Nº 74, julio-diciembre de 2020, pp. 71-84.
ISSN: 0213-4691. eISSN: 1989-9890.
DOI: 10.14198/INGEO2020.SAPE

Investigaciones Geográficas
Instituto Interuniversitario de Geografía
Universidad de Alicante

Cita bibliográfica: Sánchez-Díaz, G., Alonso-Ferreira, V., Posada de la Paz, M., & Escobar, F. (2020). Nuevas perspectivas sobre el Problema de la Unidad Espacial Modificable (PUEM) en relación con la representación cartográfica de enfermedades raras. *Investigaciones Geográficas,* (74), 71-84. https://doi.org/10.14198/INGEO2020.SAPE

Nuevas perspectivas sobre el Problema de la Unidad Espacial Modificable (PUEM) en relación con la representación cartográfica de enfermedades raras

New insights around the Modifiable Areal Unit Problem (MAUP) in its relation to cartographic representation of rare diseases

Germán Sánchez-Díaz[1]
Verónica Alonso-Ferreira[2]
Manuel Posada de la Paz[3]
Francisco Escobar[4*]

Resumen

En el marco del análisis espacio-temporal de la Enfermedades Raras (ER) en España, los efectos derivados del Problema de la Unidad Espacial Modificable (PUEM) condicionan la interpretación de los resultados. Este estudio tiene como objetivo comparar indicadores de mortalidad y visualización cartográfica para recomendar la unidad geográfica de trabajo óptima según población, área cubierta y el número de casos de la enfermedad. Se estudiaron los fallecimientos debidos a una ER (Huntington) en tres niveles de agregación espacial; el municipio, la comarca y la provincia. Se diseñó una ratio de adyacencia para observar el efecto de las relaciones de vecindad entre las tres unidades administrativas tratadas. Se calcularon indicadores epidemiológicos de mortalidad, así como indicadores locales de asociación espacial en cada nivel de agregación. Los resultados se representaron cartográficamente utilizando intervalos definidos por el usuario para comparar de forma visual las diferencias estadísticas y cartográficas. El PUEM fue particularmente visible trabajando con datos agregados de ER. Las comarcas representaron el mayor nivel de estabilidad en la ratio de adyacencia así como la unidad geográfica óptima en términos de resolución espacial y variabilidad de la información representada cartográficamente. Este proceso de ayuda en la elección de la escala de trabajo puede ser extrapolado a otras enfermedades o niveles de agregación, como paso previo a análisis epidemiológicos más avanzados.

Palabras clave: PUEM; datos agregados; epidemiología; unidad geográfica; cartografía; geografía médica; enfermedades.

Abstract

This study analyses the effect of the modifiable areal unit problem (MAUP) comparing mortality data of a specific rare disease (Huntington) in three levels of spatial aggregation in Spain. The objective is to compare mortality indicators and cartographic visualisations in order to soundly advise on the optimal

1 Instituto de Investigación de Enfermedades Raras (Instituto de Salud Carlos III); Centro de Investigación Biomédica en Red de Enfermedades Raras (Instituto de Salud Carlos III); Departamento de Geología, Geografía y Medio Ambiente (Universidad de Alcalá), España. german.sanchez@uah.es
2 Instituto de Investigación de Enfermedades Raras (Instituto de Salud Carlos III); Centro de Investigación Biomédica en Red de Enfermedades Raras (Instituto de Salud Carlos III), España. valonso@isciii.es
3 Instituto de Investigación de Enfermedades Raras (Instituto de Salud Carlos III), España. mposada@isciii.es
4 Departamento de Geología, Geografía y Medio Ambiente (Universidad de Alcalá), España. francisco.escobar@uah.es. * Autor para correspondencia

Recepción: 05/06/2020. **Aceptación:** 13/11/2020. **Publicación:** 17/12/2020.

Este trabajo se publica bajo una licencia de Creative Commons Reconocimiento 4.0 Internacional.
© Germán Sánchez-Díaz, Verónica Alonso-Ferreira, Manuel Posada de la Paz, Francisco Escobar

aggregation level according to population, covered area, and number of cases. We designed an adjacency ratio to observe the effect of neighbourhood relationships among three geographic units: province; district (comarca); municipality. For each level of aggregation, we performed epidemiological indicators of mortality as well as local indicators of spatial association. Maps were plotted with user-defined intervals to compare visual and statistical differences. MAUP-related effects are particularly noticeable in relatively infrequent events such as rare diseases. We found that district displayed the highest indicator for stability in the adjacency ratio and showed optimal characteristics for spatial resolution and the amount of information revealed through plotting. This helps in the choice of the working scale that can be used with other diseases or the levels of aggregation as a first step for more advanced epidemiological analyses.

Keywords: MAUP; aggregated data; epidemiology; geographic unit; cartography; medical geography; diseases.

1. Introducción

La relación entre las disciplinas de la geografía y la medicina queda datada desde la antigüedad, como se puede comprobar en los postulados de Hipócrates en su tratado *Sobre las aires, aguas y lugares* del siglo IV a.C (Hipocrates, Thucydides, Clifton & Ephesus, 1752). El filósofo griego asociaba el hábitat (componente geográfico) con los patógenos (componente médico) como una forma de explicar el desarrollo de las enfermedades y la muerte en el ser humano. Esta teoría ambiental supuso la base de lo que después se conocería como la geografía médica, que tuvo su despegue gracias a la institucionalización de esta subdisciplina durante el periodo de la ilustración (Urteaga, 1980). El fundador de la epidemiología moderna, John Snow, asoció la localización espacial de los pozos de agua en la ciudad de Londres con las muertes producidas por una epidemia de cólera, mediante la realización de un mapa que mostraba el emplazamiento geográfico de los pozos y los fallecimientos producidos (Snow, 1855). Este análisis espacial y las medidas preventivas efectuadas permitieron un drástico descenso de las muertes por cólera estableciendo así un importante hito en epidemiología espacial, íntimamente ligada a la geografía médica.

Ya en el siglo XX, entre los años 1950 y 1960, de la mano de la revolución cuantitativa, la geografía se redefine como una ciencia espacial en busca del desarrollo de metodologías sistemáticas. En los años 60 surgen los Sistemas de Información Geográfica (SIG) y pronto comienzan a ser progresivamente adoptados en el sector de la salud (Olaya, 2016; Shaw & McGuire, 2017). Existen numerosos ejemplos de su aplicación en ese sector, y van desde la planificación sanitaria, el análisis epidemiológico de brotes infecciosos, la prestación de servicios de emergencia a campañas de promoción de la salud (Lai, So & Chan, 2009). Sin embargo, hay que señalar, que en el caso de las enfermedades raras (ER), no se ha producido esta adopción con la misma intensidad (Gómez-Rubio & López-Quílez, 2010).

Las ER son aquellas que afectan a menos de 5 personas por cada 10.000 habitantes en la Unión Europea (UE) (Decision Nº 1295/1999/EC). Hasta la fecha hay identificadas entre las 6.000 y 8.000 ER y se estima que afectarían a entre un 6 y un 8% de la población de la UE, siendo por ello reconocidas como un problema de salud pública (Posada, Alonso & Bermejo, 2016; Groft & Posada, 2017). La integración de los SIG con las ER no está todavía en una fase avanzada debido a que la mayor parte de la investigación está enfocada a enfermedades más comunes (Alonso *et al.*, 2011; Sánchez-Díaz *et al.*, 2016). Este retraso se explica por las dificultadas de análisis que derivan de su baja frecuencia, llevando a un exceso de ceros en los resultados estadísticos al agruparlos en unidades territoriales muy desagregadas (Posada *et al.*, 2010). Esto en muchos casos supone una escasa variabilidad en las representaciones cartográficas, especialmente al realizar tareas metodológicas propias de la epidemiología espacial como la suavización de indicadores de salud. Además, los sistemas de clasificación de enfermedades mediante códigos están normalmente dirigidos a aquellas enfermedades que son más frecuentes, lo que supone un obstáculo importante para los estudios epidemiológicos de ERs al verse limitados por la disponibilidad de datos y su calidad (Rytkönen, 2004; Grady & Enander, 2009; Posada *et al.*, 2010).

La mortalidad es uno de los indicadores epidemiológicos más utilizados para el estudio de las enfermedades. Ayuda a conocer su curso y gravedad, incrementa la visibilidad de la enfermedad y pone de relieve la magnitud que tienen las desigualdades en salud pública (Posada *t al.*, 2010). En las últimas décadas, se han producido grandes avances en los estudios espacio-temporales que observan tendencias en el tiempo y comportamiento geográfico de la mortalidad, así como su relación con otras variables, con aplicaciones en diversos campos como el medio ambiente, la economía o la salud (Lawson, Banerjee, Haining & Ugarte,

2016; Meyer, Held & Höhle, 2016). Por ejemplo, en el caso de la epidemiología, los más conocidos son los estudios espacio-temporales para el monitoreo y control de brotes de enfermedades infecciosas (Waller & Gotway, 2004; Meliker & Sloan, 2011; Lawson *et al.*, 2016; Redondo-Bravo *et al.*, 2018).

Por otra parte, las jerarquías espaciales en las que se agregan los datos no reflejan siempre su naturaleza. Esta alteración es conocida como el Problema de la Unidad Espacial Modificable (PUEM) (Openshaw, 1984). De acuerdo al PUEM, cuando un evento geográfico continuo es representado mediante unidades artificialmente creadas como las secciones censales o los municipios, el tamaño del área, la configuración de su forma y, especialmente la escala, pueden ofrecer diferentes resultados, y diferentes interpretaciones (Labib, Lindley & Huck, 2020).

El PUEM ha sido ampliamente revisado en la literatura científica desde mediados de la década de los ochenta, pero respecto a las ERs hay una cierta escasez. Se ha estudiado, entre otros, en el ámbito de la salud cardiovascular (Cebrecos, Domínguez-Berjón, Duque, Franco & Escobar, 2018), en relación a la pandemia de la COVID-19 (Wang & Di, 2020) y en otros estudios que utilizan datos espaciales como la política (Menon, 2012) o la economía (Johnston, 2002). Es necesario añadir una visión sobre esta cuestión e investigar el PUEM en relación a las ER ya que su agregación de datos no da siempre respuesta a los enfoques analíticos tradicionales. En la investigación de ER, la agregación de datos es necesaria por razones que no suelen darse en otros fenómenos geográficos: por un lado, preservar la confidencialidad de las personas afectadas (Zandbergen, 2014) y por otro, la infrecuencia de casos que provoca que, a una escala con mayor detalle, los datos sean más escasos en las unidades geográficas (Posada *et al.,* 2010). Sin embargo, el hecho de pasar de datos recogidos a nivel individual (personas fallecidas) a una agregación en unidades geográficas, clasificada en intervalos, es una manifestación de la llamada falacia ecológica, y el PUEM constituye una derivación de ésta. La falacia ecológica se da cuando datos individuales son agrupados en unidades geográficas, normalmente definidas por límites administrativos, que no tienen relación directa con la variable agregada, asumiendo así una distribución homogénea a través de toda la unidad geográfica, y, como señalan Bosque, Chuvieco & Santos-Preciado (1986), pudendo dar lugar a interpretaciones incorrectas. Esto nos hace plantearnos la cuestión sobre qué nivel de agregación espacial sería el más adecuado para la representación cartográfica de una ER teniendo en cuenta el PUEM. Las unidades de gobierno local (municipio) o las unidades de código postal tienen un alto grado de heterogeneidad tanto en extensión como en población, y su adopción resultaría en un exceso de unidades geográficas sin casos. Por otro lado, un número menor de unidades que agreguen mayor cantidad de población, como las comarcas, provincias o Comunidades Autónomas, podrían debilitar el trabajo de análisis geográfico y visualización de resultados (Moore & Carpenter, 1999; Arab, 2015). En algunos estudios se ha propuesto la aplicación de algunos criterios para facilitar la selección de la unidad geográfica de análisis más adecuada en un estudio epidemiológico: población, tamaño de las unidades, distribución del evento a estudiar o su relevancia biológica (Arsenault, Michel, Berke, Ravel & Gosselin, 2013).

De acuerdo a todo lo expuesto, este estudio tiene como objetivo analizar el efecto de la agregación espacial de datos de mortalidad debida a ER en diferentes niveles utilizando la enfermedad de Huntington (EH) como ejemplo. Se busca mostrar el proceso usado para decidir el mejor nivel de agregación de acuerdo al número de casos que contenga esa ER, criterios de población y de área de cada uno de ellos mediante resultados estadísticos y de visualización cartográfica.

2. Metodología

Este estudio fue llevado a cabo en España, cuya extensión aproximada es de 500.000 km^2 y cuenta con una población aproximada de 47 millones de habitantes en 2019 según el Padrón Continuo publicado por el Instituto Nacional de Estadística (INE) (INE, 2020). El territorio español se divide en comunidades autónomas, y a su vez en 50 provincias y dos ciudades autónomas según la división administrativa que data del año 1833 (aunque se han producido posteriores modificaciones) (Moreno, 1994; Gómez-Díaz, 2008). Cada provincia contiene un número variable de municipios (desde 34 a 371) sumando más de 8.000 para el conjunto del país, los cuales constituyen el nivel más básico de gobernanza y con personalidad legal independiente en España. Las comarcas son entidades de mayor extensión que agrupan municipios vecinos de acuerdo a criterios históricos y geográficos. Sin embargo, solo cuentan con un carácter administrativo legalmente reconocido en las Comunidades Autónomas de Cataluña, Aragón, País Vasco y Castilla y León (Instituto Geográfico Nacional, 2018). A pesar de ello, es una subdivisión territorial que, a priori,

parece más conveniente, para el análisis de datos sobre ER, que la provincia, por excesivamente grande y que el municipio, por la razón contraria y por tanto ser contenedor, en su mayor parte, de cero casos de fallecimientos por ER. En este estudio se utilizó la clasificación de comarcas agrarias proporcionada por el Ministerio de Agricultura, Pesca y Alimentación (MAPAMA) (Fernández *et al.*, 2011).

La información espacial a nivel municipal a escala 1:25000 se descargó del Centro Nacional de Información Geográfica (CNIG), perteneciente al Instituto Geográfico Nacional. La capa de comarcas (326 polígonos) en formato *shapefile* de tipo vector, así como la de provincias (52 polígonos) fue generada mediante la sucesiva agregación de polígonos pertenecientes a la capa de municipios.

La Figura 1 muestra las unidades geográficas de análisis mencionadas, mostrando el modelo de adyacencias de menor a mayor detalle (desde provincias hasta municipios) y el cual será utilizado después en el proceso epidemiológico de suavizado espacial.

Figura 1. Unidades geográficas españolas (color gris) incluidas en este estudio (incluye número total)

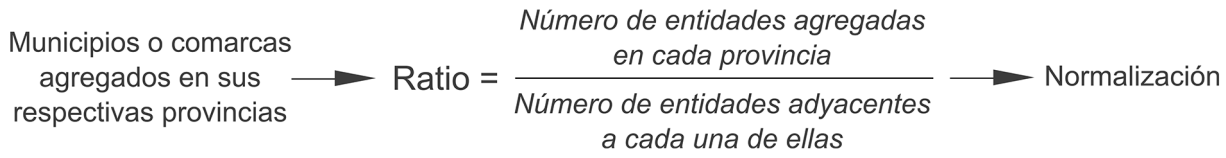
Fuente: CNIG. Elaboración propia

Los datos de mortalidad de los 15 años que abarca este estudio (1999-2013) se obtuvieron de las estadísticas de defunción del INE según la causa de muerte. La EH fue seleccionada como ejemplo de ER según la causa de fallecimiento, utilizando para ello el código G10 de la décima revisión de la Clasificación Internacional de Enfermedades (CIE-10). Esta selección dio lugar a un total de 1.090 fallecimientos para el periodo de estudio. Los datos de población para el periodo 1999-2013 fueron obtenidos del Padrón Municipal de Habitantes, publicado por el INE, y agregados posteriormente a nivel comarcal y provincial.

Los municipios y las comarcas se agruparon en provincias respectivamente y se calculó para cada uno de ellos una ratio de adyacencia (es decir, en el caso de las comarcas se dividió el número total de

comarcas contenidas en una provincia entre el número promedio de entidades adyacentes a cada comarca). Ambas ratios fueron normalizadas en una escala de -2 a 2 para poder así permitir la comparación entre las distribuciones espaciales de comarcas y municipios (Figura 2).

Figura 2. Formulación del algoritmo de cálculo de la ratio de adyacencia



Elaboración propia

La Razón de Mortalidad Estandarizada (RME) es un indicador muy utilizado en estudios de epidemiología espacial porque compara los fallecimientos observados y los que serían esperables en cada unidad espacial teniendo en cuenta la tasa de una población de referencia (en este caso la española) (Royo & Moreno, 2009). Para el estudio presente se calcularon las muertes esperadas mediante el método estandarización indirecta por grupos de edad de 5 años y posteriormente la ratio entre los fallecimientos observados y los esperados. En nuestro caso, las RME fueron calculadas para provincias, comarcas y municipios para el periodo de estudio. Los resultados obtenidos de la RME se dividieron en intervalos definidos por el usuario, siguiendo una clasificación basada en valores en torno a 1,00. El intervalo que recoge el rango de 0,80-1,20 representa el riesgo de fallecimiento normal respecto a lo esperado para España, los intervalos por encima de 1,20 mostrarán un riesgo de fallecimiento superior al esperado mientras que, aquellos por debajo de 0,80 mostrarán un riesgo inferior respecto al total español. Este tipo de intervalos predefinidos, ampliamente utilizados en epidemiología, tienen su razón de ser en motivos de visibilidad así como por la necesidad de dividir los valores de la variable en números pares de intervalos alrededor del valor medio (0,80-1,20) (Cauvin, Escobar & Serradj, 2010). Hay que tener en cuenta que también se ha añadido un intervalo extra para las RME con valor igual a cero, las cuales serán resultado de aquellas unidades geográficas que no tienen registrado ningún caso. Asimismo, se calcularon los intervalos de confianza (IC) al 95% para las RME y se evaluó la distribución de las mortalidades para cada nivel de agregación mediante un test de *chi* cuadrado (Greenwood & Nikulin, 1996).

La suavización del indicador de la RME es un método utilizado para reducir la distorsión producida por el exceso de ceros, así como los valores extremos producidos por la población de las unidades geográficas. Las RME suavizadas se estiman teniendo en cuenta no solo el número de casos de fallecimientos observados y esperados en cada unidad geográfica, sino también los valores de las unidades espaciales vecinas (criterio de adyacencia geográfica) para poder así detectar patrones espaciales de riesgo o descartarlos. Para suavizar la RME, se creó en primer lugar para cada nivel de agregación una matriz de adyacencias de tipo rook, la cual consiste en una contigüidad basada en compartir un borde físico entre entidades (Smirnov & Anselin, 2001). Tras la creación de las matrices, se utilizó el modelo condicional autorregresivo propuesto por Besag, York y Molliè para llevar a cabo la estimación del riesgo teniendo en cuenta los efectos espaciales y de heterogeneidad (Besag, York & Molliè, 1991). Para la producción cartográfica, los resultados obtenidos se clasificaron de la misma forma que la RME para facilitar la comparación visual y estadística.

También se examinó la autocorrelación espacial y la existencia de clusters para comparar los resultados en cada nivel de agregación. Se buscó conocer cuál de ellos aporta una información más relevante para posteriores estudios o para representar los datos en un software cartográfico. Para evitar sesgos en el resultado, test locales y globales solo fueron llevados a cabo con la RME y no con la RME suavizada. El índice de Moran fue calculado para evaluar la autocorrelación de los valores de la RME en provincias, comarcas y municipios, mientras que los indicadores locales de asociación espacial (LISA) se aplicaron para buscar clusters de la RME (Moran, 1950; Anselin, Syabri & Kho, 2006). El análisis local permite determinar la existencia de unidades geográficas adyacentes que comparten valores similares, pudiendo detectar, por ejemplo, áreas con altos valores o con desigualdades.

Las RME fueron calculadas con el software Stata mientras que, el posterior proceso de suavizado se llevó a cabo con el software de programación R, haciendo uso del paquete R-INLA (Integrated Nested Laplace Approximation). Este proceso de suavizado es el más rápido en computación que los métodos de MonteCarlo (McMC) y más ventajoso en lo referente al ajuste del modelo e inferencia (Rue *et al.*, 2016).

El software GeoDA se utilizó para el análisis de clusters y autocorrelación mientras que, la producción mapas se realizó con el software ArcGIS.

3. Resultados

Como se muestra en la Tabla 1, las entidades tienen diferentes características de tamaño y población para cada nivel de agregación, especialmente en provincias y municipios. Las diferencias de extensión en las unidades son importantes (de hasta 10 veces en provincias, 20 veces en comarcas y de más 5.000 veces en municipios), mostrando por tanto mayor uniformidad en provincias y mayor heterogeneidad en municipios. Del mismo modo, también fueron evidentes las diferencias en población: desde hasta 70 veces entre provincias, 30.000 veces entre comarcas y hasta 1 millón de veces entre los diferentes municipios. En resumen, la extensión y población de las unidades geográficas es muy diversa, no solo comparando entre los tres niveles de agregación sino entre las entidades que los componen internamente.

Tabla 1. Estadísticas de área y población para los tres niveles de agregación analizados

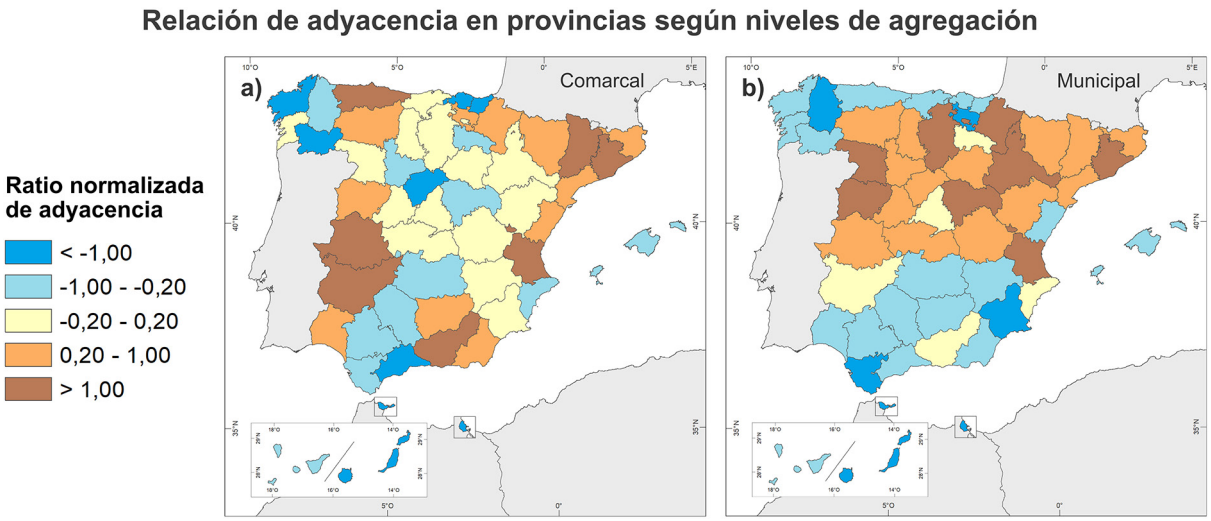
		Provincia	Comarca	Municipio
	n	52	326	8123
Área	Promedio (km^2)	9.725	1.542	62
	Desviación estándar	5.095	896	92
	Min (km^2)	13	13	0,03
	Max (km^2)	21.751	5.396	1.750
Población*	Promedio (habitantes)	853.588	136.155	5.468
	Desviación estándar	1.074.191	342.422	15.521
	Min (habitantes)	71.379	1.625	5
	Max (habitantes)	5.956.777	4.540.006	3.116.900

*Población promedio para el periodo 1999-2013

Fuente: CNIG. Elaboración propia

La comparación de las ratios de adyacencia normalizadas mostró que las comarcas son más estables que los municipios (Figura 3). Mientras que las comarcas obtuvieron ratios más equilibrados, los municipios tuvieron pocos casos en el rango intermedio (14 provincias en el análisis comarcal frente a 5 en el análisis municipal con un valor entre 0,20 y -0,20). Cuanto mayor es la diferencia en las ratios interprovinciales, más inestables serán posteriormente los resultados del suavizado espacial de la RME.

Figura 3. Ratios normalizadas de adyacencia en provincias. Agregación por: a) comarca; b) municipio



Elaboración propia

Como se puede observar en la Tabla 2, la distribución de los valores de la RME es significativamente diferente dependiendo del nivel de agregación utilizado (p<0,001). Hay que indicar que para observar la variación de los resultados en los tres niveles de agregación se han utilizado los mismos intervalos de clasificación. Además, es muy destacable el número de municipios que no han contabilizado ningún caso en los cuales el valor de la RME es cero (96,44%). Entre los intervalos seleccionados, las provincias registraron la distribución más uniforme de las RME mientras que, en comarcas el 42,94% de los resultados fueron clasificados como sin casos. El número de unidades y porcentajes fuera del IC del 95% fue del 10% y del 5% en provincias y comarcas respectivamente (p<0,001), mientras que en municipios no se obtuvieron valores fuera de los IC.

Tabla 2. Distribución de los valores de Razón de Mortalidad Estandarizada (RME) y RME suavizada después del proceso de modelado bayesiano para la EH según nivel de agregación (periodo 1999-2013). Se indica número de casos y % (entre paréntesis) para cada indicador

	Provincia		**Comarca**		**Municipio**	
Intervalo	**RME**	**RME suavizada**	**RME**	**RME suavizada**	**RME**	**RME suavizada**
Sin casos (0,00)	4 (7,69)	-	140 (42,94)	-	7835 (96,44)	-
< 0,60	7 (13,46)	1 (1,92)	31 (9,51)	7 (2,15)	11 (0,14)	0 (0,00)
0,60 – 0,80	10 (19,23)	12 (23,08)	16 (4,91)	35 (10,74)	7 (0,09)	0 (0,00)
0,80 – 1,20	15 (28,85)	30 (57,69)	39 (11,96)	202 (61,95)	17 (0,21)	8123 (100)
1,20 – 1,40	2 (3,85)	4 (7,69)	11 (3,37)	41 (12,58)	8 (0,10)	0 (0,00)
> 1,40	14 (26,92)	5 (9,62)	89 (27,31)	41 (12,58)	245 (3,02)	0 (0,00)
Total	52 (100)		326 (100)		8123 (100)	

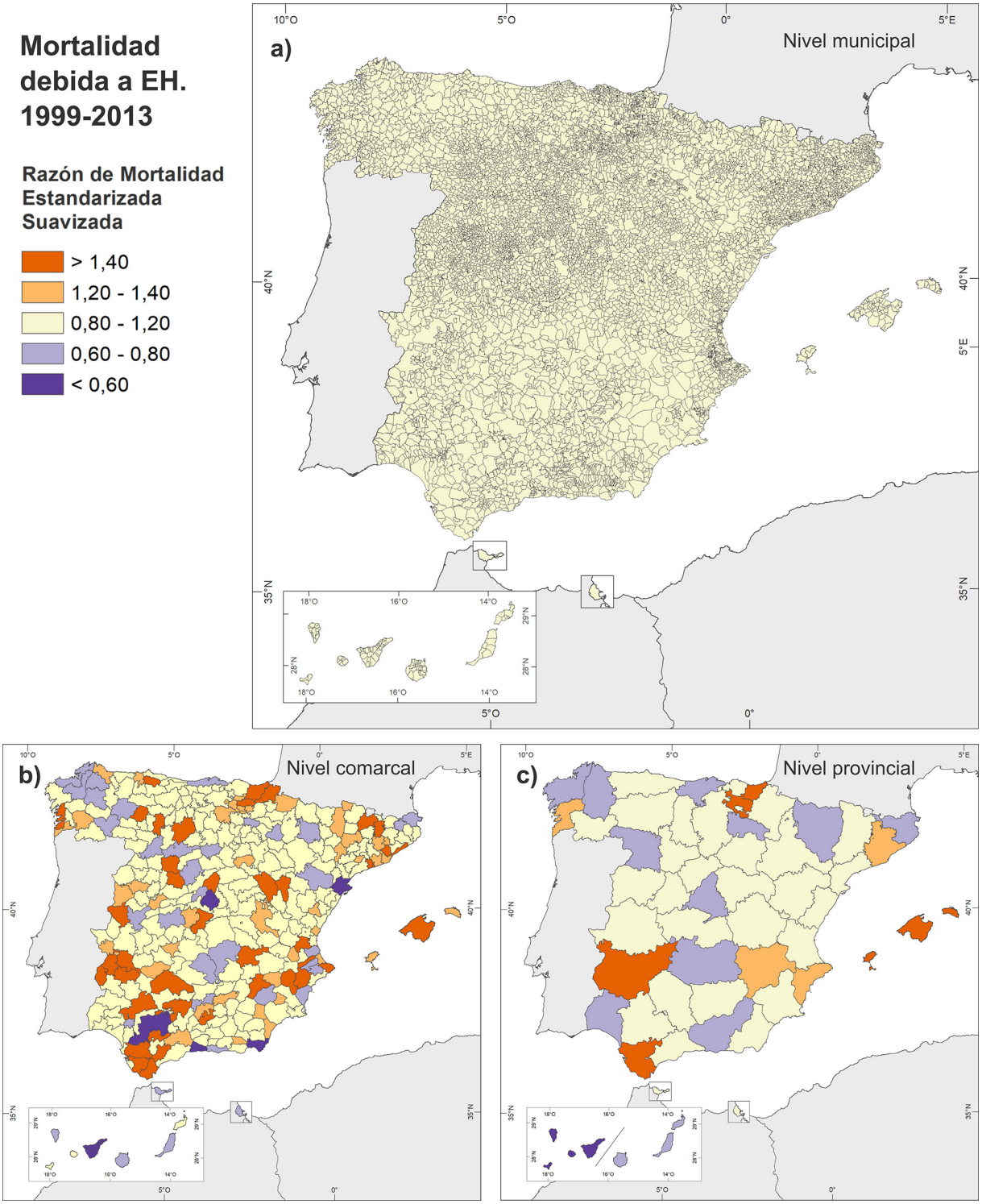
Elaboración propia

Los resultados cartográficos de las RME suavizadas para cada nivel de agregación después del proceso de modelado bayesiano se muestran en la Figura 4. En un primer análisis visual, el procedimiento de suavización de las RME difumina las diferencias de mortalidad previamente detectadas a nivel municipal. Dependiendo del tamaño de las unidades, el porcentaje de unidades con un riesgo de fallecimiento significativamente alto o bajo para las RME suavizadas varió del 26% en provincias al 19% en comarcas y al 0% en municipios (p<0,001).

Respecto al índice de Moran, no se encontró autocorrelación positiva o negativa en ninguno de los tres niveles de agregación. De hecho, la puntuación z obtenida para todos los niveles señala un patrón de dispersión en los resultados de la RME en España, no siendo significativo ninguno de los resultados. Esto significaría que no se ha encontrado para este ejemplo de la enfermedad de Huntington, en la geografía española, una agrupación de valores de riesgo superior o inferior de su riesgo de fallecimiento según la información de este índice espacial.

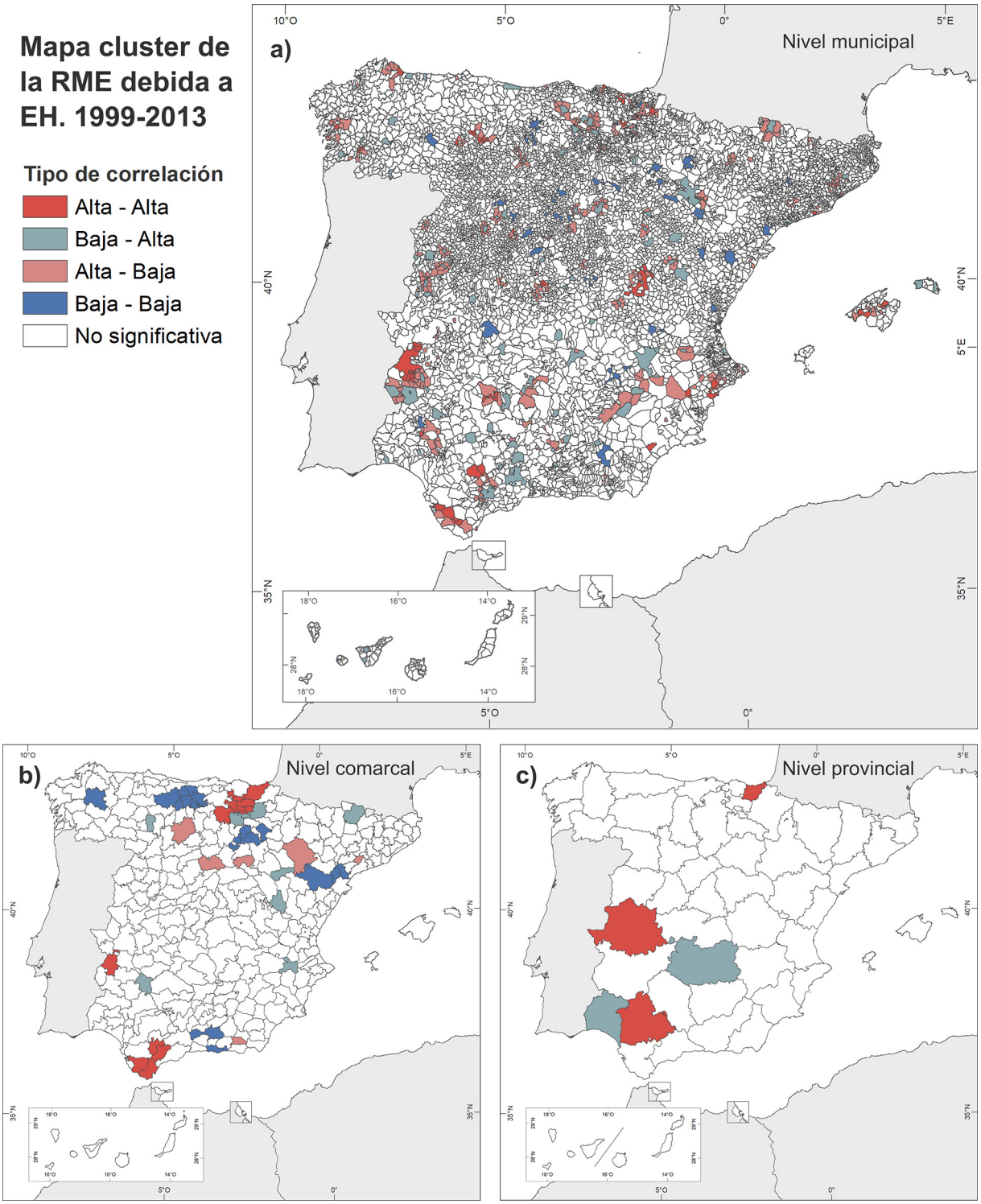
Por último, en la Figura 5, se representa cartográficamente el índice de Moran local o LISA dónde se puede observar los distintos niveles de asociación estadística de autocorrelación espacial para los valores de la RME. Los valores de autocorrelación "Alto-Alto" en el mapa municipal se han observado en el oeste del país, mientras que a nivel comarcal y provincial se ha detectado en el norte y sur. Cluster de valores "Bajo-Bajo" están distribuidos por todo el país en el mapa municipal, a nivel comarcal en el norte y suroeste mientras que, no se detectaron a nivel provincial. "Alto-Bajo" es resultado de una entidad con un valor atípico alto rodeado por valores bajos (encontrado de forma dispersa a nivel municipal, principalmente en la mitad norte a nivel comarcal, y sin representación a nivel provincial). Por último, entidades con valor de autocorrelación "Bajo-Alto" (valor significativamente bajo de la RME rodeado de altos) están repartidos aleatoriamente a nivel municipal, de forma más predominante en la mitad norte en comarcas y en el suroeste de la península a nivel provincial. Los porcentajes de valores estadísticamente significativos fueron los siguientes: 9,61% a nivel provincial (5 de 52 entidades), 13,50% a nivel comarcal (44 de 326 entidades) y 8,10% a nivel municipal (658 de 8.123 entidades).

Figura 4. RME suavizadas para la EH entre 1999-2013. Niveles: a) Municipal; b) Comarcal; c) Provincial



Elaboración propia

Figura 5. Mapa de cluster de la RME debida a EH en el periodo 1999-2013 a tres niveles: a) Municipal; b) Comarcal; c) Provincial



Elaboración propia

4. Discusión

Este estudio ha llevado a cabo un análisis estadístico y descriptivo de la mortalidad de la EH a nivel municipal, comarcal y provincial en España. Los resultados han sido mostrados en forma de tablas y

cartografía facilitando la observación de las diferencias entre los tres niveles de agregación. Además de la comparación de las tasas calculadas, saber si los resultados obtenidos son estadísticamente significativos ayuda en la toma de decisiones y de reducción de incertidumbre asociado a datos inestables. Respecto a esto, la cartografía es una parte integral de los estudios espaciales y temporales para complementar los resultados numéricos (Dykes & Brunsdon, 2007).

Como puede verse en las tablas y figuras, los datos de fallecimientos obtenidos del INE a nivel municipal fueron posteriormente añadidos en dos niveles de mayor agregación (comarca y provincia) para poder comparar los resultados. Al hacer esto, los patrones espaciales sufren cambios como consecuencia del PUEM, algo que se presenta al agregar datos en unidades administrativas de diferente distribución y tamaño (Kanaroglou, Delmelle & Páez, 2015). El PUEM, por tanto, también está presente cuando se analizan las ER desde un punto de vista geográfico, donde específicamente se visibiliza el efecto de la variación de la escala. Más que proponer soluciones al MAUP, el objetivo de este estudio ha sido su exploración e identificación para servir a otros investigadores que trabajen con sus propios conjuntos de datos, y que puedan determinar cuál es la unidad geográfica más estable dentro en su contexto, asegurando el trabajo cartográfico más óptimo para la visualización de datos. Respecto a esto, el número total de unidades geográficas es un elemento fundamental para numerosos test de tipo estadístico en un contexto geoespacial (Goodwin, Dykes, Slingsby & Turkay, 2015; Lawson, 2006). Disponer de datos geográficos de fallecimientos únicamente a nivel municipal, sin más detalles sobre la residencia del individuo, responde a la necesidad de preservar la confidencialidad de los datos personales de los fallecidos, unos datos que son especialmente sensibles en aquellos municipios poco poblados (Goodwin, *et al.,* 2015). En esos casos la anonimidad es difícil de mantener haciendo necesaria la adopción de alguna de las medidas señaladas anteriormente.

En lugar del cálculo de tasas crudas, se ha preferido estimar las RME porque la mortalidad es un indicador muy relacionado con los grupos de edad y, entidades con población muy envejecida pueden distorsionar los datos al compararlos con otras que no comparten esas mismas características. De hecho, incluso la RME puede presentar cierta inestabilidad cuando se trabaja con un bajo número de casos por unidad geográfica, como ocurre en nuestro estudio (Ugarte, Ibáñez & Militino, 2006). El enorme número de unidades geográficas sin casos y la gran heterogeneidad de la población en municipios producen muchos ceros en los resultados y valores extremos pudiendo generar una falta de robustez en los resultados (Posada *et al.*, 2010).

La ratio de adyacencia es un índice propuesto con el fin de entender la relación entre el número de entidades y sus áreas contiguas. La obtención de resultados extremos en esta ratio, alejados de los intervalos centrales, señala la irregularidad existente en las subdivisiones españolas, siendo estas desigualdades más visibles en municipios. Esto, combinado con el bajo número de casos de cada entidad generará un resultado difuminado en el suavizado espacial de la RME. En el caso de datos municipales, la variabilidad desaparece haciendo difícil la búsqueda de patrones espaciales ya que, cuando una tasa es suavizada espacialmente el resultado es influenciado por el valor de sus unidades vecinas. Teniendo en cuenta que cada municipio es adyacente con un promedio de 5,94 municipios y que, el 96% de ellos no tiene casos observados, la RME suavizada de esos municipios con valores extremos será suavizada. Mientras el número de unidades adyacentes es ligeramente inferior en comarcas y provincias (5,50 y 4,46 respectivamente), la diferencia reside en que el porcentaje de unidades sin casos observados es muy inferior. En nuestro caso, hemos observado que los municipios no serían el nivel de agregación más adecuado para el estudio de patrones espaciales de ER con este bajo número de casos en España o en otros países con una división administrativa similar. La pérdida de información es menos severa en comarcas y provincias, ya que el proceso de suavizado daría un mejor resultado con unidades de agregación menos detalladas como se señala en otros estudios (Cockings & Martin, 2005). En el caso de las provincias, el inconveniente es su baja resolución espacial ya que los datos están muy agregados en pocas entidades en comparación con los otros dos niveles analizados. También hay que tener en cuenta que los niveles de agregación disponibles no están diseñados con la intención de servir para la investigación de ER, por lo que pueden producirse distorsiones al interpretar los resultados obtenidos (Elliot & Wartenberg, 2004).

El análisis LISA mostró que el porcentaje de unidades geográficas con valores significativos fue menor en los municipios que en las provincias, presentando menor variabilidad en las áreas sin casos registrados a nivel municipal. También es cierto que, al observar una agrupación de casos de enfermedades a gran escala (como la municipal), al agregarlos en una escala más amplia pueda perderse esta información detallada.

El bajo número de casos en municipios (o secciones censales en el caso de tener disponibles esos datos) suele ser una limitación general en los estudios epidemiológicos, produciendo la necesidad de

agregar los datos a un mayor nivel. En algunos casos, el uso de una escala muy detallada podría no revelar los patrones espaciales a través de los análisis geoespaciales (Beale, Abellan, Hodgson & Jarup, 2008). En el aspecto concreto de las enfermedades raras, hay algunas de ellas comparadas con la EH (analizada en nuestro estudio) que contienen un mayor número de casos en el mismo periodo, como la esclerosis lateral amiotrófica o la silicosis. Sin embargo, el número de casos también se mostró insuficiente para conseguir un nivel de desagregación superior a la comarca (Alonso, Escobar, Sánchez-Díaz & Posada, 2019). La mortalidad de enfermedades como cáncer, dónde suele haber un mayor número de casos (aunque también hay algunos tipos muy infrecuentes), permite análisis estadísticos alcanzando incluso hasta un nivel de sección censal (Puigpinós *et al.*, 2011). Algunos de los criterios propuestos por Arsenault *et al.* (2013) acerca de la selección de la unidad de agregación han sido considerados en este estudio: Comunicabilidad de los resultados, homogeneidad intra-unidad, variación en el tamaño de población y variación en el tamaño del área. La aplicación del segundo criterio también revela que las comarcas son la unidad geográfica de análisis más adecuada, mostrando una mejor unidad interna que las provincias o municipios.

Las comarcas podrían ser las unidades geográficas más adecuadas para implementar objetivos de salud pública debido a que sus áreas son más consistentes y homogéneas en el territorio nacional. Las autoridades sanitarias en España, las consejerías de cada comunidad autónoma, podrían implementar políticas en sus comarcas correspondientes ya que tienen un mejor conocimiento de las cuestiones de salud en su propio territorio a un nivel intermedio. Sin embargo, juega en su contra la escasa implantación real, así como en el colectivo imaginario de la sociedad de la mayoría de regiones españolas de las divisiones comarcales excepto en algunas Comunidades Autónomas, (Solis-Trapero & Mohíno-Sanz, 2020; Membrado-Tena, 2013; Rodríguez, Menéndez & Cadenas, 2005). Finalmente, las comarcas también se adaptan mejor al criterio sobre población y variación de áreas. Las grandes variaciones podrían llevar a un sesgo en la identificación de patrones espaciales ya que los estudios epidemiológicos suelen basarse en supuestos distribucionales (Berke, 2004). Si bien en ciertas comunidades autónomas se ha delegado en las comarcas algunas competencias, en la mayoría de las comunidades, la comarca carece de entidad jurídica, limitándose a una agrupación de municipios en base a criterios fisiográficos, que puede también implicar un sentimiento de pertenencia entre municipios reforzado por los lazos geográficos y culturales comunes. Dotar de entidad legal a la comarca, desde todas las comunidades autónomas, sería un buen punto de partida para la recopilación de datos más estables.

Obviamente, estas unidades geográficas españolas no se corresponden exactamente con las de otros países, si bien este estudio podría ser extrapolable a sus respectivos niveles de agregación desde mayor a menor detalle. Por ejemplo, en Estados Unidos se podrían comparar municipios, condados (county) y estados (Pavković & Radan, 2007), o en el caso de Portugal freguesías, municipios y distritos (Nunes-Silva, 2017). En los países que no cuentan con una unidad de análisis equivalente a las comarcas españolas, una agregación ad hoc de unidades municipales podría ser aconsejable teniendo en cuenta las características señaladas por Arsenault *et al.* (2013).

5. Conclusiones

La utilidad de este estudio reside en la posibilidad de replicar este procedimiento con otras enfermedades o disciplinas donde la base del análisis es una variable de componente geográfica con un número bajo de casos. Los resultados presentados pueden servir como soporte en la elección de la forma más óptima de cartografiar y analizar un evento, un primer paso fundamental para llevar a cabo análisis bivariables o multivariables posteriormente. La atención debe ponerse, no solo en las características de los datos disponibles (cantidad y distribución geográfica), sino también en la heterogeneidad de las entidades espaciales que componen las diferentes unidades de agregación disponibles. En nuestro caso, las comarcas, con un tamaño medio de 1.542 km^2 y una población promedio de 136.000 habitantes por unidad se ha mostrado como la mejor opción para analizar y representar los resultados de indicadores de mortalidad con una ER específica, teniendo en cuenta el equilibrio entre su tamaño, número de habitantes por unidad y número de eventos registrados en cada una de ellas. Pese a ser unas unidades de delimitación con un respaldo variable desde el punto de vista legal e institucional por parte de las Comunidades Autónomas, en el caso de la representación espacial de indicadores de salud muestran ser adecuadas ya que se consigue desagregar las delimitaciones provinciales, mostrando resultados más detallados, y ofrecen la posibilidad de establecer políticas sanitarias que en algunos casos podrían corresponderse con las áreas de salud que gestionan los gobiernos regionales.

Financiación

Este artículo se enmarca dentro del proyecto de investigación financiado por la Acción Estratégica en Salud Intramural del Instituto de Salud Carlos III. Referencia del proyecto: TPY1238/15.

Referencias

Alonso, V., Escobar, F., Sánchez-Díaz, G., & Posada, M. (2019). *Atlas de Mortalidad debida a Enfermedades Raras en España (Rare Disease Mortality Atlas of Spain)*. Madrid: Instituto de Salud Carlos III, Instituto de Investigación de Enfermedades Raras; Universidad de Alcalá.

Alonso, V., Villaverde-Hueso, A., Hens, M., Morales-Piga, A., Abaitua, I., & Posada de la Paz, M. (2011). Increase in motor neuron disease mortality in Spain: temporal and geographical analysis (1990-2005). *Amyotrophic Lateral Sclerosis, 12*, 192-198. https://doi.org/10.3109/17482968.2010.543688

Anselin, L., Syabri, I., & Kho, Y. (2006). GeoDa: An introduction to spatial data analysis. *Geographical Analysis, 38*, 5-22. https://doi.org/10.1111/j.0016-7363.2005.00671.x

Arab, A. (2015). Spatial and spatio-temporal models for modeling epidemiological data with excess zeros. *International Journal of Environmental Research and Public Health, 12*(9), 10536-10548. https://dx.doi.org/10.3390%2Fijerph120910536

Arsenault, J., Michel, P., Berke, O., Ravel, A., & Gosselin, P. (2013). How to choose geographical units in ecological studies: proposal and application to campylobacteriosis. *Spatial and Spatiotemporal Epidemiology, 7*, 11-24. https://doi.org/10.1016/j.sste.2013.04.004

Beale, L., Abellan, J.J., Hodgson, S., & Jarup, L. (2008). Methodologic issues and approaches to spatial epidemiology. *Environmental Health Perspectives, 116*(8), 1105-1110. https://dx.doi.org/10.1289%2Fehp.10816

Berke, O. (2004). Exploratory disease mapping: kriging the spatial risk function from regional count data. *International Journal of Health Geographics, 3*, 18. https://doi.org/10.1186/1476-072X-3-18

Besag, J., York, J., & Mollìe, A. (1991). A Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics, 43*, 1-59. https://doi.org/10.1007/BF00116466

Bosque, J., Chuvieco, E., & Santos-Preciado, JM. (1986). Algunos problemas metodológicos de las técnicas cuantitativas en Geografía Humana. En Grupo de Métodos Cuantitativos (A.G.E.) (Ed.), *Geografía Teórica y Cuantitativa: concepto y métodos* (pp. 55-74). Oviedo: Universidad de Oviedo.

Cauvin, C., Escobar, F., & Serradj, A. (2010). *Thematic Cartography and Transformations*. 1st ed. Londres: Wiley.

Cebrecos, A., Domínguez-Berjón, MF., Duque, I., Franco, M., & Escobar, F. (2018). Geographic and statistic stability of deprivation aggregated measures at different spatial units in health research. *Applied Geography, 95*, 9-18. https://doi.org/10.1016/j.apgeog.2018.04.001

Cockings, S., & Martin, D. (2005). Zone design for environmental and health studies using pre-aggregated data. *Social Science & Medicine, 60*, 2729-2742. https://doi.org/10.1016/j.socscimed.2004.11.005

Decision No 1295/1999/EC of the European Parliament and of the Council of 29 April 1999 adopting a programme of Community action on rare diseases within the framework for action in the field of public health (1999–2003). (1999) Commission of the European Communities, Brussels. Retrieved from http://ec.europa.eu/health/archive/ph_overview/previous_programme/rare_diseases/raredis_wpgm99_en.pdf

Dykes, J., & Brunsdon, C. (2007). Geographically weighted visualization: Interactive graphics for scale-varying exploratory analysis. *IEEE Transactions on Visualization and Computer Graphics, 13*(6), 1161-1168. https://doi.org/10.1109/TVCG.2007.70558

Elliot, P., & Wartenberg, D. (2004). Spatial epidemiology: Current approaches and future challenges. *Environmental Health Perspectives, 112*(9), 998-1006. https://dx.doi.org/10.1289%2Fehp.6735

Fernández, J., Curt, MD., Aguado, PL., Esteban, B., Sánchez, J., Checa, M., … Romero, L. (2011). *Caracterización de las comarcas agrarias de España*. Recuperado de http://comarcasagrarias.chil.me/

Gómez-Díaz, J. (2008). División territorial de España. Provincias y partidos judiciales. 175 años. *Toletum, 55*, 151-175. Recuperado de https://realacademiatoledo.es/wp-content/uploads/2014/01/files_toletum_0055_07.pdf

Gómez-Rubio, V., & López-Quílez, A. (2010). Statistical methods for the geographical analysis of rare diseases. En M. Posada & S.C. Groft, (Eds.), *Rare Diseases Epidemiology* (pp. 151-172). https://doi.org/10.1007/978-90-481-9485-8_10

Goodwin, S., Dykes, J., Slingsby, A., & Turkay, C. (2015). Visualizing multiple variables across scale and geography. *IEEE Transactions on Visualization and Computer Graphics, 22*(1), 599-608. https://doi.org/10.1109/TVCG.2015.2467199

Grady, S.C., & Enander, H. (2009). Geographic analysis of low birthweight and infant mortality in Michigan using automated zoning methodology. *International Journal of Health Geographics, 8*, 10. https://dx.doi.org/10.1186%2F1476-072X-8-10

Greenwood, P.E., & Nikulin, M.S. (1996). *A guide to chi-squared testing*. Nueva York: Wiley.

Groft, S.C., & Posada, M. (2017). Rare Diseases: joining mainstream research and treatment based on reliable epidemiological data. En *Rare Diseases Epidemiology: Update and overview* (pp. 3-21). https://doi.org/10.1007/978-3-319-67144-4_1

Hipocrates, Thucydides, Cifton F., & Ephesus, F. (1752). *Hippocrates upon air, water and situation: upon endemical diseases and upon prognostics, in acute cases especially*. Londres: Whiston, J., & Whyte, B.

Instituto Geográfico Nacional (2018). *Serie Compendios del Atlas Nacional de España (ANE)*. Recuperado de http://www.ign.es/web/ign/portal/espana-en-mapas

INE. (2020). Instituto Nacional de Estadística. Recuperado de: http://www.ine.es/welcome.shtml

Johnston, R. (2002). Manipulating maps and winning elections: measuring the impact of malapportionment and gerrymandering. Political Geography, *21*(1), 1-31. https://doi.org/10.1016/S0962-6298(01)00070-1

Kanaroglou, P., Delmelle, E., & Páez, A. (2015). *Spatial analysis in health geography*. https://doi.org/10.4324/9781315610252

Labib, S.M., Lindley, S., & Huck, J.J. (2020). Scale effects in remotely sensed greenspace metrics and how to mitigate them for environmental health exposure assessment. *Computers, Environment and Urban Systems, 82*(2020), 101501. https://doi.org/10.1016/j.compenvurbsys.2020.101501

Lai, P., So, F., & Chan, K. (2009). *Spatial Epidemiological Approaches in Disease Mapping and Analysis*. https://doi.org/10.1201/9781420045536

Lawson, A.B., Banerjee, S., Haining, R.P., & Ugarte, M.D. (2016). *Handbook of Spatial Epidemiology*. https://doi.org/10.1201/b19470

Lawson, A.B. (2006). *Statistical Methods in Spatial Epidemiology*. Chichester: John Wiley & Sons.

Meliker, J., & Sloan, C. (2011). Spatio-temporal epidemiology: Principles and opportunities. *Spatial and Spatiotemporal Epidemiology, 2*(1), 1-9. https://doi.org/10.1016/j.sste.2010.10.001

Membrado-Tena, J.C. (2013). La división territorial valenciana: antecedentes, problemas y política de la Generalitat. *Investigaciones Geográficas, 59*(1), 5-24. https://doi.org/10.14198/INGEO2013.59.01

Menon, C. (2012). The bright side of MAUP: Defining new measures of industrial agglomeration. *Papers in Regional Science, 91*(1), 3-28. https://doi.org/10.1111/j.1435-5957.2011.00350.x

Meyer, S., Held, L., & Höhle, M. (2016). Spatio-Temporal analysis of epidemic phenomena using the R Package surveillance. *Journal of Statistical Software, 77*(11), 1-55. http://dx.doi.org/10.18637/jss.v077.i11

Moore, D., & Carpenter, T. (1999). Spatial analytical methods and Geographic Information Systems: Use in health research and epidemiology. *Epidemiological Reviews, 21*(2), 143-161. https://doi.org/10.1093/oxfordjournals.epirev.a017993

Moran, P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika, 37*(1), 17-23. https://doi.org/10.1093/biomet/37.1-2.17

Moreno, L. (1994). Ethnoterritorial Concurrence and Imperfect Federalism in Spain. En B. de Villiers (Ed.), *Evaluation Federal Systems* (pp. 162-193). Cape Town: Juta & Co/Martinus Nijhoff Publishers.

Nunez-Silva, C. (2017). Political and administrative decentralization in Portugal: Four decades of democratic local government. En C. Nunes-Silva, & J. Buček (Eds.), *Local Government and Urban*

Governance in Europe. The Urban Book Series (pp. 9-32). https://doi.org/10.1007/978-3-319-43979-2_2

Olaya, V. (2016). *Sistemas de Información Geográfica*. CreateSpace Independent Publishing Platform. Recuperado de https://volaya.github.io/libro-sig/

Openshaw, S. (1984). *The modifiable areal unit problem*. Norwich: Geo Abstracts University of East Anglia.

Posada, M., Alonso, V., & Bermejo, E. (2016). *Enfermedades raras*. Madrid: Catarata.

Posada, M., Villaverde-Hueso, A., Alonso, V., János, S., Zurriaga, O., Pollán, M., & Abaitua-Borda, I. (2010). Rare diseases epidemiology research. *Advances in Experimental Medicine and Biology, 686*, 17-39. https://doi.org/10.1007/978-90-481-9485-8_2

Puigpinós-Riera, R., Marí-Dell'Olmo, M., Gotsens, M., Borrel, C., Serral, G., Ascaso, C., … & Sánchez-Villegas, P. (2011). Cancer mortality inequalities in urban areas: a bayesian small area analysis in Spanish cities. *International Journal of Health Geographics, 10*, 6. https://doi.org/10.1186/1476-072X-10-6

Pavković A., & Radan, P. (2007). *Creating New States: Theory and Practice of Secession*. Burlington: Ashgate.

Redondo-Bravo, L., Suárez-Rodríguez, B., Fernández-Martínez, B., Simón-Soria, F., Díaz-García, O., & Sierra-Moros, J. (2018). Epidemia por virus Zika. Respuesta desde la salud púbica en España. *Revista Española de Salud Pública, 92*, e1-e16.

Rodríguez, F., Menéndez, M., & Cadenas, A. (2005). Comarcas, consorcios y otras experiencias innovadoras de cooperación territorial en España. *Boletín de la AGE, 39*, 177-199.

Royo, M.A., & Moreno, J.D. (2009). *Método epidemiológico*. Madrid: ENS - Instituto de Salud Carlos III.

Rue, H., Riebler, A., Sørbye, S., Illian, J., Simpson, D., & Lindgren, F. (2016). Bayesian Computing with INLA: A Review. *Annual Review of Statistics and Its Application, 4*(1), 1-28. https://doi.org/10.1146/annurev-statistics-060116-054045

Rytkönen, M.J. (2004). Not all maps are equal: GIS and spatial analysis in epidemiology. *Int J Circumpolar Health, 63*(1), 9-24.

Sánchez-Díaz, G., Arias-Merino, G., Villaverde-Hueso, A., Morales-Piga, A., Abaitua-Borda, I., Hens, M., … Alonso-Ferreira, V.(2016). Monitoring Huntington's disease mortality across a 30-year period: geographic and temporal patterns. *Neuroepidemiology, 47*(3-4), 155-163. https://doi.org/10.1159/000452860

Shaw, N.T., & McGuire, S.K. (2017). Understanding the use of geographical information systems (GISs) in health informatics research: a review. *BMJ Health & Care Informatics, 24*(2), 228-233. https://doi.org/10.14236/jhi.v24i2.940

Smirnov, O., & Anselin, L. (2001). Fast maximum likelihood estimation of very large spatial autoregressive models: a characteristic polynomial approach. *Computational Statistics and Data Analysis, 35*(3), 301-319. https://doi.org/10.1016/S0167-9473(00)00018-9

Snow, J. (1855). *On the mode of communication of cholera*. Londres: John Churchill.

Solís-Trapero, E., & Mohíno-Sanz, I. (2020). Los convenios de colaboración y acuerdos de cooperación entre Comunidades Autónomas. Pensar y actuar sobre nuevos territorios. *Ciudades, 23*, 95-114. https://doi.org/10.24197/ciudades.23.2020.95-114

Ugarte, M.D., Ibáñez, B., & Militino, A.F. (2006). Modelling risks in disease mapping. *Statistical Methods in Medical Research, 15*, 21-35. https://doi.org/10.1191/0962280206sm424oa

Urteaga, L. (1980). Miseria, miasmas y microbios. Las topografías médicas y el estudio del medio ambiente en el siglo XIX. *Geocrítica, Cuadernos Críticos de la Geografía Humana, 29*.

Waller, L., & Gotway, C. (2004). *Applied Spatial Statistics for Public Health Data*. Hoboken: John Wiley & Sons.

Wang, Y., & Di, Q. (2020). Modifiable areal unit problem and environmental factors of COVID-19 outbreak. *Science of the Total Environment, 740*(2020), 139984. https://doi.org/10.1016/j.scitotenv.2020.139984

Zandbergen, P.A. (2014). Ensuring confidentiality of geocoded health data: Assessing Geographic masking strategies for individual-level data. *Advances in Medicine, 2014*, 567049. http://dx.doi.org/10.1155/2014/567049