

Nota metodológica

## Estimación de indicadores de salud en áreas pequeñas a partir de datos de la Encuesta de Salud de Euskadi

Imanol Montoya<sup>a,\*</sup>, Santiago Esnaola<sup>b</sup>, Montserrat Calvo<sup>b</sup>, Elena Aldasoro<sup>b</sup>, Covadonga Audicana<sup>b</sup> y Marc Marí-Dell'Olmo<sup>c,d,e,f</sup>

<sup>a</sup> Departamento de Trabajo y Justicia, Gobierno Vasco, Vitoria-Gasteiz, España

<sup>b</sup> Departamento de Salud, Gobierno Vasco, Vitoria-Gasteiz, España

<sup>c</sup> Servicio de Calidad e Intervención Ambiental, Agència de Salut Pública de Barcelona, Barcelona, España

<sup>d</sup> CIBER de Epidemiología y Salud Pública (CIBERESP), España

<sup>e</sup> Institut d'Investigació Biomèdica Sant Pau, Barcelona, España

<sup>f</sup> Universitat Pompeu Fabra, Barcelona, España

### INFORMACIÓN DEL ARTÍCULO

*Historia del artículo:*

Recibido el 16 de enero de 2018

Aceptado el 25 de abril de 2018

On-line el xxx

*Palabras clave:*

Análisis de áreas pequeñas

Encuesta de salud

Estadística espacial

*Keywords:*

Small area analysis

Health survey

Spatial statistics

### R E S U M E N

Las encuestas de salud son una herramienta clave para la toma de decisiones en políticas de salud y para la planificación de los servicios de salud. El uso de métodos estadísticos para áreas pequeñas que utilizan información de diferentes ámbitos geográficos puede resultar útil para estimar indicadores de salud a una escala geográfica menor que la originalmente considerada en el diseño de una encuesta.

© 2018 SESPAS. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Small area estimation of health indicators based on data from the Basque Country Health Survey

#### A B S T R A C T

Health surveys are a key tool for decision-making in health policies and health services planning. The use of small-area statistical methods that use information at different geographic levels may be useful for estimating health indicators at a lower geographical level than originally considered in the survey design.

© 2018 SESPAS. Published by Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Introducción

La toma de decisiones en políticas de salud y en la planificación de los servicios de salud precisa información a distintas escalas de agregación geográfica. Los usos y las aplicaciones de los resultados de las encuestas de salud aumentan de forma considerable si los principales indicadores son descritos para áreas pequeñas.

La Encuesta de Salud de Euskadi (ESCAV) tiene un diseño muestral que permite obtener estimaciones válidas de determinados indicadores de salud (error muestral inferior al 5%) para cada una de las 12 organizaciones sanitarias integradas (OSI). La planificación de la salud y la gestión de los servicios de atención primaria precisan información más detallada para cada una de las 122 zonas de salud de Euskadi. El objetivo de este trabajo es hacer una propuesta de modelo espacial jerárquico bayesiano para obtener estimaciones

de los principales indicadores de la ESCAV en un ámbito geográfico menor, como es la zona básica de salud.

Existen otras propuestas metodológicas<sup>1-3</sup>, pero resultan más complejas de implementar que la que se propone en este trabajo.

### Método

La ESCAV es una encuesta transversal sobre la salud y sus determinantes en la población no institucionalizada residente en Euskadi. Se usaron los datos de la ESCAV de 2013. Para obtener la muestra se seleccionaron 5300 viviendas por muestreo aleatorio estratificado de las viviendas ocupadas. El diseño muestral se condicionó a que hubiera al menos 10 viviendas en cada una de las 122 zonas básicas de salud. La [tabla 1](#) muestra el número de personas encuestadas en cada una de las OSI, por sexo. Se obtuvo información de 6800 mujeres y 6195 hombres. En cuanto a las 122 zonas básicas de salud, la mediana del número de personas encuestadas fue de 48 hombres y 51 mujeres, con un mínimo de 9 y un máximo de 223. El tamaño de la muestra disponible por zonas básicas no permite obtener estimaciones fiables mediante formas de cálculo

\* Autor para correspondencia.

Correo electrónico: [imanolmontoya@gmail.com](mailto:imanolmontoya@gmail.com) (I. Montoya).

<https://doi.org/10.1016/j.gaceta.2018.04.016>

0213-9111/© 2018 SESPAS. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

**Tabla 1**  
Población encuestada por organización sanitaria integrada en la Encuesta de Salud de Euskadi 2013

OSI	Hombres	Mujeres
OSI Araba	1260	1248
OSI Ezkerra-Enkarterria	364	413
OSI Barakaldo-Sestao	300	342
OSI Bilbo-Basurtu	739	901
OSI Uribe	549	590
OSI Barrualde-Galdakao	706	770
OSI Debabarrena	210	246
OSI Debagoiena	273	282
OSI Goierri-Urola	316	327
OSI Tolosaldea	207	210
OSI Gipuzkoa	1052	1195
OSI Bidasoa	219	276
Total Euskadi	6195	6800

OSI: organización sanitaria integrada.

**Tabla 2**  
Principales indicadores analizados por áreas pequeñas de la Encuesta de Salud de Euskadi 2013

<ul style="list-style-type: none"> <li>• Autovaloración de la salud: mala salud</li> <li>• Mala salud bucodental</li> <li>• Calidad de vida relacionada con la salud</li> <li>• Índice de salud mental</li> <li>• Mala salud mental (síntomas de ansiedad-depresión)</li> <li>• Limitación crónica de la actividad</li> <li>• Necesidad de ayuda para realizar actividades de la vida cotidiana</li> <li>• Problemas crónicos</li> <li>• Lesiones no intencionadas</li> <li>• Dificultad para llegar a fin de mes</li> <li>• Satisfacción con el trabajo</li> <li>• Apoyo social afectivo bajo</li> <li>• Apoyo social confidencial bajo</li> <li>• Bebedores habituales</li> <li>• Consumo de alcohol de riesgo a largo plazo</li> <li>• Consumo intensivo de alcohol mensual o más frecuente</li> <li>• Obesidad</li> <li>• Sedentarismo</li> <li>• Bronquitis</li> <li>• Desplazamiento activo al lugar de trabajo o estudios</li> <li>• Consumo diario de frutas y verduras</li> <li>• Satisfacción con el sistema sanitario</li> <li>• Mujeres de 50-64 años que se han hecho mamografía en los 3 años previos</li> <li>• Mujeres que se han hecho una citología en los 3 años previos</li> <li>• Personas que no han consultado al dentista en los 12 meses previos</li> </ul>
--

convencionales. Puede consultarse más información acerca de la metodología utilizada en la ESCAV en otro documento ya publicado<sup>4</sup>.

En la **tabla 2** se muestran los indicadores seleccionados para su estimación en cada zona básica de salud. Todos los análisis se realizaron por separado para hombres y mujeres, y para cada uno de los indicadores.

Se utilizó un modelo en dos fases:

- En la primera fase se estimó el indicador para cada una de las 12 OSI por métodos convencionales. Para las variables dicotómicas se estimaron prevalencias estandarizadas por edad, calculadas mediante el método directo y usando como población de referencia la de Euskadi de 2011. Para las variables cuantitativas, las medias se ajustaron por edad usando el análisis de covarianza. En ambos casos se usaron los pesos muestrales.
- En la segunda fase se estimó el indicador correspondiente en cada una de las zonas básicas de salud. Para ello se usaron modelos generalizados mixtos teniendo en cuenta la naturaleza del indicador de estudio: prevalencias o medias. Por ejemplo, para el índice de salud mental se usó un modelo de regresión lineal de efectos

mixtos, mientras que para estimar la prevalencia de ansiedad y depresión se usó un modelo logístico de efectos mixtos.

Los modelos jerárquicos bayesianos propuestos tienen la ventaja de evitar los problemas de inestabilidad que surgen en la estimación para zonas donde la muestra es muy pequeña. En concreto, el modelo usado para la estimación del indicador correspondiente en cada una de las zonas básicas de salud fue un modelo generalizado jerárquico a tres niveles: el individuo, la zona básica de salud y la OSI.

Este modelo se basa en la idea de que se dispone de información relevante en los tres niveles. En cuanto al individuo, tenemos la variable principal de estudio, el sexo, la edad y su localización; de la zona básica conocemos cuáles son sus áreas vecinas y a qué OSI pertenece; y por último, en la OSI tenemos una muestra suficiente para tener un estimador válido de la variable de interés.

A continuación se especifica el modelo en el caso de que el indicador de estudio sea la prevalencia de una variable. En primer lugar, para el individuo el modelo asume que la variable de estudio para una persona  $i$  perteneciente a la zona básica de salud  $Z$  y a la OSI  $o$  sigue una distribución de Bernoulli con probabilidad  $P_{izo}$ . Esto es:

$$\text{Variable}_{izo} \sim \text{Bernoulli} [p_{izo}]$$

y

$$\log it(P_{izo}) = \alpha_{zo} + \beta_1 * \text{edad}_{25-44,i} + \beta_2 * \text{edad}_{45-64,i} + \beta_3 * \text{edad}_{65-74,i} + \beta_4 * \text{edad}_{\geq 75,i}$$

donde  $i = 1, \dots, n$  representa a los individuos incluidos en la muestra;  $z = 1, \dots, 122$  representa cada zona básica de salud; y  $o = 1, \dots, 12$  representa cada OSI. La probabilidad individual está ajustada también por la edad mediante una variable categórica con las siguientes categorías: menores de 25 años, de 25 a 44, de 45 a 64, de 65 a 74 y mayor o igual a 75 años.

Para las zonas básicas de salud, el efecto de cada zona se descompone a su vez en tres componentes, de manera similar al modelo desarrollado por Besag, York y Mollie<sup>5</sup>: un efecto común para todas las zonas correspondientes a una misma OSI  $\alpha_o$ , un efecto espacial  $S_z$  que recoge la correlación espacial entre zonas vecinas, y por último un efecto aleatorio  $H_z$  conocido como efecto heterogéneo. Es decir:

$$\alpha_{zo} = \alpha_o + S_z + H_z$$

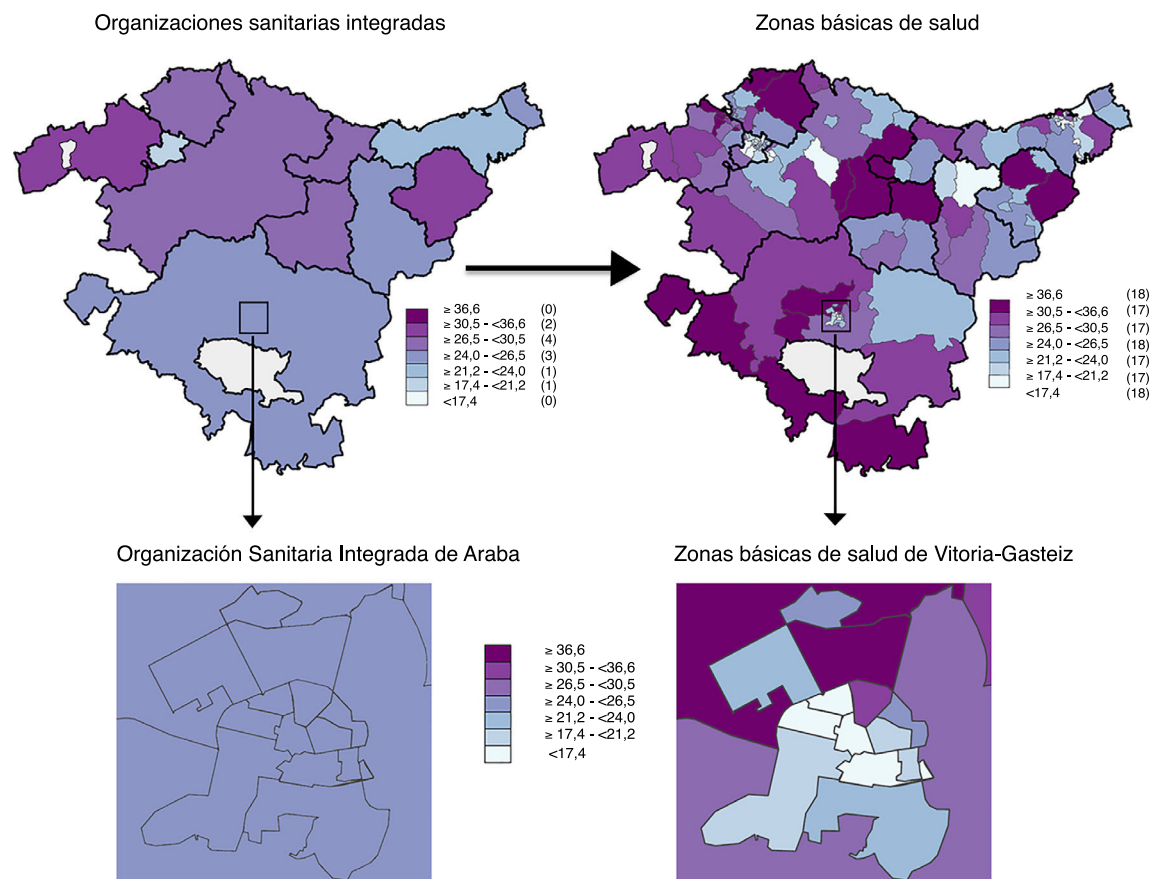
En las OSI se encuentra el intercepto general  $a$  y se introduce la prevalencia estimada mediante el método convencional para cada una de las 12 OSI. Dicha prevalencia lleva asociada una variabilidad (un error estándar) y entra en el modelo de WinBUGS (véase el **Apéndice on-line**) como variable aleatoria:

$$\alpha_o = \alpha + \log it (\widehat{\text{prevalencia estimada OSI}_o})$$

Finalmente, para estimar el parámetro de interés, la prevalencia en la zona básica de salud, se usó la siguiente ecuación:

$$\widehat{\text{Prevalencia}}_{zo} = \frac{\exp(\log it(\widehat{\text{prevalencia estimada OSI}_o}) + S_z + H_z)}{1 + \exp(\log it(\widehat{\text{prevalencia estimada OSI}_o}) + S_z + H_z)}$$

Para estimar los parámetros del modelo se usaron métodos bayesianos. Se asignó una distribución uniforme  $U(0,5)$  a las desviaciones estándar de ambos efectos aleatorios, siguiendo las indicaciones de Gelman<sup>6</sup>. La estimación y el intervalo de credibilidad de la prevalencia en la zona básica de salud se obtuvieron mediante la distribución posterior proporcionada por el modelo. Se utilizaron los mismos criterios que en el atlas de mortalidad en



**Figura 1.** Proporción estimada de hombres sedentarios por organización sanitaria integrada y por zona básica de salud en Euskadi, 2013.

áreas pequeñas de la CAPV<sup>7</sup> para valorar la convergencia de los modelos.

## Resultados

Para una interpretación correcta de los resultados es necesario tener en cuenta los métodos estadísticos utilizados en el cálculo de los indicadores. Los indicadores están ajustados por edad, lo que permite comparar áreas de poblaciones con distinta estructura de edad. Además, el valor calculado para cada zona básica de salud es una combinación de tres elementos: uno que representa al conjunto de la OSI al que pertenece la zona, otro propio de cada zona y el tercero correspondiente a la correlación espacial de la zona con las vecinas. En esta combinación, el peso relativo de cada uno de los elementos depende de la cantidad de información disponible, que a su vez es función del tamaño muestral y de la variabilidad del indicador.

La **figura 1** muestra a modo de ejemplo los resultados obtenidos para la prevalencia de hombres sedentarios, que tiene una gran variabilidad entre las zonas básicas de salud. Así, en Vitoria-Gasteiz hay zonas con una mayor prevalencia que no se habrían identificado analizando únicamente las OSI.

En la **figura 2** se muestran los resultados obtenidos para la prevalencia de bronquitis en las mujeres. En este caso no se ven apenas diferencias en la estimación por zonas básicas de salud de una misma OSI. Esto se debe a la poca variabilidad que existe en el propio indicador. Esta estimación podrá mejorarse incluyendo otras variables explicativas, como el consumo de tabaco. Los resultados para el resto de los indicadores analizados se pueden ver en el atlas de la encuesta de salud de la CAPV 2013 por áreas pequeñas<sup>8</sup>.

## Discusión

El uso en la encuesta de salud de métodos estadísticos de áreas pequeñas que utilizan información a diferentes niveles se ha mostrado útil para obtener estimaciones a una escala geográfica más pequeña que la originalmente considerada.

Una limitación de este trabajo es un posible sesgo debido a que la muestra puede resultar no representativa en algunas de las áreas con menor número de personas encuestadas. Como línea de mejora en el futuro podría considerarse un único modelo para hombres y mujeres añadiendo el sexo como covariable con sus diferentes interacciones. De esta manera se dispondría de más información en las zonas básicas de salud, en caso de que no haya grandes diferencias por sexo. También sería necesario evaluar el modelo propuesto mediante un estudio de simulación que permita determinar su validez para la estimación de indicadores en las zonas básicas de salud en distintos escenarios, y compararla con la de otras propuestas de la literatura con el mismo fin<sup>1-3</sup>.

Sin embargo, se ha podido observar que, cuando hay variabilidad suficiente en el indicador analizado, estos métodos permiten encontrar diferencias entre las zonas de una misma OSI que hubieran quedado ocultas. Esto permitirá mejorar las decisiones en políticas en salud y en la planificación y la gestión de los servicios de salud.

## Editora responsable del artículo

María Victoria Zunzunegui.

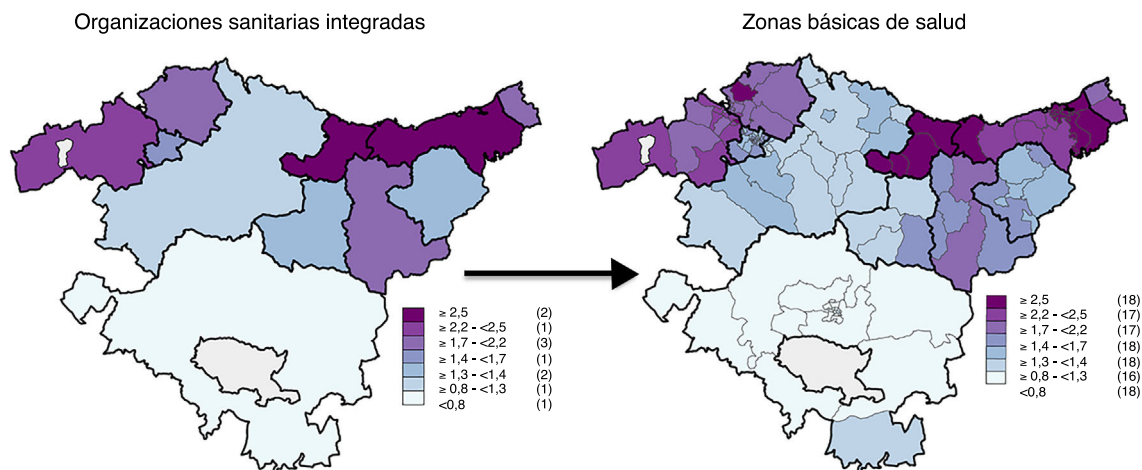


Figura 2. Prevalencia estimada de bronquitis en mujeres por organización sanitaria integrada y por zona básica de salud en Euskadi. 2013.

### Contribuciones de autoría

Concepción y diseño del trabajo: I. Montoya y S. Esnaola. Recogida de datos: M. Calvo y C. Audicana. Análisis de los datos: I. Montoya. Interpretación de los resultados: I. Montoya, S. Esnaola, M. Calvo, E. Aldasoro, C. Audicana y M. Marí-Dell'Olmo. Escritura del artículo y revisión crítica con importantes contribuciones intelectuales: I. Montoya, S. Esnaola, M. Calvo, E. Aldasoro, C. Audicana y M. Marí-Dell'Olmo. Todas las personas firmantes han aprobado la versión final para su publicación.

### Financiación

Ninguna.

### Conflicto de intereses

Ninguno.

### Anexo. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en [doi:10.1016/j.gaceta.2018.04.016](https://doi.org/10.1016/j.gaceta.2018.04.016)

### Bibliografía

- Chen C, Wakefield J, Lumely T. The use of sampling weights in Bayesian hierarchical models for small area estimation. *Spatial and Spatio-Temporal Epidemiology*. 2014;11:33-43.
- Mercer L, Wakefield J, Chen C, et al. A comparison of spatial smoothing methods for small area estimation with sampling weights. *Spatial Statistics*. 2016;8:69-85.
- Vandendijck Y, Faes C, Kirby RS, et al. Model-based inference for small area estimation with sampling weights. *Spatial Statistics*. 2016;18:455-73.
- Metodología de la encuesta de salud ESCAV 2013. Departamento de Salud del Gobierno Vasco. (Consultado el 8/2/2018.) Disponible en: [http://www.euskadi.eus/contenidos/informacion/enc.salud.2013.metodologia/es.def/adjuntos/Metodologia\\_ESCAV\\_2013.pdf](http://www.euskadi.eus/contenidos/informacion/enc.salud.2013.metodologia/es.def/adjuntos/Metodologia_ESCAV_2013.pdf)
- Besag J, York J, Mollie A. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*. 1991;43:1-59.
- Gelman A. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*. 2016;1:1-19.
- Esnaola S, Montoya I, Calvo M, et al. Atlas de mortalidad en áreas pequeñas de la CAPV (1996-2003). Vitoria-Gasteiz: Departamento de Sanidad y Consumo; 2010. p. 165.
- Esnaola S, Montoya I, de Diego M, et al. Atlas de la Encuesta de Salud de la CAPV 2013 por áreas pequeñas. Vitoria-Gasteiz: Departamento de Salud, Servicio de Estudios e Investigación Sanitaria; 2015 (Consultado el 16/1/2018.) Disponible en: <http://www.euskadi.eus/contenidos/informacion/enc.salud.2013.publicaciones/es.def/adjuntos/AtlasEremuTxiki.pdf>