén



Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

FIGURA 26. Índice Socio Económico de las áreas programáticas según coordenadas espaciales. Provincia de Neuquén.



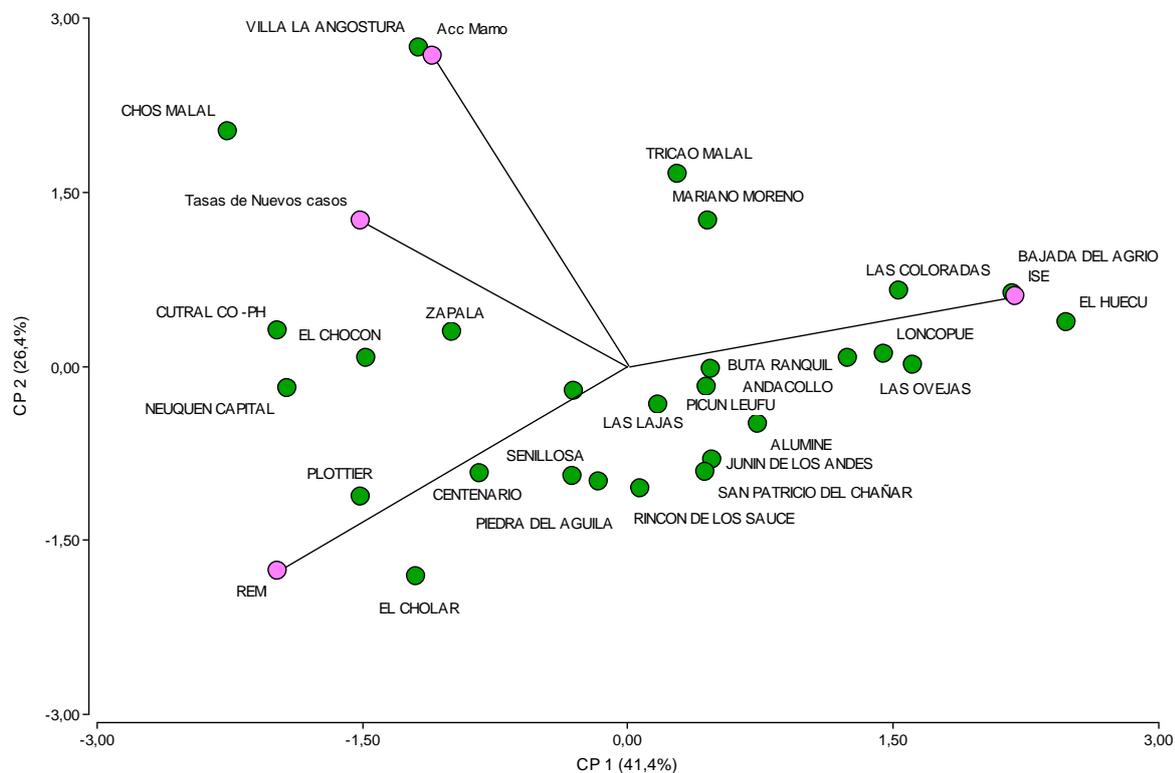
Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

En función de los diagramas de dispersión se incorporan los datos originales relevados sin necesidad de remover las tendencias sistemáticas en el análisis espacial (Bruno *et al.*, 2014).

5.7. Análisis de Componentes Principales espacial (ACPe)

Como primera instancia se realiza un análisis de componentes principales (ACP) sin tener en cuenta la componente espacial de las variables que se consideran para el estudio: Tasas de nuevos casos de cáncer de mama en Salud pública, Razón de Mortalidad Estandarizada (REM), Accesibilidad al Mamógrafo e Índice Socio Económico (ISE). El ACP es una técnica de reducción de dimensión que maximiza simultáneamente la varianza de la proyección de las observaciones sobre la componente principal y la suma de cuadrados de la correlación entre la componente principal y las variables (Bruno *et al.*, 2014).

FIGURA 27. Biplot de las áreas programáticas generado según ACP. Provincia de Neuquén



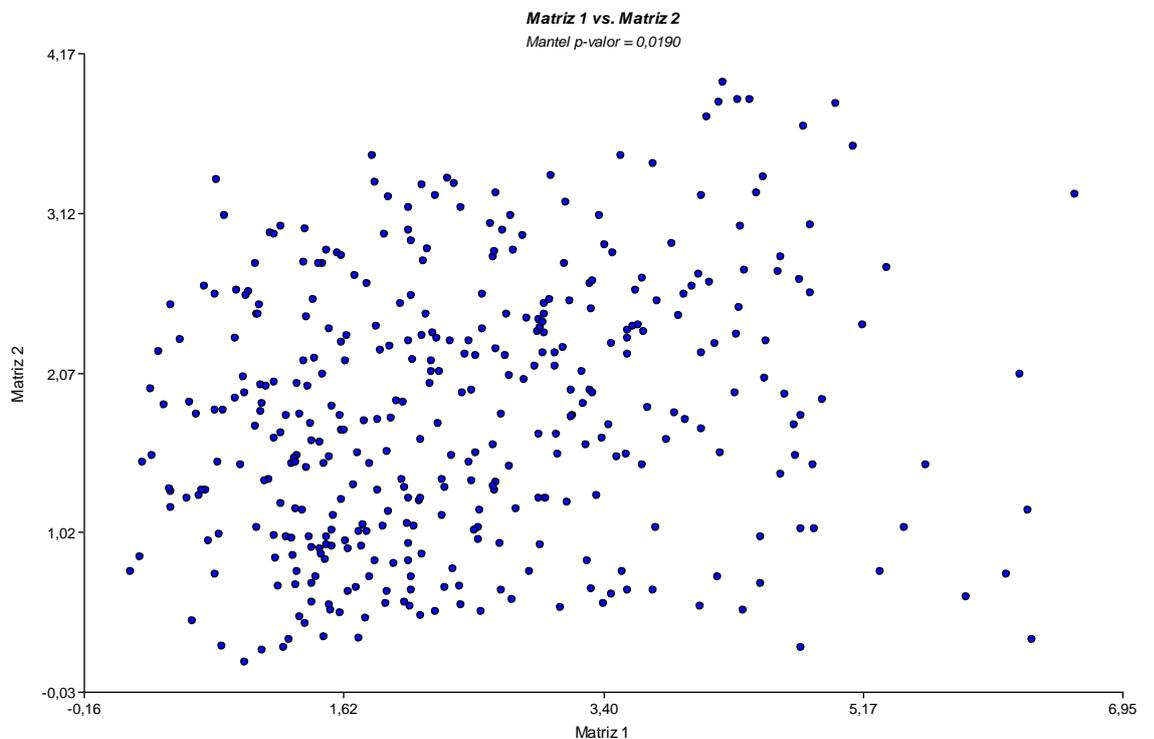
Nota: Acc Mamo: Accesibilidad al Mamógrafo, ISE: Índice Socio Económico, REM: Razón de Mortalidad Estandarizada, Tasas de Nuevos casos: Tasas de nuevos casos de cáncer de mama en Salud Pública.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

Del análisis de componentes principales (ACP) los dos primeros ejes explican 67,8% de la variabilidad de los datos, sin tener en cuenta la componente espacial. En la primera componente principal se observa un aporte del índice socio Económico (ISE) que se contrapone a la Razón de Mortalidad Estandarizada (REM) diferenciando áreas más vulnerables socioeconómicamente con bajas REMs como Bajada del Agrio, El Huecu, Las Coloradas, Loncopue, Las Ovejas. También se puede observar que no hay correlación entre el ISE y la accesibilidad al mamógrafo. En la segunda componente aportan las variables referidas a la salud, áreas programas con altas alta tasa de nuevos casos de cáncer de mama se asocian con mayor accesibilidad al mamógrafos y altas REM. En términos generales las áreas programáticas presentan una alta variabilidad en términos de las variables analizadas.

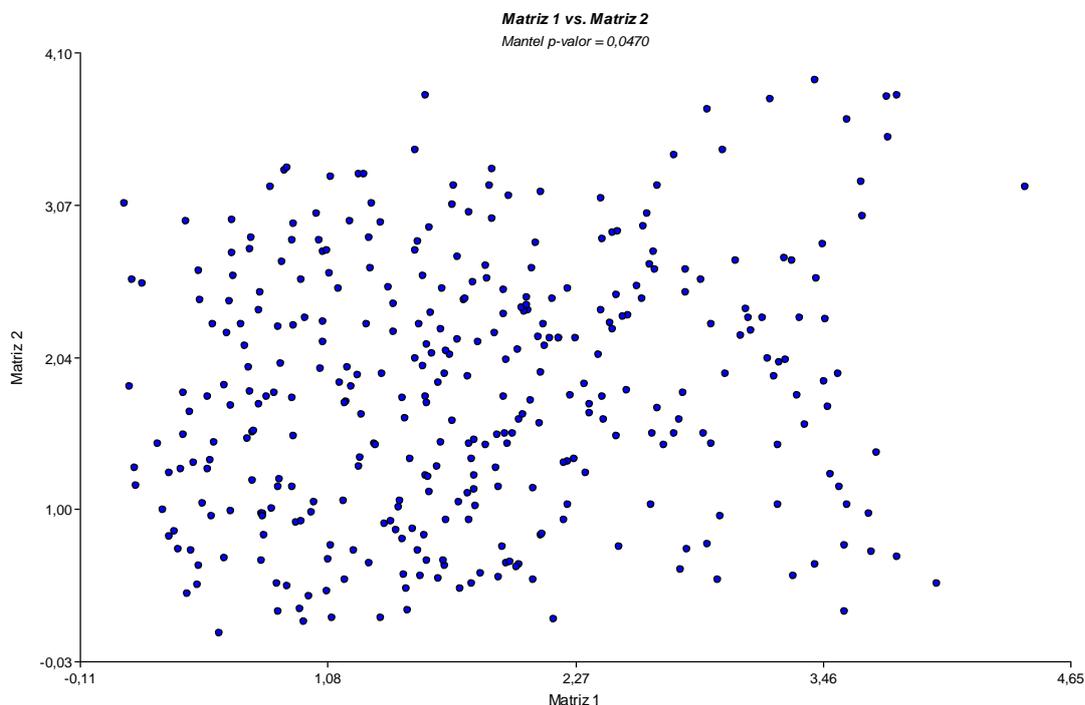
El test de Mantel estima el grado de correlación existente entre dos matrices X e Y . La hipótesis nula de esta técnica (H_0) postula que las distancias/similitudes entre las variables de la matriz respuesta Y no están linealmente correlacionados con las correspondientes distancias/similitudes en la matriz modelo X . Se trata, por tanto, de evaluar si la asociación (positiva o negativa) es más robusta de lo que cabría esperar por puro azar. La inclusión del espacio en este proceso es sencilla, basta incluir como matriz modelo X , una matriz de distancias. De este modo puede evaluarse si la distancia en el espacio determinado está correlacionada con la distancia en el espacio geográfico Mantel (1967).

FIGURA 28. Test de Mantel: Correlación entre las matrices distancia geográfica y distancia Euclídea de los datos. Provincia de Neuquén



Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

FIGURA 29. Test de Mantel: Correlación entre las matrices distancia geográfica y CP1_ CP2 (ACP). Provincia de Neuquén



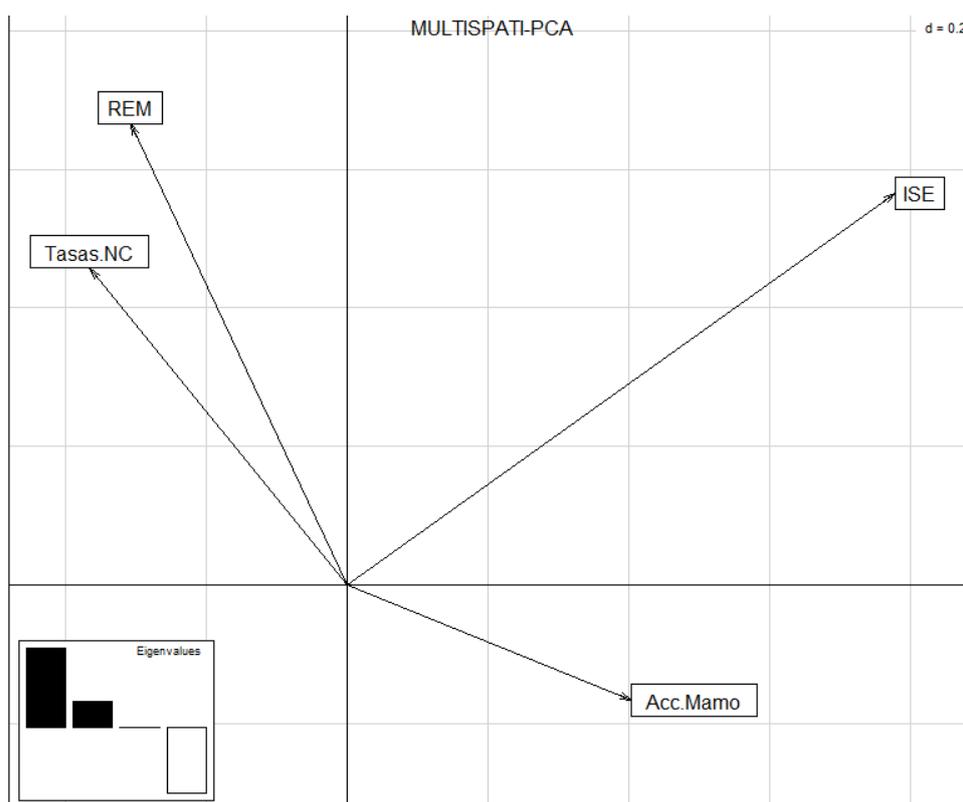
Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

De las figuras 28 y 29 resulta significativo el test de Mantel, lo que implica incorporar la estructura espacial en los datos.

El Análisis de Componentes principales (ACPe) surge como un análisis que pretende integrar el análisis multivariado con datos georreferenciados. Esta integración se realiza a través de dos análisis un primer análisis que resume la información multivariada como el ACP clásico y un segundo paso basado en una técnica univariada de estadística espacial que es aplicada sobre las variables sintéticas obtenidas desde el ACP. Una vez obtenido el ACP y la matriz de pesos espaciales sobre los cuales se podría realizar un estudio de autocorrelación espacial con el coeficiente de Morán la función “multispati” permite realizar un ACPe que incorpora la restricción dada por los datos espaciales. Los autovalores del ACPe son el producto entre la varianza y la autocorrelación espacial. Por el contrario, el ACP clásico sólo maximiza la varianza (Bruno et al., 2014).

A continuación, se aplica un ACPe en las variables propuesta para este trabajo, teniendo en cuenta la estructura espacial CCTDR para el cálculo del coeficiente de Morán.

FIGURA 30. Biplot de las áreas programáticas generado según ACPe. Provincia de Neuquén



Nota: Acc Mamo: Accesibilidad al Mamógrafo, ISE: Índice Socio Económico, REM: Razón de Mortalidad Estandarizada, Tasas de Nuevos casos: Tasas de nuevos casos de cáncer de mama en Salud Pública.

Fuente: Elaboración propia en base a información de la Subsecretaría de Salud de la provincia de Neuquén

Como se observa en la figura 29 en el margen inferior izquierdo (Eigenvalues) detectó una correlación espacial positiva en los dos primeros ejes, teniendo en cuenta la componente espacial (Bruno *et al.*, 2014).

En la primer componente el Índice Socio Económico (ISE) es la variable que más aporta a la construcción del mismo, además se detecta una correlación positiva con la variable accesibilidad al mamógrafo (Acc. Mamo), es decir que personas más vulnerables socio económicamente poseen accesibilidad al mamógrafo público.

Otro aporte significativo, es que no hay correlación entre el nivel socio económico y las tasas de nuevos casos de cáncer de mama, como tampoco con las REMs (Novoa, 2009).

Como también altas tasas de casos nuevos de cáncer de mama y REMs se condicen a la no accesibilidad al Mamógrafo, es decir que la falta de accesibilidad al mamógrafo conlleva a la imposibilidad de detección temprana del CM, detectando casos de Cáncer de mama en estadios avanzados en Salud Pública (Novoa, 2009; Viniegra, 2010).

CAPÍTULO 6

CONCLUSIÓN

En el análisis topológico de las redes modeladas a partir de los criterios de contigüidad propuestos (CCTR, CC4V y CCTDR), se observó que los mamógrafos en Salud Pública se encuentran en los nodos (áreas programáticas) más influyentes: Cutral Có-Plaza Huinca, Chos Malal y Zapala de acuerdo a los diversos indicadores de centralidad utilizados. Es importante recordar que la centralidad en este trabajo se midió sólo en términos de contigüidad geográfica, lo cual no garantiza la accesibilidad de todas las áreas programáticas al mamógrafo. Así como tampoco contempla variables epidemiológicas, económicas y/o sociodemográficas.

En este trabajo de tesis se llevó a cabo un estudio epidemiológico del CM combinando indicadores de la enfermedad con información geográfica provista por la topología de las redes estudiadas. En función de los resultados obtenidos se identificó que desde el punto de vista de “estructura espacial” la distribución de las variables analizadas en la Provincia de Neuquén, siguen un patrón no aleatorio. Es decir que de acuerdo al criterio de contigüidad seleccionado, se observaron diferencias entre las distintas áreas programáticas reflejándose una asimetría en cuanto a lo espacial, respecto a las oportunidades de acceder al mamógrafo, registros de mortalidad y detección de nuevos casos de CM como también en su vulnerabilidad socioeconómica.

Luego de incorporar la restricción espacial al análisis estadístico (ACPe) se detectó una correlación positiva entre el indicador de vulnerabilidad socioeconómica (ISE) y la accesibilidad al mamógrafo que no mostró el ACP clásico (que no contempla la estructura espacial). Aunque en ambos análisis se observó que ISE fue la variable de mayor influencia, por el peso que ejerce sobre el eje principal (primer componente), en el ACPe se detectó que áreas programáticas más vulnerables, desde el punto de vista socioeconómico, también tienen accesibilidad al mamógrafo. Esta accesibilidad destaca la alta potencialidad del sistema de salud Neuquino, en cuanto a la capacidad de brindar servicio de atención mamario en distintos puntos geográficos.

En concordancia con lo planteado por Novoa (2009), nuestros resultados indicaron que el CM afecta a mujeres sin diferenciar nivel socioeconómico. La alta correlación positiva de REM con la variable Tasas NC (la detección de nuevos casos) sugiere que la detección de nuevos casos se produce en estadios avanzados del CM. La detección tardía puede deberse a la dificultad de acceder al mamógrafo con la frecuencia necesaria para una detección temprana del mismo.

BIBLIOGRAFÍA

- Alazraqui M., Mota E. y Spinelli H. (2007). El abordaje epidemiológico de las desigualdades en salud a nivel local. *Cad. Saúde Pública*. Rio de Janeiro, 23(2):321-330.
- Ancot, L., Paelinck, J., Klaassen, L. y W. Molle (1982). Topics in Regional Development Modelling”, en M. Albegov, A. Andersson y F. Snickars (eds), *Regional Development Modelling in Theory and Practice*, Amsterdam, North Holland.
- Anselin, L. (1980). Estimation Methods for Spatial Autoregressive Structures. Regional Science Dissertation and Monograph Series, Ithaca, New York.
- Anselin, L. (1988). Spatial econometrics: Methods and Models. Kluwer Academic Publishers.
- Anselin, L. (2002). Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models. *Agricultural Economics*, 17, pp. 247–267.
- Arriaga, E. (2011). “Análisis demográfico de la mortalidad”. Córdoba: Centro de Investigaciones y Estudios sobre Cultura y Sociedad- CONICET. UNC.
- Bavaud F. (1998). Models for spatial weights: A systematic look. *Geographical Analysis*. 30:153–171.
- Beuguitte L. (2014) Les indices Kansky avec R : premières fonctions. Grouperfmr.
<http://grouperfmr.hypotheses.org/tag/igraph>
- Braicovich T., Caro P., Cerda V., Oropeza M., Osio E. y Reyes C (2009). Introducción a la Teoría de Grafos (1a.ed.). Neuquén: Educo – Universidad Nacional del Comahue. ISBN 978-987-604-147-8.
- Bruno, C. y Balzarini, M. (2014) Aplicaciones del Análisis Multivariado. Cátedra de estadística y Biometría. Facultad de Ciencias Agropecuarias. Universidad Nacional de Córdoba, Argentina.
- Case, A., Rosen, H. y Hines, J. (1993). Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States. *Journal of Public Economics*, 52, pp. 285-307.
- Cliff, A. D. y Ord J. K. (1981). *Spatial Processes, Models and Applications*. London: Pion
- Comité Provincial de Biotecnología de Neuquén (2014). Asociación Trastuzumab + Emtansine (T-DM1) para tratamiento del cáncer de mama Her-2 positivo en estadiometastásico, refractario a quimioterapia de primeras líneas incluyendo Trastuzumab y un Taxano. Ministerio de Salud de la Provincia de Neuquén.
- Córdoba M., Bruno C., Aguate F., Tablada M. y Balzarini M. (2014). Análisis de la variabilidad espacial en lotes agrícolas. Manual de Buenas Prácticas. Ed. Balzarini, M. Eudecor. Córdoba, Argentina.

- Córdoba, M., Bruno, C., Costa, J. & Balzarini, M. (2013) Subfield management class delineation using cluster analysis from spatial principal components of soil variables. *Computers and Electronics in Agriculture*, 97, 6-14.
- Coro C. (2003). *Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales*. Consejería de Economía e Innovación Tecnológica. Príncipe de Vergara, 132. 28002 Madrid (España). ISBN 84-451-2442-0
- Coro C. (2003). *Métodos gráficos del análisis exploratorio de datos espaciales*. Instituto L.R. Klein-Dpto. de Economía Aplicada. Universidad Autónoma de Madrid.
- Cressie N. (1993). *Statistics for spatial data*. John Wiley, New York.
- Dacey M. (1969). Similarities in the Areal Distributions of Houses in Japan and Puerto Rico. *Area*, 3, pp. 35-37. The Eigenvector Approach. *Environment and Planning A*, 39(5), pp. 1193-1221.
- Delaunay, B. (1934) Sur la sphère vide. *Izvestia Akademii Nauk SSSR Otdelenie Matematicheskikh i Estestvennyy Nauk. Bulletin of Academy of Sciences of the USSR*. 7, págs. 793-800, 1934
- Diggle, P. (2003). *Statistical analysis of spatial point patterns (2nd.ed.)* London: Arnold.
- Dray, S., Saïd, S. & Débias, F. (2008) Spatial ordination of vegetation data using a generalization of Wartenberg's multivariate spatial correlation. *Journal of Vegetation Science*, 19, 45-56.
- Durand, G., Zannotti, M., Stefanazi, V., Moscoso, N., Blanco, A. y Bandoni, A. (2013). *Planeamiento óptimo de la accesibilidad a servicios de mamografía en la región sanitaria I de la provincia de Buenos Aires*.
- Elliot P., Wakefield J., Best N., Briggs D., editors. (2000). *Spatial Epidemiology Methods and Applications*. Oxford University Press.
- Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C. y Parkin, D. M. *Globocan (2008), Cancer Incidence and Mortality Worldwide: IARC Cancer Base No. 10* [Internet]: Lyon, France: International Agency for Research on Cancer. Disponible en: <http://globocan.iarc.fr>
- González, M. (2015). *Modelos extendidos para el análisis espacial en epidemiología del cáncer*. Córdoba
- Goodchild, M. (1987). A spatial analytical perspective on geographical information systems. *International Journal of Geographical Information Systems*, 1. Pp.327-334.
- Goodchild, M. (2008) Spatial autocorrelation. In KEMP, Karen (Ed.) *Encyclopedia of Geographic Information Science*. Thousand Oaks, California: SAGE, 2008, pp.397-398.
- Griffin, P.M., Scherrer, C.R. y Swann, J.L (2007). *Optimization of community Health Center Locations and Service Offering with Stastical Need Estimation*. IIE Transactions.

- Haenszel W., Loveland D.B. y Sirken M.G. (1962). *J. Natl. Cáncer Inst.* , 28 , 947 – 1101.
- Haining R. (1990) *Spatial data analysis in the social and environmental sciences*. Cambridge.
- Herrera M., Mur J. y Ruiz M. ¿Cuál matriz de pesos espaciales? Un enfoque sobre selección de modelos. Departamento de Análisis Económico. Universidad de Zaragoza. Gran vía 2-4 (50005). Zaragoza. España.
- Huairan Ye, Hyun Kim (2014). *Measuring Spatial Health Disparity Using a Network-Based Accessibility Index Method in a GIS Environment: A Case Study of Hillsborough County, Florida*.
- Jombart, T., Devillard, S., y Dufour, A.-B.(2009a). "Finding essential scales of spatial variation in ecological data: a multivariate approach". *Ecography* 32(1): 161-168.
- Kansky K., Danscoide P. (1989). Measures of network structure. In: Flux, número special. pp.89-121;doi:10.34006/flux.1989.913.
- Lamfre L., y Hasdeu S. (2016) Construcción de un indicador de privación como marco de referencia para guiar intervenciones sanitarias en la ciudad de Neuquén. *Rev Argent Salud Pública*. 2016;
- Lawson A., Browne W., Rodeiro C. (2003). *Disease Mapping with WinBUGS and MLwiN*. Wiley Ed.
- Lawson A. (2001). *Statistical Methods in Spatial Epidemiology*. John Wiley&Sons, Ltd.
- López Abente Ortega G., Ibáñez Martí C. (2000) *Aplicación de técnicas de análisis espacial a la mortalidad por cáncer en Madrid, Dirección General de Salud Pública*. Madrid. ISBN: 84-451-2030-1
- LopezHernandez F., Palacios Sanchez M. (2000). *Distintos modelos de dependencia espacial. Análisis de autocorrelación*. Anales de Economía Aplicada. XIV ReuniónASEPELTEspana. ISBN: 84-699-2357-9
- Macchiavelli, Rossana (2013). Saneamiento ambiental y salud en una población urbano-marginal de Córdoba, Argentina. Riesgos de enfermedades pediátricas. Facultad de Ciencias Médicas. Maestría en Salud Pública. UNC.
- Mahar, S., Bretthauer, K. M., Salzarulo, P.A (2011). Innovative Applications of O.R. Locating specialized service capacity in a multi-hospital network.
- Mantel, N. A. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.*, 27:209-220.
- Ministerio de Salud de la Nación, (2009). *Boletín de Vigilancia. Enfermedades No Transmisibles y Factores de Riesgo*. Dirección de Promoción y Protección de la Salud: Buenos Aires.
- Ministerio de Salud de Neuquén. (2011). Una mirada en la provincia del Neuquén desde la Epidemiología que se llevó a cabo en un encuentro Interjurisdiccional para el logro de acuerdos para un Programa de abordaje del Cáncer de Mama. Sala de Situación.

- Newman, M. E. J. (2008) Mathematics of networks. The New Palgrave Encyclopedia of Economics, 2nd edition, L. E. Blume and S. N. Durlauf (eds.), Palgrave Macmillan, Basingstoke.
- Noori, A. (2011) On the relation between centrality measures and consensus algorithms. Proceedings of International Conference on High Performance Computing and Simulation, pp. 225-232.
- Novoa A. (2009). Red provincial de Patología Mamaria de Prevención y control. Ministerio de Salud de la provincia de Neuquén.
- O'Loughlin, J. y Luc, A. (1992), Geography of International Conflict and Cooperation: Theory and Methods, en Michael Don Ward (ed.), The New Geopolitics, Gordon and Breach, Filadelfia, pp. 11-38.
- Ord, K. (1975). Estimation Methods for Models of Spatial Interaction. Journal of the American Statistical Association, 70, pp. 120-126.
- Pina M.F., Ferreira Alves S., Correira Ribero A.S., Castro Olhero A. (2010). Epidemiología espacial: Nuevos enfoques para viejas preguntas. Univ. Odontológica. 29(23):47-65. ISSN 0120-4319.
- PhD. Nigenda, G., MSc. Gonzalez Robledo M.C, PhD. Gonzalez Robledo, L.M., BA. Bejarano R.M (2011) Disparidades al descubierto: un estudio comparado sobre políticas de cáncer de mama en América Latina. Unión Latinoamericana contra el cáncer de la mujer (ULACCAM) y American Cancer Society (ACS). Cuernavaca, Morelos.
- Quesada Rico, J.A. (2015). Análisis de la mortalidad evitable prevenible mediante procesos puntuales espacio-temporales en la ciudad de Alicante en el periodo 2001-2011. Tesis Doctoral. Universidad de Alicante.
- Riboli, E; (2002); Estudio prospectivo europeo sobre cáncer y nutrición (EPIC): poblaciones de estudio y recopilación de datos. Nutr Salud Pública. 5 (6B): 1113-1124. Francia. ntr@iarc.fr
- Ripley, B.D. (1981). Spatial Statistics Wiley, New York.
- Stewart B. W., Kleihues P. (2003) (Eds): World Cancer Report. IARC Press. Lyon.
- Tiefelsdorf, M. (2000). Modelling Spatial Processes Springer-Verlag.
- Tiefelsdorf M. y Griffith D. (2007). Semiparametric Filtering of Spatial Autocorrelation: (Falta editorial)
- Tisnés A. (2012); Análisis de la mortalidad en la ciudad de Tandil utilizando Sistema de Información Geográfica (SIG), Revista Universitaria de Geografía. 21,89-111/ISSN 03268373.
- Tobler, W. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, 46, pp. 234-240.
- Umadevi, V. (2013). Case Study: Centrality Measure Analysis on Co- Authorships Network. Journal of Global Research in computer Science.

- Uscanga-Sánchez, S., Torres-Mejía, G., Ángeles-Llerenas, A., Domínguez-Malpica, R. y Lazcano-Ponce, E. (2014) Indicadores de procesos de tamizaje de cáncer de mama en México: un estudio de caso. *Salud Pública de México*. Vol. 56. N° 5. pp. 528-537.
- Vilalta y Perdomo, Carlos Javier (2005). Cómo enseñar autocorrelación espacial. *Economía, Sociedad y Territorio*, vol. V, núm. 18, pp. 323-333. El Colegio Mexiquense, A.C.Toluca, México.
- Viniegra, M ; Paolino, M ; Arrossi, S . (2010). *Cáncer de mama en Argentina: organización, cobertura y calidad de las acciones de prevención y control: Informe final julio 2010: diagnóstico de situación del Programa Nacional y Programas Provinciales*. 1a ed. - Buenos Aires: Organización Panamericana de la Salud - OPS. ISBN 978-950-710-124-3.
- Whang, F. (2012). Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review, *Annals of the Association of American Geographers*, 102:5, 1104-1112.
- Whang, F., Tang, Q. (2013). Planning toward equal accessibility to services: a quadratic programming approach.
- Wilson, R. (1979). *Introduction of Graph Theory*. Longman. Nueva York.