

ANÁLISIS ESPACIAL DE LA MORTALIDAD POR CÁNCER EN TANDIL 2003-2005 UTILIZANDO MÉTODOS BAYESIANOS

TISNÉS, Adela^()*

Resumen

El uso del espacio como categoría de análisis en general, y para el análisis de la desigual distribución de la mortalidad en este caso, no está exento de problemas teórico-metodológicos. El cálculo de indicadores a nivel municipal, será problemático debido al tamaño de las unidades de estudio. La precisión de los resultados en las áreas con menor población será mucho menor que en las regiones más pobladas, por lo tanto, pueden existir problemas de estimación de indicadores cuando se trabaja con áreas pequeñas. Se presenta una aplicación utilizando métodos bayesianos, que permiten “corregir” los mapas y las estimaciones, dada la estabilidad de los estimadores calculados, logrando estructuras que no eran visibles con los procedimientos estadísticos clásicos. En este sentido se desarrollará la mencionada metodología para analizar la desigual distribución de la mortalidad por cáncer en la ciudad de Tandil en el trienio 2003-2005.

Palabras clave: Análisis espacial; Ajuste bayesiano; Mortalidad por cáncer; Áreas pequeñas

^(*) Doctora en Demografía. Centro de Investigaciones Geográficas CIG-IGECHS-CONICET/UNCPBA. Universidad Nacional del Centro de la Provincia de Buenos Aires. Campus Universitario. Paraje Arroyo Seco s/n (7000) Tandil, Buenos Aires, Argentina atisnes@fch.unicen.edu.ar

SPATIAL ANALYSIS OF CANCER MORTALITY IN TANDIL 2003-2005 USING BAYESIAN METHODS

Abstract

The use of space as a category of analysis in general, and for the analysis of the unequal distribution of mortality in this case, which is not free of theoretical and methodological problems. The calculation of indicators at the municipal level will be problematic due to the size of the units of study. The accuracy of the results in areas with smaller populations will be much lower than in the more populated regions. Therefore there may be indicators estimation problems when working with small areas. An application is presented using Bayesian methods that allow "correct" maps and estimates, given the stability of the estimates calculated, obtaining structures that were not visible with classic statistical methods. This methodology will be developed to analyze the unequal distribution of cancer mortality in the city of Tandil in 2003-2005.

Keywords: Spatial analysis; Bayesian adjustment; Cancer mortality; Small areas

Introducción

La variabilidad espacial está presente en cualquier investigación geográfica en particular y en cualquier ciencia vinculada al espacio en general (Ecología, Epidemiología, Agricultura, Geología, entre otras). La Estadística Espacial fue brevemente presentada por Fisher en la década de 1930 en una investigación estadística aplicada a la Agricultura. El desafío es modelar la variabilidad espacial, dado que la aleatoriedad espacial tiene un impacto directo sobre el análisis de la mortalidad analizada desde su impronta espacial. La distribución aleatoria de los fallecimientos según causa (cáncer en este caso), como se presenta en este artículo, justifica realizar un análisis de la varianza a fin de contrastar las diferencias espaciales. Sin embargo, controlar el sesgo existente implica un esfuerzo más en términos de la complejización y la eficiencia del análisis. En este trabajo se analizará la variabilidad espacial, expresada a través de la desigual distribución de la mortalidad por edad, sexo y causa de muerte por cáncer en el espacio geográfico de la ciudad de Tandil en el trienio 2003-2005. El objetivo principal es presentar y aplicar una serie de procedimientos que podrían ser replicados en otras áreas geográficas del país y otros momentos históricos, permitiendo realizar estudios acerca de la variabilidad del riesgo de morir a lo largo del espacio, utilizando un método bayesiano de ajuste.

Para ello se recolectaron de los certificados de defunción de los Registros Civiles de la ciudad, los datos referidos a las muertes ocurridas en esos años. Luego se introdujeron en un sistema de información geográfica, que contaba con un plano con la digitalización de las calles de la ciudad, donde se dispuso la distribución de la mortalidad, a partir del dato captado de los certificados de defunción referidos al domicilio en el que vivía la persona fallecida. Se calcularon y analizaron las tasas de mortalidad de tres años calendario consecutivos (calculadas como un promedio simple de las defunciones de esos años y relacionadas con la población media del trienio), de manera de eliminar o disminuir los problemas asociados a la estacionalidad y a la aleatoriedad de la mortalidad. Además, se trabajó con tasas de mortalidad tipificadas, con la finalidad de neutralizar el efecto de la estructura etaria particular de las defunciones en general, y por cada causa de muerte en particular.

El análisis espacial o geográfico de la mortalidad por causas es una buena técnica para sugerir estudios posteriores más específicos que expliquen las diferencias encontradas entre las áreas geográficas estudiadas. Esto se puede realizar principalmente a través de estudios de diseño ecológico, también conocidos como de correlación geográfica, los cuales ayudan a describir la

relación entre las variaciones geográficas encontradas en las enfermedades estudiadas, y la concomitante variación en el grado de exposición a factores, generalmente estilos de vida, dieta etc.; o mediante los estudios geográficos de movimientos migratorios que ayudan a determinar si los riesgos de enfermar y morir de los emigrantes de una región de alto riesgo a otra de bajo riesgo, cambia con la migración, es decir se trata de separar el efecto del lugar del de las características personales.

¿Mapas de salud o mapas de enfermedad?

La construcción de mapas de enfermedad ha sido central para describir procesos epidemiológicos a través de la historia. La utilización de mapas para representar eventos de salud, se inicia aproximadamente en la década de 1970 (Barret, 2000). Sin embargo, hasta hace poco, se dispone sólo de un pequeño número de eventos y datos para áreas grandes. Los estudios desarrollados sobre áreas pequeñas deben esperar al desarrollo de métodos estadísticos avanzados para surgir y establecerse definitivamente (Havulinna, 2011).

El punto de vista espacial y los métodos gráficos de la Geografía fueron adoptados rápidamente por los estadísticos. Algunos de los primeros trabajos se ocupaban más que nada de la Geografía Física descriptiva. Para ilustrar esto, pueden citarse autores como Tulloch (1838) o Hind (1864). Uno de los principales contribuyentes en temas relacionados con la Geografía Humana, fue Ravenstein (1875), quién consideró de suma importancia el uso de mapas como medio de representación visual de los datos estadísticos. Ya en el siglo XX, Student (1914) fue el primer estadístico que mostró real preocupación acerca de los problemas espaciales. En 1907, en su análisis de recuento de celdas contiguas utilizando un hemacirómetro, se preocupó por el supuesto referido a cuántas celdas contiguas de una cuadrícula rectangular podían considerarse variables aleatorias independientes. Nace así el primer “coeficiente de autocorrelación espacial”. Luego Student (1914), discute la eliminación de las correlaciones espúreas debido a la posición de las observaciones en tiempo y espacio, haciendo uso de polinomios.

La representación gráfica y la cuantificación de los datos, exige en este punto, la aplicación de estadísticas adecuadas para no caer en mecanismos interpretativos rápidos y seguramente erróneos, por ejemplo, a partir de la vista simple de la distribución de puntos en una superficie. La estadística descriptiva, el cálculo de tasas generales o específicas, brutas y estandarizadas, razones, porcentajes u otras estadísticas, en general, necesitan para su cálculo, un denominador poblacional.

La posibilidad de representar acontecimientos relevantes para la salud (la distribución de los fallecimientos en este caso) por medio de mapas, permite un acercamiento visual al fenómeno que supera la simple observación de los datos organizados en las tablas.

Sin embargo, el uso del espacio como categoría de análisis no está exento de problemas teórico-metodológicos. Los estudios de lugares con mayor nivel de resolución (mayor desagregación geográfica), proporcionarán resultados que permitirán otra aproximación a los estudios individuales. En este sentido, el cálculo de indicadores a nivel municipal, por ejemplo, será problemático debido al tamaño de las unidades de estudio. La precisión de los resultados en las áreas con menor población será mucho menor que en las regiones más pobladas. Por lo tanto pueden existir problemas de estimación de indicadores cuando trabajamos con áreas pequeñas.

El concepto área pequeña no tiene una definición específica. En principio, no existe ningún umbral poblacional o territorial consensuado para definir con precisión qué es un área pequeña. Rao, las define como: “Cualquier región cuyo tamaño muestral en un estudio es demasiado pequeño como para producir estimaciones a un nivel de precisión aceptable” (Rao, 1999); (Rao, 2003:2). Por tanto el concepto de área pequeña que se manejará no será territorial o poblacional, sino de origen estadístico.

Pueden, sin embargo, indicarse algunos factores que ayuden a detectar un área pequeña. Ellos son:

- Tamaño poblacional.
- Frecuencia del evento de salud que se desea medir.
- Indicador que se desee estimar (en concreto su variabilidad).

Por lo tanto, si bien la construcción de mapas de enfermedades o muertes ha sido una herramienta central en el desarrollo de la Geografía de la salud, los métodos estadísticos con que se estudian, han cambiado y evolucionado recientemente.

Existen básicamente dos tipos de mapas de enfermedad o defunciones: los mapas de tasas estandarizadas de mortalidad y los mapas que muestran la significación estadística de la diferencia entre el riesgo de la enfermedad en cada área, y el riesgo general de la media sobre todo el mapa (López Abente, 2005). A simple vista el primer enfoque podría ser el más adecuado, porque aporta una estimación del parámetro de interés (tasa de las muertes por causa), pero en realidad, es el enfoque que se está dejando de utilizar (Clayton, 1992); (Martínez Benito, 2007).

Tradicionalmente, para superar el problema ocasionado por la utilización de poblaciones pequeñas, se han representado en los mapas las tasas

estandarizadas de mortalidad por el método indirecto (REM). Pero, debido a los inconvenientes que se plantean, es muy factible que los resultados que se obtengan de las tasas estandarizadas de mortalidad no representen la realidad correctamente (por calcularse un evento de interés con muy pocos casos o sobre unidades espaciales pequeñas), obteniéndose estimaciones del riesgo relativo muy extremas (con valores muy altos o muy bajos respecto de las demás unidades espaciales), lo que distorsionará la interpretación en el mapa y difícilmente sea útil su lectura.

Esto se debe principalmente al hecho de que el cáncer es relativamente poco común, y un mapa de tasas estandarizadas no es el más adecuado, ya que la variabilidad de las tasas refleja sólo en parte la heterogeneidad geográfica. La varianza restante (la que no queda explicada por las tasas), puede modelarse y explicarse con la variabilidad de Poisson (1) (se ampliará en el apartado Ajuste bayesiano) (Silva Aycaguer, 2003). Esto conduce a varios problemas, entre ellos y en particular que las tasas más extremas tienden a currir en las áreas con poblaciones más pequeñas, de modo que el principal interés en el mapa es aportado por los datos menos confiables (Cliff, 1975). El objetivo principal de los estudios basados en el análisis de causas de muerte en áreas pequeñas utilizando tasas estandarizadas y suavizadas, es obtener una estimación de los índices sintéticos particulares y confiables de salud para cada una de las áreas geográficas pequeñas en que se divide la región de estudio. Así, se podrán comparar las áreas en términos de indicadores de salud. Estos son principalmente estudios descriptivos, y se basan en el uso de mapas para poder ser representados (English, 1992). La obtención de medidas de resumen para cada área geográfica que luego se representan en los mapas correspondientes, revelan patrones espaciales de los indicadores de salud en un conjunto de áreas geográficas, principalmente el riesgo de enfermedad o muerte (Elliot, 2000). Pero además, permite comparar los mapas de indicadores de salud con mapas de riesgo potenciales, para determinar las áreas de alto y bajo riesgo de enfermedad o mortalidad. De manera paralela, es posible analizar la evolución temporal de los indicadores de salud para un conjunto de áreas geográficas (2).

Esta perspectiva viene a complementar las representaciones clásicas, ya que de su aplicación resulta una expresión visual de un fenómeno que representa su comportamiento. En otras palabras, se mapean patrones de organización espacial de tales sucesos. Greeland y Morgenstern llaman la atención acerca de la posible afectación a este enfoque por el llamado sesgo ecológico, el cual refiere al problema que presentan las asociaciones entre los diferentes niveles ecológicos para reflejar adecuadamente las asociaciones a nivel individual (Greeland y Morgenstern, citados por Richardson, 1987). Así, las

diferencias regionales obtenidas en las tasas que se analizan pueden producir confusión debido al sesgo ecológico (Greeland, 1989). Sin embargo, es imprescindible comprender que no siempre se quieren “trasladar” las conclusiones propias del ámbito grupal al individual, en cuyo caso tal sesgo no es motivo de preocupación.

Estos análisis de hechos referidos a la salud de las poblaciones, representados en mapas, de donde se obtiene información y se concluyen resultados, pueden ser pensados como metodologías de tipo descriptivas o como estudios de tipo exploratorios. Como metodología descriptiva, permite detectar la existencia de agrupaciones espaciales, y contribuye a cuantificar en términos absolutos y relativos la importancia de las diferencias presentes en un territorio (Wakefield, 2001). Este tipo de análisis ha sido muy utilizado para estudiar también, por ejemplo, la mortalidad por cáncer (3). Como metodología exploratoria, permite inferir algunas posibles explicaciones a las distribuciones obtenidas.

Estudiar cómo varía el riesgo en el espacio geográfico para una causa de muerte, permite formular hipótesis acerca de sus determinantes y dar indicios para planear futuros estudios de investigación en procura de explicaciones más refinadas (Lawson, 1999).

No debe perderse de vista lo que se mencionaba anteriormente: cuando se calculan indicadores con valores poblacionales pequeños (tanto en el numerador como en el denominador), las tasas que se calculan tienden a ser muy lábiles, inestables, y al momento de ser mapeadas muestran un comportamiento difícil de interpretar y sin patrones, o pueden tenerse valores razonables pero no confiables, dado el escaso número de casos (Clayton, 1987).

La solución que se plantea a estos problemas, surge a partir de la aplicación de los métodos bayesianos. Estos métodos, permiten “corregir” los mapas, y, gracias a la estabilidad de los estimadores calculados, lograr estructuras que no eran visibles con los procedimientos estadísticos clásicos (Wakefield, 2001; Silva Aycaguer, 2003). Con la aplicación de estos métodos, se pueden controlar los efectos de la edad sobre las tasas, y paralelamente, observar la correlación espacial de las áreas bajo estudio. En este sentido, parecería lógico esperar que la mortalidad en cada unidad territorial sea más parecida a la de las áreas circundantes, que a la de las áreas distantes.

Los métodos de ajuste tradicionales

Cuando se trabaja con datos de población y mortalidad para una serie de unidades espaciales, se dispone generalmente de la información en una

matriz o tabla (Cuadro N° 1). En las filas de la matriz, se ubican las unidades espaciales (UE1,...,UE_n), y en las columnas de la matriz, las variables a estudiar (Var1,...,Var_n). Así, en cada una de las celdas que surgen de la intersección de las filas y las columnas, se consigna el valor que adopta la Variable_m en la unidad espacial_n.

Cuadro N° 1. Matriz de datos geográficos

	Var1	Var2	Var3	Var4	...	Var_n
UE1	UE1Var1	UE1Var2	UE1Var3	UE1Var4		
UE2	UE2Var1	...				
UE3	UE3Var1					
UE4	UE4Var1					
UE5	UE5Var1					
...	
UE_n						UE_nVar_m

Fuente: elaboración personal

Si bien los datos son organizados de manera clara y precisa, la información que puede concluirse de la lectura de las tablas es muy básica; pero además, si estas matrices llegaran a contener gran cantidad de datos, serían difíciles de interpretar. Surge entonces la necesidad de buscar recursos expresivos más sencillos, que ayuden en un segundo momento a la toma de decisiones. Para simplificar la discusión, en lo sucesivo consideraremos que tenemos n unidades espaciales en que se puede desagregar un espacio dado (en este caso la ciudad de Tandil) y que se va a analizar la mortalidad por causa: 'cáncer'.

Si la intención final del estudio fuera el análisis de la mortalidad por esa causa en la ciudad de Tandil, se podría comenzar calculando las tasas brutas de mortalidad en cada una de las unidades espaciales. Pero atendiendo a las cuestiones planteadas hasta el momento, respecto de la inestabilidad de los indicadores en unidades espaciales pequeñas por un lado, y la influencia que tenía la estructura por edad de la población en el valor de la tasa, es muy posible que rápidamente se arribe a conclusiones erróneas. Por otra parte, operar con tasas específicas por grupos de edad, es una alternativa para contemplar esta realidad que no permite un juicio sintético, ya que si, por ejemplo, se cotejaron dos unidades, tendrían que establecerse comparaciones entre tantos pares de valores como grupos de edad se

consideren, y las diferencias entre los valores de cada par podrían ser de distinta magnitud y signo. De modo que, con el objetivo de neutralizar el efecto de las distintas composiciones por edad en el cálculo de las tasas brutas, con frecuencia se acude al ajuste de tasas por edad para cada unidad territorial considerada (Silva Aycaguer, 2003).

Mediante el ajuste indirecto, se compara el número observado de muertes en cada unidad, con el número esperado que se obtiene de aplicar las tasas específicas de una población de referencia (en este caso, la población del Total País, para el año 2001), a la estructura de edad de cada una de las unidades espaciales. La razón o proporción entre estos dos números (el observado y el esperado), es lo que se denomina razón estandarizada de mortalidad (REM), y se deriva de la tipificación indirecta que ya se ha mencionado.

Según palabras de (Silva Aycaguer, 2003: 315) puede afirmarse que: “entre estas dos variantes, con frecuencia se prefieren las REM, en buena medida porque consienten una interpretación ágil y sencilla: por ejemplo, una REM mayor que uno para una unidad geográfica dada, significa que la mortalidad para esa unidad es mayor que la que le correspondería si su patrón de mortalidad fuese similar al de la población de referencia”.

Pese a esto, el método no ha estado exento de críticas y cuestionamientos: por un lado, debido a que las REM pueden ser dependientes de los tamaños poblacionales con los que se esté intentando trabajar, provocaría un impacto notorio sobre la variabilidad de las estimaciones. Este fenómeno ocurre con más frecuencia cuando las unidades geográficas sobre las que se quieren representar las tasas son pequeñas (Clayton, 1987). Y en este sentido, afirman los autores Silva Aycaguer, Benavidez, Rodríguez Vidal y Rodeiro: “cuando los datos son escasos, se pueden producir estimaciones inestables para las REM en cada área. Típicamente, las estimaciones extremas tienden a ubicarse en aquellas áreas cuyos tamaños poblacionales son más pequeños; puede ocurrir, por ejemplo, que una REM alta en cierta unidad no necesariamente responda a la presencia de una verdadera singularidad en ese sitio, sino más bien a simple azar, que se expresa gracias a la reducida magnitud de los valores involucrados en su cómputo” (Silva Aycaguer, 2003: 316).

Las ventajas de utilizar los métodos bayesianos para estimar el riesgo relativo en áreas pequeñas

Como se vio anteriormente, la información que proporciona el enfoque tradicional no permite que se controle la inestabilidad estadística de las REM. Para controlar este hecho, hay que incorporar información adicional

de modo que, cuando se quiera estimar el riesgo relativo de un área, se controle analizando si ésta posee una población pequeña o grande. Para lograr este objetivo, los métodos bayesianos establecen una ponderación entre la información sobre la zona en la que se quiere estimar el riesgo relativo, y la información proporcionada por las otras áreas (Clayton, 1987). Si un área particular de interés tiene una población reducida, entonces, la estimación del riesgo relativo toma prestados valores de la información proporcionada por todas las áreas de mayor peso, para lograr que la información proporcionada por esta zona se suavice. Si, por el contrario, el área de interés tiene una gran población, no es necesario asignar más peso a la información que ella proporciona (Clayton, 1987).

Esto significa que la estimación de riesgo relativo para cada área es suavizada de tal manera que cuando un área tiene una pequeña población, su riesgo relativo tenderá hacia el valor medio de todos los riesgos relativos, que es por lo general 1 (sin embargo, si se obtiene la población de referencia externa no tienen que estar necesariamente centradas alrededor de 1 (Clayton, 1992). Por el contrario, si el área tiene una gran población, la estimación del riesgo relativo se acercará a su propia REM. Tales métodos se pueden extender cuando existe evidencia de un patrón espacial en la mortalidad. En este caso, la información suministrada por el resto de las áreas será de carácter local en vez de global, y el riesgo relativo de un área particular de interés no va a tender a la media global, sino más bien a un valor similar al riesgo relativo de las zonas adyacentes. Los enfoques bayesianos recopilan información sobre el resto de las zonas o sobre las áreas adyacentes en una distribución de probabilidad considerada en los riesgos relativos (Clayton, 1987). Esto se conoce como la distribución a priori e incorpora la variación de los riesgos relativos de las áreas geográficas de la región en estudio (Clayton, 1992).

El hecho de aplicar metodologías específicas novedosas en nuestro país (se cuenta con dos antecedentes: (Pujol, 2012); (Ministerio de salud y Ambiente de la Nación, 2004)) diseñadas concretamente para el procesamiento de datos en unidades espaciales pequeñas (Anselin, 1989); (Silva Aycaguer, 2003); (García Ballesteros, 2000); (Barcellos, 2003); (Benach, 2007), permitirá realizar un análisis exhaustivo, profundo y a nivel local acerca de cuáles son los factores que se conjugan favoreciendo la existencia, en ciertos lugares dentro de la ciudad, de tasas de mortalidad más o menos elevadas respecto de otras.

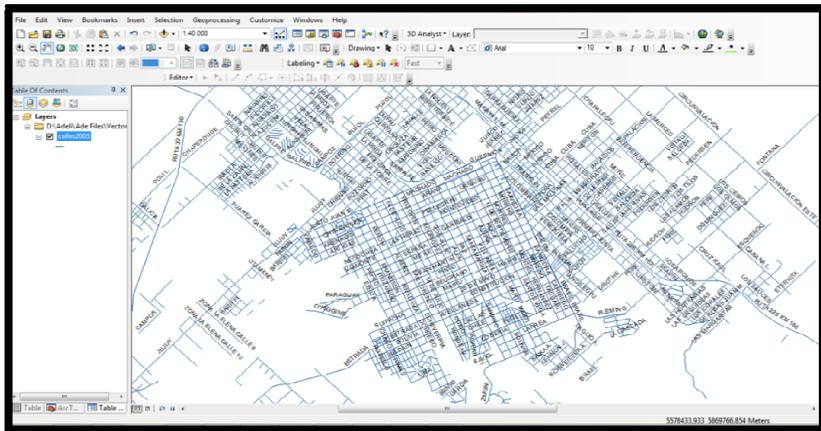
Aplicación del método de ajuste bayesiano para la ciudad de Tandil 2003-2005

Para este trabajo, se recolectó información proveniente de los certificados de defunción existentes en los Registros Civiles de la ciudad: fecha de defunción; sexo de la persona; lugar de ocurrencia; domicilio habitual de la persona fallecida; fecha de nacimiento; causa de la muerte. Esos datos se organizaron en una planilla de cálculo, y luego se ingresaron en un Sistema de Información Geográfica (SIG) (4) ArcGis 10.0.

De esta manera, y haciendo uso del dato correspondiente al 'Domicilio habitual de la persona fallecida', podemos tener el atributo de localización. Así, cada uno de los registros correspondientes a cada una de las personas fallecidas, pueden plasmarse cartográficamente y son susceptibles de ser tratados geográficamente (Moreno Jimenez, 2006). Esta tarea de incorporar el atributo de localización de una entidad o suceso en el espacio, se denomina georreferenciación o geocodificación. En este caso, se le están asignando coordenadas a las entidades (cada caso de fallecimiento), de las cuales conocemos un dato importante: la dirección postal. Esto aporta dos componentes: una información cartográfica correctamente localizada (portadora de coordenadas), y una información temática sobre los sucesos, capaz de ser vinculada de alguna forma a la base cartográfica. En la práctica, esto significa disponer del nombre y del número (altura) de la calle (5).

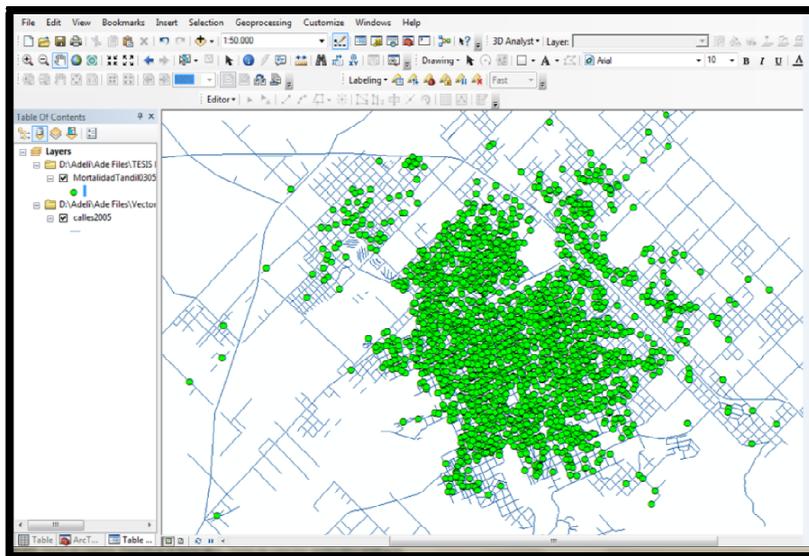
La base cartográfica digital que puede ser utilizada para la geocodificación suele ser una capa de líneas que representan las calles. Cada calle estará formada por una línea o conjunto de ellas y puede estar descompuesta o no, en varios tramos o arcos. Cada tramo (segmento), necesariamente deberá tener dos datos mínimos: el nombre de la calle y su numeración (para cada lado de la calle) y el sentido de la misma.

En el Mapa N° 1 se muestra la disposición del callejero de la ciudad de Tandil.

Mapa N° 1. Callejero de la ciudad de Tandil

Fuente: elaboración personal

El paso siguiente será realizar la geocodificación automática de las direcciones postales de las personas fallecidas, que quedarán de esta manera, incorporadas en el callejero de la ciudad. Esta operación consiste en la creación de una nueva capa de puntos, que poseerá una tabla de atributos asociada (con la información de cada uno de los registros de defunciones indicados anteriormente). Mediante este procedimiento, se tiene que cada registro o línea de la tabla estará asociado a un punto, cuya ubicación dilucida el sistema de información geográfica de manera casi automática (6) (Mapa N° 2).

Mapa N° 2. Distribución de los fallecimientos. Tandil 2003-2005

Fuente: elaboración personal

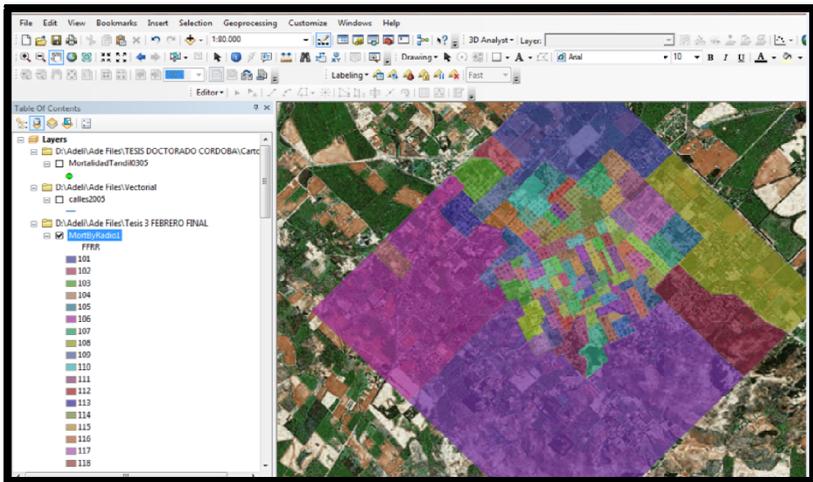
El procedimiento siguiente, será generar una nueva capa de información geográfica. Esta vez, de unidades espaciales a las que se les asignarán los casos de fallecimientos que se ubiquen dentro de cada uno de ellas. Estas unidades espaciales son los radios censales. Los radios censales son las unidades censales (y espaciales) mínimas, de naturaleza operativa, para las cuales existe una buena base cartográfica y poblacional establecida y disponible por el INDEC (Indec, 2010).

De esta manera, cada uno de los casos de mortalidad quedará agregado en el radio censal al que pertenezca, y tendrá asociada la información proveniente del Censo de Población y Viviendas para el año 2001. El hecho de utilizar datos agregados puede estar condicionada por dos razones fundamentales: por un lado, la disponibilidad de información a nivel agregado, y por otro, el interés particular de los estudios que tengan en cuenta como unidad de análisis las áreas geográficas en las que las áreas de estudio está dividida. La primera razón se debe a la imposibilidad de obtener datos a nivel individual, lo que puede estar explicado por tres razones básicas: 1) la recolección de los datos no es factible debido a que las fuentes de información disponibles no lo contienen, 2) los datos pueden estar disponibles, pero ser muy pocos, o muy incompletos, y 3) los datos se

disponen en forma agregada por razones de secreto estadístico. Sabiendo que las áreas con mayores tasas de desempleo, o una peor calidad de la vivienda, tienen, en general, mayores tasas de mortalidad, la detección de zonas con alta mortalidad puede ayudar a la administración pública a establecer y orientar la política social y de salud, adaptándolas a cada necesidad, para ayudar a reducir la mortalidad en las zonas más problemáticas.

Para el Censo del año 2001 (Censo con el que se trabaja en esta investigación, dado el período de estudio considerado), la ciudad de Tandil poseía 107 radios, dispuestos de la manera en que se presentan a continuación (Mapa N° 3)

Mapa N° 3. Radios Censales de la ciudad de Tandil. 2001



Fuente: elaboración personal en base a datos Indec, 2001

Sobre ese mapa, se agregan los puntos generados, correspondientes a las personas fallecidas, y se obtiene como resultado un mapa de radios censales que posee ahora un nuevo dato y tiene que ver con la cantidad de personas fallecidas (y todas las características asociadas a cada registro ya mencionadas), en cada una de estas unidades espaciales.

En primer lugar, se calcularán las tasas de mortalidad general según sexo por edades quinquenales. Se considerarán los datos de mortalidad general para Tandil durante el trienio 2003-2005 para las unidades espaciales indicadas (107 radios censales). La información referida a la mortalidad es la obtenida de los Certificados de Defunción y conformarán el numerador

de la REM (defunciones observadas). El denominador de la expresión, se obtuvo de la información disponible para el año 2001 proveniente del Censo Nacional de Población y Viviendas.

Se tendrá la siguiente expresión para cada radio censal:

$$REM = \sum_1^{20} \frac{d_j}{n_j t_j}$$

- $d_{\{j\}}$ representa el número de defunciones observadas en el j-ésimo grupo de edad
- $n_{\{j\}}$ es el total de personas por año; $t_{\{j\}}$ representa la tasa de mortalidad en Argentina en cada grupo de edad
- $C_{\{j\}}$ es el número de defunciones en el grupo de edad j-ésimo para todo el país
- $N_{\{j\}}$ el total poblacional en ese grupo ($t_{\{j\}} = \frac{C_{\{j\}}}{N_{\{j\}}}$);
- $n_{\{j\}} t_{\{j\}}$ expresa el número esperado de muertes en el j-ésimo grupo de edad.

Las unidades espaciales con REM mayores que 1, indicarán un exceso de riesgo, ya que el número de muertes observadas, será superior al valor que se esperaba. Las RME calculadas para cada grupo de edad, eliminan el efecto de este factor de la distribución geográfica del riesgo.

Para el caso de la ciudad de Tandil, se trabajará con los radios censales, por lo tanto, tendríamos, por ejemplo, para la fracción 02, radio 05, la siguiente tabla de distribución de fallecidos según grupos de edad quinquenales (Cuadro N° 2):

Cuadro N° 2: Fallecidos según grupos de edad quinquenales. Tandil 2003-2005

	Fracción 02 Radio 05		Tandil	
	Defunciones en radio 0205 (dj)	Pob. Total en radio 0205 (nj)	Defunciones en Tandil (Cj)	Pob. Total en Tandil (Nj)
0-4	1	64	57	7600
5 a 9	0	78	0	7874
10 a 14	0	88	2	8561
15 a 19	0	96	7	8447
20 a 24	0	83	11	8775
25 a 29	0	85	15	7205
30 a 34	0	72	12	6437
35 a 39	0	78	22	6230
40 a 44	1	73	31	6107
45 a 49	0	65	73	5701
50 a 54	5	81	99	5654
55 a 59	1	57	122	4946
60 a 64	1	65	168	4318
65 a 69	4	78	263	3778
70 a 74	3	60	362	3657
75 a 79	10	41	459	2752
80 a 84	5	19	500	1703
85 a 89	9	12	412	861
90 y mas	6	9	632	380

Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

Las defunciones observadas en la fracción 02, radio 05 de la ciudad de Tandil, son las que están ubicadas en la segunda columna. Puede observarse que en el grupo de edad de 0 a 4 años se produjo una sola muerte en el trienio. La suma de esa columna constituye el numerador de la REM:

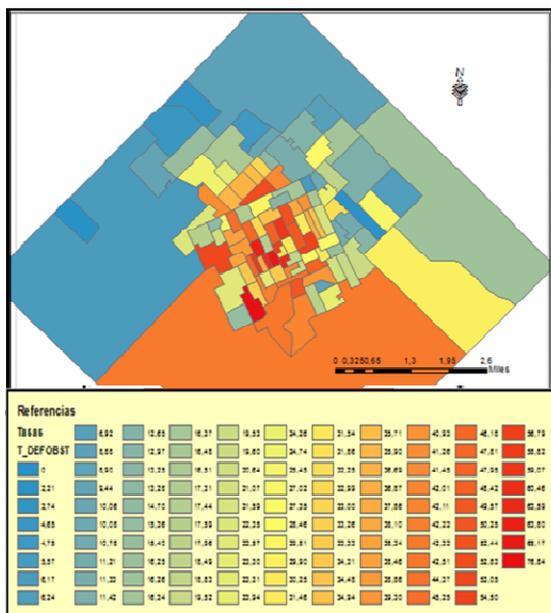
$$\sum_{j=1}^{19} d_j = 46$$

Como se mostró anteriormente, el denominador de la REM es el total de casos esperados. Para calcularlo se divide el total de defunciones de cada grupo de edad ($C_{\{j\}}$) entre la población total de ese grupo ($N_{\{j\}}$) y se multiplica el resultado por la población de la fracción 02, radio 05 del grupo de edad correspondiente. El valor esperado de muertes en el municipio es la suma de dichos productos para todos los grupos de edad:

$$\sum_{j=1}^{19} t_j = 52,32$$

Por lo tanto, el valor de la REM para la fracción 02, radio 05 es igual a 0,879, que es el resultado de dividir 46 entre 52,32. En este caso, no aparece con claridad un patrón espacial definido de mortalidad utilizando esta metodología. Esta situación se debe a la presencia de valores extremos, acaso debidos a fluctuaciones aleatorias no estructurales, que tienen un engañoso impacto en las áreas más pequeñas (Mapa N° 4).

Mapa N° 4. Mortalidad general Tandil 2003-2005. Método tradicional



Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

Habitualmente, los valores más extremos de las REM se producen en las áreas más pequeñas. Como puede observarse en el Mapa N° 4, no existe un patrón claro de distribución, pero además, las áreas más pequeñas son las que poseen los valores más altos (coloraciones rojizas). Estas áreas coinciden también en la ciudad de Tandil, con las zonas donde hay más

población y esta población tiene una estructura etaria envejecida. Por lo tanto, debería descartarse este tipo de ajuste para esta investigación.

Para solucionar este problema (eliminar o mitigar el efecto distorsionado de fluctuaciones aleatorias), una posibilidad consiste en calcular las “REM suavizadas” mediante un modelo jerárquico bayesiano.

Ajuste bayesiano

A medida que aumentamos la resolución geográfica, el problema de los valores poblacionales pequeños, también aumenta. Lo que vemos en un mapa crudo estará abrumado por el ruido aleatorio. Como se adelantaba, una solución posible es utilizar estimadores bayesianos, que aproximan las observaciones sobre la base de cifras pequeñas hacia la media global. Cuando haya motivos para creer que las observaciones son espacialmente dependientes, y por lo general así sucede, estos datos deben tenerse en cuenta. La solución general es llevar a cabo algún tipo de suavizamiento por encima del mapa. Sin embargo, los métodos de procesamiento de imágenes normales no son lo suficientemente potentes. Por otra parte, las opciones tales como la selección de los parámetros de suavizamiento son muy subjetivas. Estas preocupaciones han sembrado el campo de la cartografía de la enfermedad. La decisión acerca de utilizar o no el suavizamiento, depende naturalmente de la problemática que se analice y del lugar concreto que se esté investigando (Silva Aycaguer, 2003). Se propone el siguiente ejemplo: un tomador de decisiones del municipio, podría querer considerar el número bruto real de casos de fallecidos a causa de una enfermedad, para la revisión, por ejemplo, de la asignación de recursos de atención médica. Sin embargo, si estuviera interesado en prevenir lo que va a pasar el año que viene, habría un alto grado de incertidumbre debido a los pequeños números. En este caso, su verdadero interés sería la predicción basada en la tasa de incidencia de la enfermedad subyacente estimada (Clayton, 1987).

Los métodos estadísticos tradicionales utilizan para sus cálculos, información conseguida de manera empírica. El avance que produce este enfoque, tiene que ver con el tipo de información que utiliza para obtener una estimación. De esta manera, se combinan criterios que, a priori ha obtenido el investigador, surgidos de conocimientos y estudios previos, de reflexiones racionales y juicios razonablemente conformados (7). El resultado de esa integración, que se realiza por conducto del Teorema de Bayes (8), es una visión a posteriori que constituye la base de las inferencias subsiguientes (Silva Aycaguer, 2003).

Este enfoque, denominado bayesiano, opera con dos tipos de información relevantes para la estimación de las tasas de mortalidad o morbilidad en un

área determinada. El primer tipo de información, se obtendrá de la cantidad de personas-tiempo considerado en un área geográfica determinada. El segundo es el número de defunciones o casos observados de la enfermedad estudiada. Estos datos permiten el cálculo de las estimaciones convencionales de las tasas por máxima verosimilitud bajo el supuesto de que la variable estudiada (por ejemplo, número de defunciones observadas en un lapso dado), sigue una distribución Poisson (Alamilla, 2008); (Silva Aycaguer, 2003). Como anteriormente se mencionaba, cuando se trabaja con áreas pequeñas, es muy posible que se obtengan estimaciones extremas de las tasas en algunas de ellas. Inmediatamente asociada a la expresión área pequeña, aparece el problema de la ausencia de una muestra significativa de datos. Pero también por el costo inusitado y en la complejidad de los diseños muestrales que busquen alcanzar cotas aceptables de calidad en las estimaciones en todas las áreas o dominios de interés para los investigadores (Mancho Corcuera, 2002).

Por lo tanto, no es extraño pensar y comprender que las estimaciones obtenidas dependen, fundamentalmente del conocimiento o experiencia previa sobre esas áreas, pero además, sobre qué otras áreas pudieran tener tasas similares a las de aquellas en que hay pocos casos. Para esto, habitualmente se definen unas “adyacencias” a partir de criterios tales como proximidad geográfica, similitud en materia económica, demográfica, etc. Esto es lo que constituye la información a priori. El caso más simple de llevar a cabo y de menos costo operativo, (simula en realidad una total ignorancia sobre el lugar que se está analizando), consiste en atribuir a las áreas 'conflictivas' el valor medio de todas las áreas presentes en el estudio. Claro que si los casos observados son pocos y la cantidad de personas-tiempo reducido, entonces el peso específico que adquirirá la información a priori será determinante; en cambio, si la información que se posee inicialmente sobre esas áreas es alto, entonces el peso de la información local será tanto o más importante que el de la contextual. La creencia a priori sobre las distintas áreas geográficas puede ser representada por una distribución de probabilidad que estará centrada en el valor que en principio sea más creíble, y tendrá una variabilidad inversamente proporcional al grado de certidumbre que quepa atribuir a dicha creencia (Silva Aycaguer, 2003).

Este enfoque, entonces, se ajusta a la teoría de los modelos jerárquicos bayesianos y en este sentido, el teorema de Bayes permite obtener la distribución a posteriori para las tasas poco confiables, como una magnitud proporcional al producto de la distribución a priori y la verosimilitud de los datos. Por lo tanto, la estimación de la tasa se convierte en una medida central de la distribución a posteriori.

Debe tenerse en cuenta que, si bien las REM pueden interpretarse como estimaciones de máxima verosimilitud del riesgo relativo (RR) bajo un modelo de Poisson, para enfermedades raras y áreas pequeñas, en cuyo contexto los riesgos individuales son heterogéneos, la variabilidad del riesgo relativo dentro de cada área excede al que se podía esperar para una distribución de Poisson. Esta variación “extra-Poisson” se puede manejar considerando a los riesgos relativos dentro de cada área, como una variable. Los métodos bayesianos se pueden usar con tal finalidad, y producir así estimaciones suavizadas de las REM. Clayton y Kaldor hicieron una propuesta bayesiana para la modelación de los riesgos relativos y evitar así la mencionada inestabilidad de las REM brutas que se ubicarían en el mapa. La idea básica consiste en imponer una estructura a los riesgos relativos modelándolos colectivamente como un proceso estocástico espacial. En la modelación bayesiana esto significa que los riesgos relativos se suponen con una distribución a priori multivariante, cuyos parámetros determinan aspectos tales como el nivel global del riesgo y la interdependencia geográfica entre los valores correspondientes a las áreas (Clayton, 1987).

La distribución a priori reúne información de todas las áreas del mapa; posteriormente, para cada área se produce una estimación del riesgo relativo, establecida a partir de una relación entre la REM bruta y la información que se obtiene de las áreas adyacentes o que la rodean. Las fluctuaciones de las REM bruta son así reducidas, y el mapa se “suaviza” (es decir, se elimina la variación de Poisson).

Si se supone que el área total que es objeto de estudio está dividida en “n” áreas contiguas representadas mediante los índices $i=1, \dots, n$. Se denominará $O = (O_{\{1\}}, \dots, O_{\{n\}})$ al vector que tiene por coordenada genérica $O_{\{i\}}$ el número de casos de cierta enfermedad (o número de muertes) que se produjeron durante el período de estudio para la i -ésima área geográfica. El vector de casos esperados $E = (E_{\{1\}}, \dots, E_{\{n\}})$ se calcula aplicando a la población las tasas específicas por edad y sexo, asumiendo que éstas son constantes durante todo el período. Se denominará $x_i_{\{i\}}$ al riesgo relativo, desconocido, correspondiente a la i -ésima área, y se denominará $x_i = (x_{i_{\{1\}}}, \dots, x_{i_{\{n\}}})$ al vector de los n riesgos relativos.

Lo que se quiere es tener una visión probabilística que combine lo observado con una apreciación propia probabilística a priori. Es decir, los métodos bayesianos combinan los dos tipos de información: la que provee cada área a través de las muertes acaecidas (lo que permite calcular la verosimilitud suponiendo válido el modelo de Poisson) y la información a priori sobre los riesgos relativos. Coherentemente con el teorema de Bayes, la distribución a posteriori de los riesgos relativos viene dada por:

$$\left[\frac{\xi}{o} \right] \propto \left[\frac{o}{\xi} \right] \chi[\xi]$$

donde $\{x_i\}$ es la verosimilitud y ξ_i la distribución a priori que se atribuya a los riesgos relativos. La función de verosimilitud es el producto de n distribuciones independientes de Poisson puesto que los $O_{\{i\}}$ ($i=1, \dots, n$) son mutuamente independientes y siguen una distribución de Poisson con media $E_{\{i\}} \xi_{\{i\}}$.

$$\left[\frac{o}{\xi} \right] = \prod \left[\frac{o_i}{\xi_i} \right] = \prod \frac{e^{E_i \xi_i} (E_i \xi_i)^{o_i}}{o_i!}$$

Bajo estas condiciones, el estimador de máxima verosimilitud de $\xi_{\{i\}}$ coincide con REM:

$$\hat{\xi} = REM = \frac{o_i}{E_i}$$

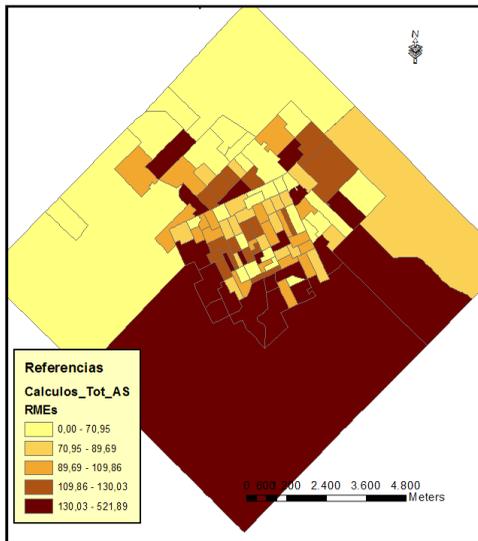
Para especificar la estructura a priori para $\xi_{\{i\}}$ existen diversas posibilidades, que van desde el modelo más simple, en el cual se considera un valor medio hacia el cual tiende cada área, hasta los que postulan una compleja variante de interdependencia entre las áreas. En este trabajo se considera una estructura en la cual la estimación de los riesgos relativos en cada área solo depende de los riesgos relativos en las áreas contiguas o adyacentes. Puesto que se parte de la creencia de que existen correlaciones espaciales entre los riesgos relativos (por ejemplo, qué áreas geográficamente próximas tienden a tener riesgos relativos similares), se considera una estructura espacial de tal manera que aquellas áreas que rodean a una cierta área i tengan influencia sobre la estimación del riesgo relativo en ella (9).

El proceso de establecimiento de los valores a posteriori para los $\xi_{\{i\}}$ es considerablemente complejo desde el punto de vista computacional. Para obtenerlo, se deben estimar los parámetros del modelo jerárquico para lo cual se utilizan técnicas de simulación; en este caso se utiliza el Gibbs Sampler, que pertenece a la familia de métodos de Monte Carlo y Cadenas de Markov (MCMC), susceptible de ser aplicado usando el programa WinBugs.

Al aplicar este complejo algoritmo, se obtienen las estimaciones a posteriori de los riesgos relativos, es decir, las REM suavizadas (Mapas N° 5, N° 6 y N° 7).

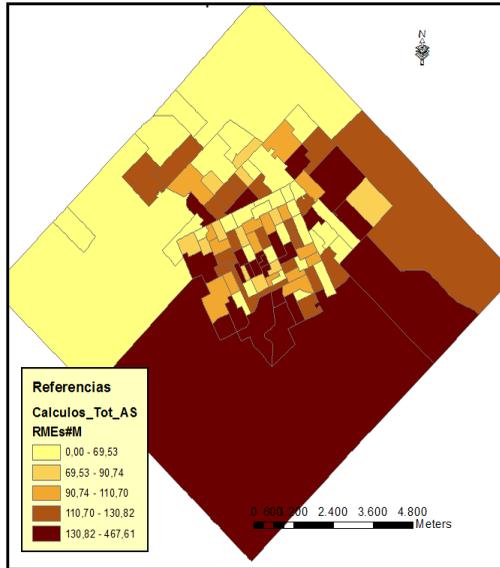
Los Mapas N° 5, N° 6 y N° 7 presentan la distribución geográfica de las REM calculadas según el método bayesiano explicado para la ciudad de Tandil. En este caso se han coloreado con tonalidades oscuras aquellos radios con valores de Razones Estandarizadas de Mortalidad suavizadas más elevados. A medida que los colores se aclaran, los valores de las REM's, descienden. Es interesante observar cómo las tasas ahora presentan un patrón marcado sobre el sur de la ciudad. La categorización se hizo de acuerdo a las sugerencias bibliográficas (Silva Aycaguer, 2003); (Clayton, 1992).

Mapa N° 5. REM's Población Total. Mortalidad por todas las causas



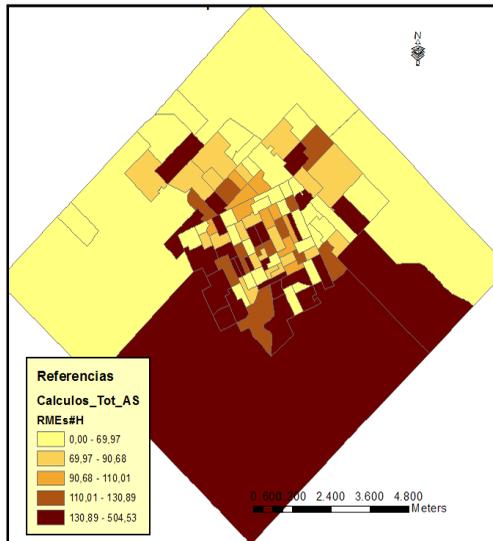
Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

Mapa N° 6. REM's Mujeres. Mortalidad por todas las causas



Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

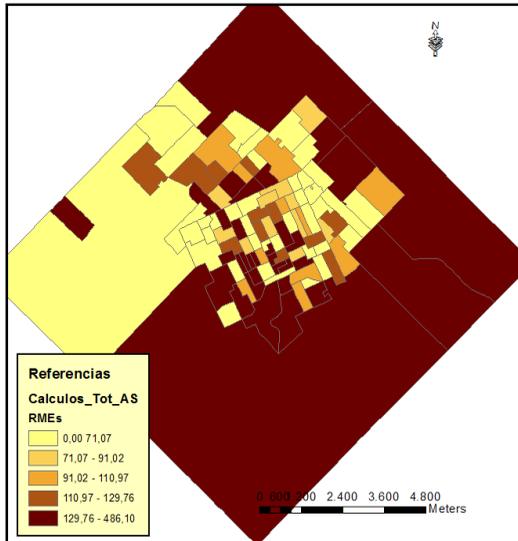
Mapa N° 7. REM's Varones. Mortalidad por todas las causas



Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

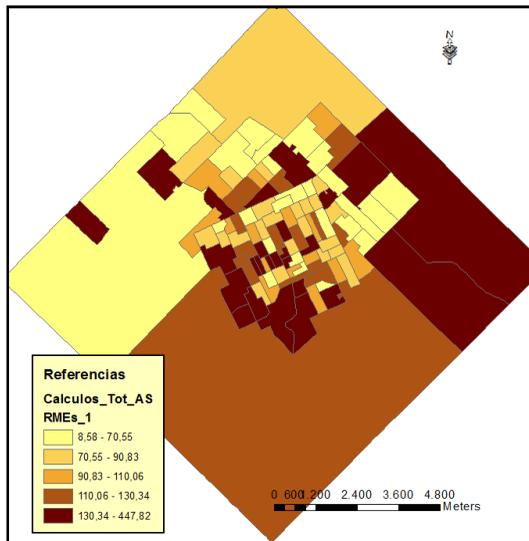
Si ahora se seleccionan sólo las muertes producidas por 'cáncer' y se replica el procedimiento, se obtienen los siguientes mapas, para la población total, varones y mujeres fallecidos por 'cáncer' (Mapas N° 8, N° 9 y N° 10).

Mapa N° 8. REM's Población Total. Mortalidad por cáncer



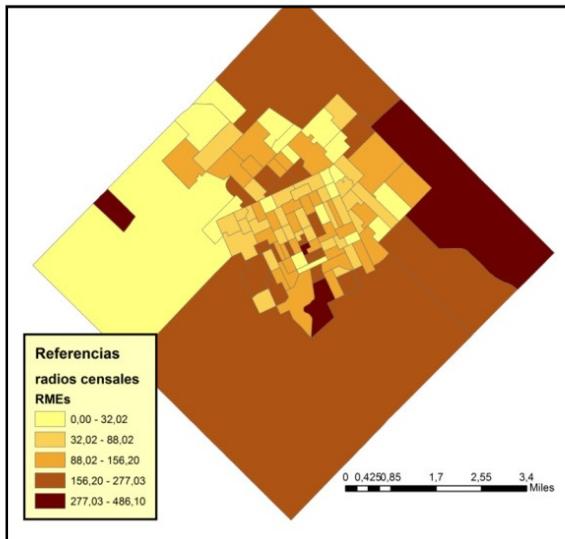
Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

Mapa N° 9. REM's Mujeres. Mortalidad por cáncer



Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

Mapa N° 10. REM's Varones. Mortalidad por cáncer



Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

Los mapas se realizaron con la RME ajustada o suavizada (RMEs) por el modelo descrito anteriormente, y fue necesario categorizarlas para su representación en los mapas. Se utilizó una escala relativa de cinco intervalos, agrupando la RME's de acuerdo a un criterio de riesgo. En la elección de las categorías se utilizó como referencia el Atlas de mortalidad de la Comunidad Valenciana de España (Martinez Benito, 2007). Para elegir el punto de corte de las categorías de mayor riesgo se realizó un análisis ROC (Receiver Operating Characteristic) (Dominguez, 2002), comparando distintos puntos de corte en relación con aquellas RME's que fueron significativamente mayores a 100. Del análisis anterior, surgen las cinco categorías de riesgo (Cuadro N° 3), donde Bajo Riesgo representa los radios en los que el riesgo es un 30% menor a lo esperado; Riesgo Moderado Bajo, representa los radios donde el riesgo es entre 29% y 10% menor a lo esperado. La categoría Riesgo Moderado Alto, representa los radios donde el riesgo es entre 11% y 29% mayor a lo esperado y, por último, Alto Riesgo representa riesgo mayor al 30% de lo esperado. En el Cuadro N° 3 se presenta la distribución de riesgo de la patología analizada. Para la representación de las categorías en el mapa, se eligieron cinco colores en una misma escala de tono marrón, donde el marrón intenso indica alto riesgo, así hasta el marrón claro que indica bajo riesgo.

Cuadro N° 3. Categorías de RME's y frecuencia porcentual observada por causa de mortalidad 'cáncer' según Población Total, Varones y Mujeres

		Categorías de RME's y frecuencia porcentual observada por causa de mortalidad 'Tumores' según Población Total, Hombres y Mujeres		
RME's	Categoría	Pob Total	Varones	Mujeres
<70	Bajo Riesgo (< al 30%)	37,38	37,38	24,30
71-90	Riesgo Moderado Bajo (Entre 29 y 10% menor)	10,28	10,28	21,50
91-110	Riesgo Esperado	10,28	10,28	14,95
111-130	Riesgo Moderado Alto (Entre 10 y 29% mayor)	13,08	13,08	14,95
>130	Alto Riesgo (>al 30%)	28,97	28,97	24,30

Fuente: elaboración personal sobre la base de datos Registros Civiles y DEIS

Conclusiones

Debido a que Tandil es una ciudad media, el total de población que posee, ocasiona inestabilidad estadística en las tasas que se calculan sobre toda la población o en alguna proporción de ella (tasas sobre el total de mujeres, de hombres, o sobre algún grupo etario particular). Además, no presenta hasta el momento, rupturas espaciales tan abruptas como se comienza a ver en las

grandes ciudades. Dado el escaso número de defunciones en algunos tramos etarios, es que, se debió ajustar la tasa de mortalidad por cáncer para hombres y mujeres por medio de un método Bayesiano, de manera de conseguir tasas estadísticamente estables.

Una vez que se obtuvieron las tasas de mortalidad por cáncer ajustadas, fue posible comenzar a aplicar las metodologías demográficas y estadísticas necesarias para alcanzar el objetivo principal.

El análisis espacial o geográfico de la mortalidad por causas es una buena técnica para sugerir estudios posteriores más específicos que expliquen las diferencias encontradas entre las áreas geográficas estudiadas. El hecho de estudiar cómo varía el riesgo en el espacio geográfico, permite luego esbozar hipótesis sobre los determinantes que inciden en la disposición espacial del riesgo, pudiendo así dar indicios para planear futuros estudios de investigación, y arribar a explicaciones más certeras.

Esto se puede realizar principalmente a través de estudios de diseño ecológico, también conocidos como de correlación geográfica, que ayudan a describir la relación entre las variaciones geográficas encontradas en las enfermedades estudiadas y la concomitante variación en el grado de exposición a factores, generalmente estilos de vida, dieta etc.; o mediante los estudios geográficos de movimientos migratorios que ayudan a determinar si los riesgos de enfermar y morir de los emigrantes de una región de alto riesgo a otra de bajo riesgo cambia con la migración, es decir se trata de separar el efecto del lugar del de las características personales.

Es interesante mostrar que, si bien a veces no es posible por cuestiones de tipo metodológicas y de disponibilidad de datos, el hecho de poseer información con un gran nivel de desagregación, permite observar “con lupa” el espacio geográfico y atender puntualmente a las necesidades sociales, demográficas y económicas. No debe perderse de vista la inestabilidad de las tasas que se obtienen, y al momento de ser mapeadas muestran un comportamiento difícil de interpretar y sin patrones, o pueden tenerse valores razonables pero no confiables, dado el escaso número de casos. La solución que se plantea surge a partir de la aplicación de los métodos bayesianos que permiten “corregir” los mapas, y, gracias a la estabilidad de los estimadores calculados

La información disponible (tanto de mortalidad, como de información sociodemográfica), desagregada geográficamente a nivel de radio, permitió un acercamiento muy puntual acerca del comportamiento del fenómeno que se intentaba explicar y de qué manera, el contexto local, favorecía o no, el aumento en los valores de la tasa de mortalidad por cáncer. Se destaca en este sentido, el importantísimo aporte que realizan a los análisis locales y globales, las herramientas espaciales, sumadas a la aplicación de métodos

estadísticos y bayesianos. Estos permiten sin duda alcanzar niveles de estudio y análisis a escala geográfica mucho más pequeña y en alguna medida complejos, a los que sería imposible llegar sin echar mano a estos métodos de ajuste. Sin embargo, debe hacerse notar que no debe caerse en aplicaciones estadísticas ligeras, ni alejadas del sentido común al de la visión crítica y aguda de los investigadores que las utilizan. Los métodos bayesianos utilizados, son una poderosísima herramienta de suavización comprobada internacionalmente, sin la cual es imposible pensar en niveles de desagregación.

Pero esos métodos están muy lejos de ser mágicos o de resolver todos los problemas que plantean las áreas pequeñas. Sólo con la observación de los valores 'outliers' encontrados, se acepta que no ha habido una suavización perfecta o límite, y es necesario aceptar esa realidad, para abordar cualquier investigación.

Notas

(1) La distribución de Poisson es un modelo discreto, pero en el que el conjunto de valores con probabilidad no nula es numerable. Esta distribución se utiliza generalmente para conteos de números de individuos (fallecidos por una causa determinada en este caso) por unidad de tiempo (el trienio 2003-2005) y de espacio (la ciudad de Tandil en este caso)

(2) Debe aclararse que esto no es posible hacerlo en este trabajo, debido a que no se pudo obtener información actualizada del Registro Civil como la correspondiente al período 2003-2005, lo que hubiera permitido el análisis de la evolución temporal.

(3) Sobre esta temática, pueden citarse trabajos como el de (Lopez Abente, 2001); (Benito Martinez, 2007); (Barcelo, 2008)

(4) Los Sistemas de Información Geográfica se han posicionado como una tecnología básica, imprescindible, y poderosa, para capturar, almacenar, manipular, analizar, modelar y presentar datos espacialmente referenciados (Moreno Jimenez, 2006)

(5) Se aclara que en ningún momento se ha violado el secreto estadístico, dado que los datos del domicilio de las personas se utilizaron sólo para el mapeo de la información.

(6) La labor de geocodificación está sujeta a muchos factores e incidencias, lo que conduce a que aparezcan inexactitudes, a que ciertos registros no sean encontrados, etc. En definitiva, es una operación algo delicada que

exige cierta prudencia, debido a que pueden obtenerse resultados inutilizables, si es que existiera un grado alto de error. Es necesario entonces, una labor de evaluación previa de la calidad, tanto de los datos de salida como los resultados obtenidos.

(7) La filosofía bayesiana no ha gozado de mucha aceptación hasta hace pocos años, en parte por el componente de subjetividad en relación a la distribución a priori que hay que introducir y en parte por la dificultad de su implementación en coste computacional. El desarrollo de los métodos bayesianos ha sido importante a partir de la década de los 80, favorecido en gran medida por los avances en reducción del coste y mayor rapidez de las computadoras. Actualmente los métodos bayesianos se han introducido en una gran cantidad de campos de la teoría y la práctica estadísticas. La virtud principal de estos métodos radica en que los procedimientos que los soportan consiguen un equilibrio adecuado entre sesgo y varianza, las dos componentes de toda estimación estadística.

(8) Aunque el teorema de Bayes es conceptualmente simple, su articulación dentro de un problema puede ser engorrosa, y explicarla en detalle desborda el propósito de la presente investigación. De hecho, a los efectos de este artículo, lo que básicamente interesa comprender es que se tiene a priori una opinión acerca de una proporción π (reflejo del conocimiento antes de realizar un experimento), que una vez realizado éste, el conocimiento acerca de π se modifica, y que el mecanismo utilizado para ello es el teorema de Bayes.

(9) Cabe aclarar que esta afirmación es válida para las ciudades pequeñas e intermedias, como la que se analiza en este trabajo. Los procesos espaciales suelen no poseer contigüidad y semejanza espacial en las ciudades grandes, donde la fragmentación y desigualdad espacial es más abrupta y severa.

Bibliografía

ALAMILLA LOPEZ, Norma Edith; JIMENEZ HERNANDEZ, Jose del Carmen: **Análisis bayesiano del modelo de poisson**. Temas de Ciencia y Tecnología, 12, 35, 2008:19-26

ANSELIN, Luc: **What is special about spatial data? Alternative perspectives on spatial data analysis**, 1989, pág. 63-77

BARCELLOS, Christovam: **Unidades y escalas en los análisis espaciales en salud**. Revista Cubana de Salud Pública, 29, 2003.

BARRET, Frank: **Finke's 1792 map of human diseases: the first world disease map?** Social Science & Medicine, 50, 2000, pág. 915-920.

BENACH, Joan (dir); MARTINEZ, Miguel Angel; BORRELL, Carmen et.al.: **Estudio geográfico de la mortalidad en España. Análisis de las tendencias temporales en municipios o agregados de municipios.** Fundación BBVA, 2007.

CLAYTON, David; BERNARDINELLI, Luisa: **Bayesian methods for mapping disease risk, en Geographical and Environmental Epidemiology: Methods for Small area studies.** Published on behalf of the world health organization regional office for europe by Oxford University Press, 1992.

CLAYTON, David; KALDOR, John: **Empirical bayes estimates of age-standardized relative risks for use in disease mapping.** Biometrics, 43 1987, pág. 671-681.

CLIFF, Andrew; ORD, Keith: **Model building and the analysis of spatial pattern in human geography.** Journal of the Royal Statistical Society: Series B. (Statistical Methodology), 37, 1975, pág. 297-348.

DOMINGUEZ ALONSO, Emma; GONZALEZ SUÁREZ, Roberto: **Análisis de las curvas receiver- operating characteristic: un método útil para evaluar procedimientos diagnósticos.** Rev. Cubana Endocrinología, 13 (2), 2002, pág. 173-253.

ELLIOT, Paul; WAKEFIELD, Jon et.al.: **Spatial epidemiology: methods and applications.** In: Elliot P, Wakefield Jc, Best NG, Briggs DJ (editors), 2000.

ENGLISH, David: **Geographical epidemiology and ecological studies,** en: Elliot, Paul; Cuzick John et al (editors), 1992

GARCÍA BALLESTEROS, Aurora. **La recuperación de la escala local en geografía de la población.** Investigaciones Geográficas (Mx) 43,2000, pág. 76-87.

GREENLAND, Sander; Henry MORGENSTERN: **Ecological bias, confounding, and effect modification.** International Journal of Epidemiology. International Epidemiological Association, 18 Nro 1, 1989, pág. 269-274.

HAVULINNA, Aki: **Bayesian Spatial and Temporal Epidemiology of Non- communicable diseases and mortality,** PhD Thesis, Depto of

Biomedical Engineering and Computational Science. Aalto University, 2011

HIND, Henry: **On the commercial progress and resources of the central british america; the lake winnipeg and saskatch ewan district.** J.R Statist Soc., 27 1864, pág. 82-105.

INSTITUTO NACIONAL DE ESTADÍSTICA Y CENSOS (INDEC). Censo Nacional de Población y Viviendas, 2010.

LAWSON, Andrew; BOHING, Dankmar; BIGGIERI, Annibale, et.al.: **Disease mapping and its uses**, en: A.B Lawson; D. Boehing; E. Lasaffree; A. Biggeri; Viel JF; Bertolline R (eds). **Disease Mapping and risk assesment for Public Health**, 1999.

LOPEZ ABENTE ORTEGA, Gonzalo; POLLAN SANTAMARÍA, Mariana: **Atlas de mortalidad de la comunidad valenciana 1991-2000.** Generalitat Valenciana. Censelleria de Sanitat, 2005.

MANCHO CORCUERA, Jesús: **Técnicas de estimación en áreas pequeñas.** Eustat, Junio 2002.

MARTINEZ BENITO, Miguel Angel; BORREL, Carmen et.al.: **Estudio geográfico de la mortalidad en españa. Análisis de tendencias temporales en municipios o agregados de municipios**, 2007.

MINISTERIO DE SALUD Y AMBIENTE DE LA NACIÓN. **Atlas de mortalidad de Argentina: trienios 1990-1992/1999-2001** Dirección de Epidemiología. Gestión y Políticas de Salud. Organización Panamericana de Salud. Representación Argentina. [CD- Rom] Universidad Nacional de Lanús. Maestría en Epidemiología, 2004

MORENO JIMENEZ, Antonio (coord): **Sistemas de análisis de la información geográfica. Manual de Autoaprendizaje con ArcGis.** Ed. Ra-Ma, España 2006.

PUJOL, Celine. **La relación plaguicidas- salud- ambiente, en un caso testigo: Isla Verde, Provincia de Córdoba.** Tesis para aspirar al grado de Doctor en Demografía, dirigida por la Dra. María Francis Álvarez. Facultad de Ciencias Económicas. Universidad Nacional de Córdoba, 2012

RAO, Jon: **Some recent advances in model- based small area estimation.** Survey Methodology, 25, Nro 2,1999, pág. 175-186.

RAO, Jon: **Small area estimation.** Jonh Wiley & Sons. SAS institute, inc. Annals of statistics, 6, 2003, pág. 461-464.

RAVENSTEIN, Ernest: **Statistics at the graphical congress of paris.** Journal of the statistical Society of London, 38 1875, pág. 422-429.

RICHARDSON, Sylvia, et. Al.: **Comparision of relative risks obtained in ecological and individual studies.** Int J Epidemiol, 16 1987, pág. 11-20.

SILVA AYÇAGUER, Luis Carlos; BENAVIDES RODRÍGUEZ, Alina; VIDAL RODEIRO, Carmen Lucía: **Análisis espacial de la mortalidad en áreas pequeñas. El enfoque bayesiano.** Revista cubana de Salud Pública, 29, 2003, pág. 314-322.

STUDENT: **The elimination of spurious correlation due to position in time or space.** Biometrika, doi: 10.1093/biomet/10.1.179 -1980

TULLOCH, Alexander. **On the sickness and mortality among the troops in the west indies.** J. R. Statist, Soc., 1, 1838, 129-14.

WAKEFIELD, John; MORRIS, Sara: *The Bayesian modeling of disease risk in relation to a point source*, en: **Journal of the American Statistical Association**, (2001), 96(453), 77-91.

Fecha de recepción: 13 de marzo de 2014
Fecha de aprobación: 12 de mayo de 2014