

Impacto de la COVID-19 en los centros residenciales de Euskadi: un análisis multinivel de la importancia de los factores de riesgo relativos al centro

Madalen Saizarbitoria

SIIS Servicio de Información e Investigación
msaizarbitoria@siis.net

Oraintsuko ikerlan baten barruan egindako maila anitzeko analisiaren berri ematen du artikulua honek. Ikerlanean zera ondorioztatzen da: 2020ko martxotik urrira bitartean Euskadiko bizitegi-zentroetan zeuden adineko pertsonen SARS-COV-2 koronabirusa elkarri kutsatzeko arriskuari dagokionez, zentro batetik bestera aurkitutako aldean % 55 zentro bakoitzaren ezaugarriak esplikatzen dutela. Bizitegi-zentro horietan artatzen zituzten eta Osakidetza bere esku zituen 20.000 lagun baino gehiagoren datuak aztertu ahal izan dituzte analisiaren egileek. Datuen arabera, kutsatutako pertsonen hiltzeko arriskua % 20 eta % 57 artean igo omen zen aztergai hartutako aldi horretan. Zentro bakoitzean prebentzio modura esku hartzea oso eraginkorra den arren, emaitzek argi erakusten dute ikertzen jarraitu beharra dagoela: batetik, zentro bakoitzean kutsatzeen kopurua murrizte aldera zer alderdi espezifiko aldatu behar diren jakiteko; eta, bestetik, zentro bakoitzaren ezaugarri estrukturalak eta langileen antolaketak zer-nolako inpaktua duten aztertzeke.

GAKO-HITZAK:

Adineko pertsonak, menpekotasuna, bizitegi-zentroak, COVID-19a, erikortasuna, arrisku-faktoreak.

El presente artículo describe el análisis multinivel realizado en el marco de un reciente estudio que concluye que el centro residencial y sus características pueden explicar hasta un 55 % de las diferencias en el riesgo de contagio por SARS-CoV-2 entre personas que vivían en centros residenciales para personas mayores entre marzo y octubre de 2020 en Euskadi. El análisis, para el que se pudieron analizar datos de Osakidetza de más de 20.000 personas atendidas en esos centros, revela que las personas contagiadas tuvieron entre un 22 % y un 57 % más de riesgo de fallecer en el periodo de estudio. Los resultados indican que la intervención preventiva a nivel de centro puede resultar efectiva, pero que se necesita seguir investigando para conocer los factores específicos del centro que deben modificarse para reducir los contagios. El impacto de las características estructurales de los centros y la organización del personal son aspectos que habría que analizar de forma prioritaria.

PALABRAS CLAVE:

Personas mayores, dependencia, residencias, COVID-19, morbilidad, mortalidad, factores de riesgo.

1. Introducción

Desde el inicio de la pandemia de la COVID-19, la situación en los centros residenciales para personas mayores ha sido objeto de especial atención en muchos países. Esto se ha debido, por una parte, al temor —no infundado— de que podría tratarse de estructuras altamente vulnerables frente al virus, tanto por las características de la población atendida como por la previsión de que, al tratarse de estructuras creadas para fomentar la convivencia y la socialización entre las personas usuarias, se pudiesen plantear dificultades específicas a la hora de aplicar las medidas de control necesarias para frenar la transmisión una vez detectado algún caso entre la población usuaria o trabajadora, dificultades que se añadían, por otra parte, a la falta generalizada de pruebas diagnósticas sobre las que basar las decisiones respecto a las medidas que adoptar en cada caso. Este temor se fue acrecentando a medida que aparecían noticias en los medios de comunicación advirtiendo de la escasez de equipos de protección individual para las personas profesionales de centros sociosanitarios, y salían a la luz informaciones que apuntaban a un número elevado de fallecimientos en algunos centros españoles (Médicos Sin Fronteras, 2020).

Los primeros informes internacionales que abordaron la cuestión de la mortalidad por COVID-19 en los centros residenciales para personas mayores apuntaban, desde luego, a un impacto muy importante en países como el Reino Unido, Francia o España, donde se estimaba que entre un 40 % y un 50 % de la mortalidad total por COVID-19 se había producido entre la población usuaria de residencias para personas mayores (Comas-Herrera *et al.*, 2020). Informes posteriores rebajaron algo estas estimaciones, reconociendo que el mayor impacto en los centros residenciales se había producido durante las primeras semanas de la pandemia y que, posteriormente, se había podido tomar cierto control sobre la situación. La última actualización del informe elaborado por las personas expertas pertenecientes a la International Long-Term-Care Policy Network indicaba que, a 1 de febrero de 2021, en torno al 40 % de los fallecimientos por COVID-19 en España se correspondían con usuarios/as de residencias; no obstante, hay que tener en cuenta que esa estimación se basaba todavía en datos no oficiales, recopilados por RTVE a partir de la información que habían ido publicando los Gobiernos autonómicos (Comas-Herrera *et al.*, 2021).

En efecto, a pesar del interés suscitado por la situación vivida en los centros residenciales para personas mayores durante la pandemia, muchos países no han podido disponer de información oficial, completa y actualizada sobre el número de contagios y fallecimientos por COVID-19 en centros residenciales hasta muy recientemente. En España, por ejemplo, el Gobierno no hizo públicos los primeros datos oficiales sobre la situación en los centros residenciales hasta finales de febrero

de 2021, cuando el Imsero publicó el primer reporte semanal de los datos referidos a centros residenciales para personas mayores, centros para personas con discapacidad y otros alojamientos permanentes de servicios sociales dirigidos a esos colectivos. A partir de esa fecha, los datos se han ido publicando periódicamente¹. El último informe consultado, actualizado a 19 de diciembre de 2021, habla de 101.151 casos confirmados por prueba diagnóstica entre la población residente en centros residenciales para personas mayores de España, 6.999 de ellos en Euskadi. Por otra parte, habrían fallecido con COVID-19 confirmado por prueba diagnóstica 20.300 residentes en España y 1.166 en Euskadi, lo que arroja una tasa de letalidad media para los años 2020 y 2021 del 20,1 % y el 16,7 %, respectivamente.

En Euskadi, el Ararteko ha publicado recientemente un estudio —cuyo informe base fue realizado por el SIIS— en el que se realiza un completo análisis de la morbilidad y la mortalidad por COVID-19 en los centros residenciales para personas mayores durante el año 2020. El estudio tiene varios aspectos que cabe destacar:

- En primer lugar, incorpora información sobre la existencia de contagios y fallecimientos por COVID-19 para la totalidad de los centros residenciales de atención secundaria de Euskadi, a través de los datos facilitados semanalmente por los propios centros para informar a la Fiscalía Superior del País Vasco.
- En segundo lugar, incorpora información individual anonimizada, procedente del Sistema de Información de Osakidetza —Osakidetza Business Intelligence-Oracle Analytic Server (OBI-OAS)—, sobre las más de 20.000 personas que se alojaban en las residencias de Euskadi entre el 1 de marzo y el 31 de octubre de 2020.
- Por último, incorpora un análisis más cualitativo, a partir de una encuesta a la que respondieron aproximadamente la mitad de los centros residenciales de Euskadi y un grupo de discusión con personas expertas, en la que se recoge la valoración de las y los directores de centro respecto a la utilidad de las medidas preventivas aplicadas, las principales dificultades a las que se han enfrentado durante los meses de pandemia y las lecciones que se pueden extraer de cara a abordar situaciones similares en el futuro.

Desde el punto de vista metodológico, el estudio realizado por el Ararteko resulta asimismo innovador, al aplicar a los datos individuales aportados por Osakidetza una técnica de análisis —la regresión logística multinivel— que presenta algunas ventajas respecto a los modelos de regresión tradicionales que se han aplicado en la mayoría de las investigaciones realizadas hasta ahora sobre los

¹ Véanse en <https://www.imsero.es/imsero_01/mas_informacion/serv_soc/sem_cr/index.htm>.

factores asociados con la morbilidad y la mortalidad por COVID-19 en los centros residenciales:

- Por una parte, los modelos multinivel incorporan al análisis la necesaria diferenciación entre el plano individual, al que pertenece en sentido estricto el riesgo que tiene una persona de ser infectada por el virus o de fallecer a causa de él, y el plano grupal (el relativo al centro residencial) sobre cuya influencia nos estamos preguntando, lo que reduce la probabilidad de incurrir en diversos tipos de error que pueden surgir al analizar datos agrupados o jerarquizados mediante las técnicas habituales de regresión².
- Por otra parte, permiten descomponer la varianza total en la respuesta de interés (en nuestro caso, los contagios por COVID-19 o la mortalidad) en diferentes componentes, de acuerdo con la estructura de los datos en niveles o grupos. De esta manera, es posible conocer qué parte de la variabilidad que se observa en el fenómeno de interés se corresponde con factores individuales, que varían de una persona a otra dentro de cada grupo, y qué parte se corresponde con diferencias entre los grupos. Cuanto mayor sea la proporción de la varianza correspondiente a las diferencias entre grupos, mayor será la relevancia de estos grupos o contextos para explicar el fenómeno de interés y, por lo tanto, mayor la efectividad de intervenir preventivamente sobre ellos.

Dado el interés que presenta esta aproximación en particular, en el presente artículo reproducimos el análisis multinivel realizado en el marco del estudio del Ararteko y contrastamos los resultados obtenidos a la luz de la evidencia disponible en la literatura especializada, instando al lector interesado en los aspectos más descriptivos o valorativos a consultar el informe completo (SIIS, 2021). El resto del artículo se estructura en cuatro apartados: en el primero de ellos, se realiza un rápido repaso de las cifras globales que ha dejado la pandemia en

² En los análisis de regresión tradicionales, el problema de los datos jerarquizados (personas anidadas en centros o estudiantes en aulas) se suele obviar recurriendo a la agregación de las variables individuales a nivel de grupo (por ejemplo, utilizando como variable independiente la edad media de las personas usuarias del centro, en lugar de la variable individual 'edad'), o bien asignando las características del grupo a cada uno de los individuos que pertenece a él (por ejemplo, atribuyendo la titularidad del centro a todas las personas que lo utilizan). En el primer caso, se desaprovecha toda la riqueza de los datos originales, se pierde poder estadístico y aumenta la probabilidad de error de tipo II. En el segundo caso, no se cumple el supuesto de independencia de las observaciones para las variables referidas al centro, con lo que se incrementa la probabilidad de cometer un error de tipo I cuando se analizan las diferencias entre centros, o bien se corre el riesgo de ser excesivamente conservador al analizar las diferencias entre individuos. Los modelos multinivel o jerarquizados resuelven este problema incluyendo en el modelo efectos aleatorios, es decir, añadiendo a la estimación de los coeficientes de regresión (que reflejan la asociación media entre las variables dependiente e independiente para todos los individuos) una parte aleatoria que refleja las diferencias que existen entre los grupos (en este caso, los centros) en la asociación medida por la ecuación de regresión (la varianza a nivel de grupo en los coeficientes de regresión).

los centros, a partir de las diferentes fuentes de las que se ha podido disponer para el estudio; en el segundo apartado, de contenido fundamentalmente metodológico, se describen de forma detallada los datos individuales obtenidos de Osakidetza y la estrategia utilizada en el análisis de regresión multinivel; el tercer apartado se centra en los resultados de los modelos; finalmente, en el cuarto y último apartado se revisan los resultados obtenidos a la luz de la evidencia científica disponible en la literatura especializada sobre factores asociados con contagios y mortalidad por COVID-19 en los centros residenciales.

2. Cifras globales del primer año de pandemia en los centros residenciales para personas mayores de Euskadi

La información aportada por los propios centros residenciales a la Fiscalía Superior del País Vasco, remitida por el Departamento de Salud al Ararteko, permite dimensionar a grandes rasgos el impacto que ha tenido la pandemia en la población usuaria de estos centros durante el año 2020. De acuerdo con esta información, 199 de los 298 centros residenciales de atención secundaria existentes en Euskadi³ (dos tercios del total) tuvieron algún contagio entre su población usuaria y 141 (el 47,5%) tuvieron algún fallecimiento atribuible a la enfermedad causada por el virus hasta el 10 de enero de 2021⁴.

La mayoría de los centros afectados por contagios a lo largo del año 2020 en Álava y Bizkaia tuvieron ya algún caso durante la primera ola de la pandemia, entre marzo y junio: concretamente, seis de cada diez en Álava y siete de cada diez en Bizkaia (Tabla 1). En Gipuzkoa, la proporción de centros que se vieron afectados en los primeros meses fue notablemente menor (del 43%).

³ Los datos remitidos a la Fiscalía, recopilados por las diputaciones forales a partir de información aportada por los propios centros, no incluyen datos relativos a centros residenciales de atención primaria, dirigidos a la población autónoma o con dependencia moderada (que son competencia de los ayuntamientos o agrupaciones locales). En cambio, sí incluyen información relativa a todas las personas alojadas en los centros de atención secundaria referidos, independientemente de si ocupan plazas pertenecientes al Sistema Vasco de Servicios Sociales (plazas públicas, concertadas o financiadas públicamente) o plazas privadas (sin financiación pública).

⁴ La información se refiere a fallecimientos de personas usuarias en cuyo certificado de defunción se menciona alguna causa relacionada con la COVID-19.

Tabla 1. Número de centros con casos confirmados de COVID-19 entre la población usuaria, por periodo y territorio histórico

	Año 2020*		1ª ola (marzo -junio 2020)		
	N	% del total	N	% del total	% de los centros afectados en 2020
Álava	45	55,6	27	33,3	60,0
Bizkaia	112	72,7	78	50,6	69,6
Gipuzkoa	42	66,7	18	28,6	42,9
Euskadi	199	66,8	123	41,3	61,8

* Datos hasta el 10 de enero de 2021.

Fuente: Informes semanales de indicadores de la COVID-19 remitidos a la Fiscalía Superior del País Vasco.

Respecto a los fallecimientos, aunque en términos absolutos el número de centros afectados fue mayor en la primavera que a partir del verano (86 centros registraron fallecimientos por COVID-19 en los tres meses de primavera, frente a 90 en los seis meses posteriores), la proporción de centros con algún contagio que registraron fallecimientos resultó similar, de en torno al 70 %, en ambos periodos. De la información recogida en la Tabla 2, cabe asimismo destacar que la proporción de centros afectados en los que se produjeron defunciones atribuibles al virus fue algo menor en Bizkaia que en los otros dos territorios históricos, particularmente durante la primera ola: mientras que algo más de ocho de cada diez centros afectados por casos de COVID-19 entre marzo y junio registraron alguna defunción en Álava y Gipuzkoa, solamente seis de cada diez tuvieron fallecimientos

atribuidos al virus en Bizkaia. Para finales de año, los datos de Álava y Bizkaia se habían igualado (hubo defunciones en prácticamente siete de cada diez centros con casos confirmados), y los de Gipuzkoa seguían estando por encima (con fallecimientos atribuibles al virus en más de ocho de cada diez centros que habían tenido algún caso durante el año).

En lo referente a la proporción de la población usuaria de los centros residenciales para personas mayores que se ha visto afectada por la pandemia, a partir de los datos que aportan los centros a la Fiscalía no es posible conocer la tasa acumulada de contagios (la información no está armonizada para los tres territorios y no se dispone de la cifra de personas únicas que han tenido un test positivo en Álava y Bizkaia), pero sí es posible comparar la mortalidad total y la atribuida a la COVID-19 durante el primer año de pandemia.

En total, entre el 1 de marzo de 2020 y el 10 de enero de 2021, 938 personas mayores alojadas en centros residenciales para personas mayores fallecieron por causas atribuibles a la COVID-19, y 3.393 más murieron por otras causas, lo que da una tasa de mortalidad total del 25,1 % respecto a la población usuaria estimada a 1 de marzo de 2020, y una tasa de mortalidad atribuible a la COVID-19 del 5,4 % (Tabla 3). La tasa de mortalidad por cualquier causa es llamativamente menor en los centros de Gipuzkoa que en los alaveses y vizcaínos; en cambio, la mortalidad atribuible a la COVID-19 es similar en Gipuzkoa y en Álava (en torno al 6 %), mientras que la tasa en los centros de Bizkaia resulta más baja, algo más de un punto porcentual menor.

Tabla 2. Número de centros con defunciones de usuarios/as atribuibles a la COVID-19, por periodo y territorio histórico

	Año 2020*			1ª ola (marzo-junio 2020)		
	N	% del total	% de centros con casos	N	% del total	% de centros con casos
Álava	32	40,0	71,1	22	27,5	81,5
Bizkaia	75	48,7	67,0	49	31,8	62,8
Gipuzkoa	34	54,0	81,0	15	24,2	83,3
Euskadi	141	47,5	70,9	86	29,1	69,9

* Datos hasta el 10 de enero de 2021.

Fuente: Informes semanales de indicadores de la COVID-19 remitidos a la Fiscalía Superior del País Vasco.

Tabla 3. Número de fallecimientos por cualquier causa y atribuibles a la COVID-19 en centros residenciales para personas mayores*. 1-3-2020 al 10-1-2021

	Nº estimado de usuarios/as en marzo**	Fallecimientos por COVID-19	Fallecimientos por otras causas	Tasa mortalidad total (%)	Tasa mortalidad atribuible a la COVID-19 (%)
Álava	2.722	172	607	28,7	6,33
Bizkaia	9.360	455	2.025	26,5	4,86
Gipuzkoa	5.202	311	761	20,7	6,01
Euskadi	17.284	938	3.393	25,1	5,44

* La información disponible se limita a los centros residenciales de atención secundaria (no incluye datos de centros de competencia municipal).

** Se calcula a partir de la cifra de personas usuarias de los servicios recogida en el primer informe, correspondiente a la semana del 27 de abril, sumando a esa cifra el número de personas fallecidas hasta esa fecha.

Fuente: SIIS (2021).

En resumen, los datos aportados a la Fiscalía indican que prácticamente dos de cada tres centros residenciales para personas mayores de atención secundaria de Euskadi tuvieron casos de COVID-19 entre su población usuaria a lo largo de 2020 y en torno al 70 % de esos centros registraron algún fallecimiento relacionado con el virus. En total, 938 personas usuarias (de un total de 17.284 estimadas al inicio de la pandemia) fallecieron como consecuencia de la enfermedad del coronavirus y prácticamente 3.400 más por otras causas, con lo que la tasa de mortalidad total fue del 25 % en diez meses.

3. Factores asociados con contagios y mortalidad en los centros: un análisis multinivel

Para el análisis de los factores que podrían estar asociados a los contagios y a la mortalidad por COVID-19 en los centros residenciales de Euskadi, se planteó un estudio de cohorte transversal y retrospectivo a partir de datos de Osakidetza extraídos del Sistema de Información Osakidetza Business Intelligence (OBI), donde quedan registradas las peticiones de pruebas de laboratorio por parte de los centros residenciales. Este sistema recoge información del 98 % de las residencias para personas mayores o con discapacidad que tienen más de 15 plazas, así como de las viviendas comunitarias de menos de 15 plazas que disponen de personal sanitario, por lo que los datos obtenidos cubren la gran mayoría de las plazas residenciales dirigidas a la población mayor dependiente.

Dado que nuestro interés se centraba en conocer la influencia que ejercen factores del contexto (como las características del centro residencial o la zona de salud en la que se ubica la residencia) en el riesgo de contagio o mortalidad por COVID-19 para las personas residentes, se optó por utilizar la técnica de análisis de regresión multinivel, en la línea de lo que proponen diferentes autores del ámbito de la educación o la sociología para analizar datos agrupados (alumnos en escuelas, personas en barrios o ciudades) y que se ha planteado asimismo como herramienta útil en epidemiología social para medir el efecto de factores socioculturales y del entorno en distintas patologías. Tal y como explica Ana Diez Roux (2008: 87):

Los estudios multinivel examinan simultáneamente grupos (o muestras de grupos) e individuos dentro de ellos (o muestras de individuos dentro de ellos). La variabilidad a niveles grupal e individual se examina simultáneamente y así puede obtenerse información sobre el efecto de las variables de nivel grupal y de nivel individual en explicar la variación entre los individuos y entre los grupos.

Por lo tanto, el análisis multinivel tiene como objetivo modelizar estadísticamente la influencia

de variables contextuales sobre las actitudes o los comportamientos medidos a nivel individual (Andreu y Abela, 2011), y, en este sentido, constituye una herramienta idónea para responder a las preguntas que se plantean en este estudio.

A continuación, se describen los sujetos y variables analizadas y los métodos de análisis aplicados en esta parte del estudio.

3.1. Sujetos y variables de estudio

Mediante una petición realizada por el Ararteko al Departamento de Salud, y previa evaluación favorable del proyecto de investigación por parte del Comité de Ética en Investigación con Medicamentos de Euskadi⁵, se extrajo la siguiente información de OBI para la cohorte formada por las personas que habían estado alojadas en un centro residencial entre el 1 de marzo y el 31 de octubre de 2020:

- Paciente: número ficticio que identifica a cada persona.
- Sexo.
- Fecha de nacimiento.
- Fecha de fallecimiento (en caso de haber fallecido en el periodo de estudio).
- Número de residencias: dado que cada persona puede haber estado en más de una residencia, se especifica en cuántas durante el periodo de estudio.
- Primera residencia: nombre de la primera residencia en la que ha estado.
- Segunda residencia: nombre de la última residencia en la que ha estado (si ha estado en más de una).
- Test (sí/no): indica si a la persona se le ha realizado algún test diagnóstico para COVID-19 (PCR, antígeno o anticuerpos).
- Fecha test positivo: en el caso que la persona haya dado positivo en cualquiera de los test, se especifica la fecha correspondiente.

Por otra parte, a partir de los registros forales de servicios sociales y del Directorio de Servicios Sociales del Gobierno Vasco⁶, se obtuvo información sobre el número de plazas, la titularidad (pública, privada sin fin de lucro o privada mercantil) y la ubicación (zona de salud, organización sanitaria integrada —OSI— y territorio histórico) de cada centro residencial incluido en el censo inicial aportado por las diputaciones forales al Ararteko.

Por último, se obtuvo información del Departamento de Salud sobre la tasa acumulada de contagios por COVID-19 y la mortalidad asociada en la población general de las zonas de salud de Euskadi. Esta base de

⁵ El Comité es un órgano independiente adscrito a la Dirección de Investigación e Innovación Sanitarias del Departamento de Salud del Gobierno Vasco, y se encarga de velar por la protección de los derechos, la seguridad y el bienestar de las personas que participan en proyectos de investigación biomédica.

⁶ <https://www.direktorioa.net>.

datos aporta información sobre el número acumulado de contagios y de fallecimientos en las 135 zonas de salud de la comunidad autónoma a 28 de junio, a 1 de noviembre de 2020 y a 10 de enero de 2021.

3.2. Métodos

Dado que el interés del estudio se centraba fundamentalmente en estimar el impacto que ejercen las características del centro residencial en el riesgo de contagio o mortalidad por COVID-19 para las personas usuarias, el análisis debía incorporar al menos dos niveles: personas residentes (nivel 1) y centros residenciales (nivel 2). Por otra parte, se valoró que la zona de salud (nivel 3) y la OSI (nivel 4) a la que pertenecen los centros podrían asimismo resultar relevantes, al delimitar zonas geográficas dentro de las cuales los centros residenciales comparten recursos del sistema de salud (capacidad para realizar pruebas diagnósticas, disponibilidad de camas hospitalarias), los cuales pueden variar de una zona a otra.

En consecuencia, inicialmente se definió una estructura en cuatro niveles, con las personas residentes en el nivel inferior (nivel 1) y las OSI en el nivel superior (nivel 4)⁷. La primera fase del análisis consistió en valorar la necesidad de aplicar modelos multinivel diferenciando los cuatro niveles, mediante un análisis de componentes de la varianza, consistente en estimar modelos vacíos —únicamente con efectos aleatorios para los niveles 2 a 4— para cada una de las dos variables respuesta de interés: (1) si la persona tuvo un resultado positivo en una prueba de diagnóstico para SARS-CoV-2 en el periodo analizado; y (2) si falleció por cualquier causa. En los modelos finales, únicamente se incluyeron aquellos niveles que capturaban una parte significativa de la varianza total en el riesgo, y que, por lo tanto, al ser tenidos en cuenta, podían mejorar sustancialmente la capacidad predictiva del modelo.

La definición de los modelos se hizo siguiendo una estrategia de análisis en tres pasos descrita por Merlo *et al.* (2016). Explicamos, en primer lugar, los detalles de los modelos relativos a las pruebas positivas para COVID-19 y, después, las peculiaridades de los modelos de mortalidad. Tanto para la definición de los modelos como para su estimación y para la interpretación de los resultados, se contó con el asesoramiento y la colaboración del Dr. Juan Merlo, catedrático de Salud Pública y jefe de la Unidad de Epidemiología Social de la Facultad de Medicina de la Universidad de Lund (Suecia), y de Raquel Pérez Vicente, bioestadística de la misma Unidad. Agradecemos a ambos su inestimable colaboración.

⁷ A la hora de aplicar los modelos se utilizaron cuatro niveles (persona, centro, zona de salud y OSI), considerando el quinto, el correspondiente a los tres territorios históricos, como efecto fijo, puesto que el análisis de efectos aleatorios requiere contar con un número mínimo de grupos, y la agrupación en tres territorios resulta insuficiente.

3.2.1. Modelos para la variable test COVID-19 positivo

Sea que y_i denota la respuesta binaria de interés (en este caso, resultado positivo en una prueba de COVID-19) para el individuo i ($i = 1, \dots, n_i$) en el centro residencial j ($j = 1, \dots, J$).

Paso 1: el modelo de efectos individuales

El paso 1 consiste simplemente en ajustar una regresión logística convencional de un solo nivel para y_i incluyendo solo las covariables de nivel individual edad y sexo:

$$y_i \sim \text{Binomial}(1, \pi_i) \quad (1)$$

$$\text{logit}(\pi_i) = \beta_0 + \beta_1 \text{edad}_i + \beta_2 \text{sexo}_i \quad (2)$$

donde π_i denota la probabilidad de que el individuo i tenga un resultado positivo en una prueba de SARS-CoV-2, dadas sus características individuales de edad, y sexo. En el análisis, estimamos primero un modelo que incluye solo la edad (modelo 1) y otro que incluye la edad y el sexo (modelo 2).

Los coeficientes de regresión β_1 y β_2 miden la asociación entre estas variables independientes o factores y el logaritmo de la razón de probabilidades de tener un resultado positivo en una prueba de SARS-CoV-2 y, cuando se exponencian, se traducen en *odds ratios* (OR). Para facilitar la ilustración, hemos introducido la edad en el modelo de forma lineal, pero en el análisis esta variable se incluye como categórica. Tras la estimación, se calculan las probabilidades previstas $\hat{\pi}_{ij}$ para cada individuo y se utilizan para calcular el área bajo la curva COR (AUC) del modelo.

El AUC se construye trazando la fracción de verdaderos positivos (FVP) (es decir, la sensibilidad) contra la fracción de falsos positivos (FFP) (es decir, $1 - \text{especificidad}$) para diferentes umbrales de clasificación binaria de las probabilidades obtenidas a partir de la predicción. El AUC mide la capacidad del modelo para clasificar correctamente a los individuos (en nuestro caso, para predecir si han tenido o no una prueba positiva para SARS-CoV-2) en función de las probabilidades estimadas a partir del AUC⁸. El AUC toma un valor de entre 1 y 0,5, donde 1 es la discriminación perfecta y 0,5 tendría un carácter tan informativo como lanzar una moneda al aire (es decir, indicaría que las covariables no tienen poder predictivo). El AUC del modelo del paso 1 cuantifica la precisión de utilizar únicamente información de nivel individual (referida a la persona) para identificar a los individuos con el resultado (en nuestro caso, personas con un test positivo) (Pepe *et al.*, 2004).

⁸ Por ejemplo, si se establece el umbral de clasificación en 0,5, las personas cuya probabilidad estimada a partir del modelo fuera superior a ese valor serían clasificadas como positivas, y aquellas cuya probabilidad quedase por debajo, como negativas. Comparando esta predicción con el “estatus” real positivo o negativo de cada persona, se obtendrían los valores de sensibilidad y especificidad para calcular el AUC.

Paso 2: el modelo de efectos contextuales generales

El paso 2 consiste en ampliar el modelo del paso 1 de un modelo convencional de regresión logística de un solo nivel a un modelo de regresión logística multinivel. Para la variable ‘resultado positivo en una prueba de SARS-CoV-2’, el análisis de componentes de la varianza inicial indicaba la necesidad de analizar los niveles 2 a 4 como efectos aleatorios (véase la sección de resultados). Por lo tanto, realizamos un modelo multinivel compuesto por cuatro niveles con los individuos en el nivel inferior.

Sea que y_{ij} denota la respuesta binaria ‘prueba para SARS-CoV-2 positiva’, para el individuo i ($i = 1, \dots, n_j$) en el centro residencial j ($j = 1, \dots, J$), dentro de la zona de salud k ($k = 1, \dots, K$), dentro de la OSI l ($l = 1, \dots, L$).

Este modelo ampliado se escribe mediante las siguientes ecuaciones:

$$y_{ij} \sim \text{Binomial}(1, \pi_{ijkl}), \quad (3)$$

$$\text{logit}(\pi_{ijkl}) = \beta_0 + \beta_1 \text{edad}_{ijkl} + \beta_2 \text{sexo}_{ijkl} + u_j + u_k + u_l \quad (4)$$

$$u_j \sim N(0, \sigma_{uj}^2),$$

$$u_k \sim N(0, \sigma_{uk}^2),$$

$$u_l \sim N(0, \sigma_{ul}^2), \quad (5)$$

donde u denota el efecto aleatorio para los diferentes niveles. Estos efectos se suponen normalmente distribuidos con media cero y varianza σ_u^2 , con el parámetro por estimar.

Tras la estimación, se pueden asignar valores a estos efectos mediante la predicción empírica de Bayes. Estas predicciones \hat{u} se denominan, en ocasiones, estimaciones de contracción, ya que sus valores se contraen hacia la media de la población (de cero) mediante un factor de contracción proporcional a la cantidad de información disponible en cada nivel. La contracción es deseable para evitar que se sobreinterpreten las predicciones asociadas a unidades muy pequeñas, que suelen ser a menudo extremas. La incertidumbre estadística que rodea a estas predicciones también puede calcularse y comunicarse mediante barras de error (por ejemplo, intervalos de confianza del 95%). Esta incertidumbre debe tenerse en cuenta, por ejemplo, a la hora de realizar rankings respecto a las unidades de nivel superior (Leckie y Goldstein, 2009, 2011a y 2011b). Además, la interpretación de los rankings debe asimismo hacerse en relación con el efecto contextual general (Merlo *et al.*, 2012).

El efecto contextual general se evalúa utilizando la varianza estimada entre unidades de nivel superior (por ejemplo, la varianza de los centros residenciales $\hat{\sigma}_u^2$), dado que cuantifica la variabilidad debida a todos aquellos factores inobservados que influyen en la variable respuesta, y que son comunes a todos

los individuos que viven en los mismos centros. Por lo tanto, suponiendo que no hay diferencias de composición en las características individuales no modeladas (por ejemplo, la selección no observada de los individuos a cada centro), se espera que $\hat{\sigma}_u^2$ refleje cualquier efecto general del contexto de los centros, sin indicar ninguna característica contextual específica, sino sólo el número de identificación del centro residencial. No obstante, para estimar este efecto contextual general necesitamos poner la información referida a la varianza entre unidades de segundo, tercer o cuarto nivel, en relación con la varianza total proporcionada por el modelo. Calculamos dos medidas diferentes de los efectos contextuales generales: el coeficiente de correlación intraclase y el incremento del AUC en el paso 2 comparado con el paso 1.

(i) EL COEFICIENTE DE CORRELACIÓN INTRACLASE (CCI)

Elegimos calcular el coeficiente de correlación intraclase (CCI) basándonos en la formulación de respuesta latente del modelo, ya que es el enfoque adoptado de una forma más amplia en los trabajos aplicados. Esta formulación supone que una respuesta continua latente subyace a la respuesta binaria observada, y es esta respuesta latente para la que calcula e interpreta el CCI (Goldstein, Browne y Rasbash, 2002). Cuanto más alto sea este, más relevante será el contexto del centro residencial para comprender la variación de la respuesta latente individual (Merlo *et al.*, 2005, 2006).

$$CCI_{\text{nivel } 4} = \frac{\sigma_{ul}^2}{\sigma_{uj}^2 + \sigma_{uk}^2 + \sigma_{ul}^2 + \frac{\pi^2}{3}} \quad (6)$$

$$CCI_{\text{nivel } 3} = \frac{\sigma_{uk}^2 + \sigma_{ul}^2}{\sigma_{uj}^2 + \sigma_{uk}^2 + \sigma_{ul}^2 + \frac{\pi^2}{3}} \quad (6)$$

$$CCI_{\text{nivel } 2} = \frac{\sigma_{uj}^2 + \sigma_{uk}^2 + \sigma_{ul}^2}{\sigma_{uj}^2 + \sigma_{uk}^2 + \sigma_{ul}^2 + \frac{\pi^2}{3}} \quad (6)$$

donde π es la constante 3,1416, y $\frac{\pi^2}{3}$ denota la varianza de una distribución logística estándar ($\sim 3,29$). Por ejemplo, el $CCI_{\text{nivel } 2}$ puede interpretarse como la correlación en la propensión latente de tener un resultado positivo en una prueba para COVID-19 entre dos individuos seleccionados aleatoriamente del mismo centro, de la misma zona de salud y OSI.

También calculamos el coeficiente de partición de la varianza (*variance partition coefficient*, o VPC por sus siglas en inglés). El VPC no es más que la relación entre la varianza de un nivel específico y la varianza individual total. Puesto que la estructura de la información es jerárquica, el CCI (Fórmula 6) del nivel 4 es también el VPC. Sin embargo, cuando se calcula el VPC para los niveles 3 y 2, los numeradores solo incluyen σ_{uk}^2 y σ_{uj}^2 , respectivamente.

(II) EL INCREMENTO DEL AUC EN EL PASO 2 COMPARADO CON EL PASO 1 (ΔAUC)

En este paso 2, también se estiman las probabilidades previstas o el riesgo absoluto (RA) de las unidades con intervalos de credibilidad (IC) del 96 %. Por ejemplo, el riesgo absoluto para los centros residenciales (RA_{centro}) se calcularía mediante la siguiente fórmula:

$$RA_{centro} = \pi_{jkl} = \text{logit}^{-1}(\beta_0 + u_j + u_k + u_l) = \frac{\exp(\beta_0 + \beta_1 \text{edad}_{ij} + \beta_2 \text{sexo}_{ij} + u_j + u_k + u_l)}{1 + \exp(\beta_0 + \beta_1 \text{edad}_{ij} + \beta_2 \text{sexo}_{ij} + u_j + u_k + u_l)} \quad (7)$$

Utilizando esta información, podemos estudiar las diferencias de riesgo entre los centros residenciales que pertenecen a diferentes zonas de salud y OSI. Utilizamos esta información para representar las diferencias entre centros en la prevalencia de casos positivos de COVID-19.

Las probabilidades de predicción del modelo del paso 2 se basan en las covariables a nivel individual, como en el modelo 1, y en los efectos aleatorios \hat{u}_j , \hat{u}_k y \hat{u}_l (Fórmula 7). Por consiguiente, el AUC del modelo del paso 2 (AUC_2) se puede comparar con el del paso 1 (AUC_1) para cuantificar el incremento del AUC (ΔAUC) obtenido al incorporar en el modelo información sobre la estructura de los datos en niveles. Es decir, nos ofrece información sobre cómo varía la capacidad predictiva del modelo cuando se tiene en cuenta la jerarquía de los datos en centros, zonas de salud y OSI, en lugar de considerar toda la información a un mismo nivel (como ocurre con los modelos de regresión convencionales):

$$\Delta AUC = AUC_1 - AUC_2 \quad (8)$$

Paso 3: el modelo de efectos contextuales específicos

El paso 3 consiste en añadir al modelo las covariables relativas a las características de los centros, para estimar la OR específica de esas variables contextuales. Nos centramos en medir la asociación entre la variable dependiente ‘prueba positiva para SARS-CoV-2’ y tres factores: el territorio histórico en el que se ubica el centro, la titularidad (personalidad jurídica de la entidad titular del centro) y el tamaño (número de plazas). El modelo 4 del paso 3 puede escribirse como:

$$y_{ij} \sim \text{Binomial}(1, \pi_{ij}), \quad (9)$$

$$\text{logit}(\pi_{ijkl}) = \beta_0 + \beta_1 \text{edad}_{ijkl} + \beta_2 \text{sexo}_{ijkl} + \beta_3 \text{territorio}_{jkl} + \beta_4 \text{titular}_{jkl} + u_j + u_k + u_l \quad (10)$$

Los efectos contextuales específicos se evalúan midiendo las asociaciones (es decir, las OR) entre las variables contextuales específicas (características de los centros residenciales) y la variable respuesta a nivel individual. Con todo, en la regresión multinivel, las OR de las variables contextuales no tienen la interpretación habitual de

‘asociación media en la población estudiada’, sino una interpretación específica dentro de cada grupo (Larsen *et al.*, 2000; Larsen y Merlo, 2005; Merlo *et al.*, 2006). Es importante destacar que los efectos contextuales específicos deben interpretarse junto a los efectos contextuales generales para llegar a una interpretación correcta del impacto que ejerce el contexto en la variable respuesta (Merlo *et al.*, 2018).

El paso 3 también permite comprender el mecanismo que subyace a los efectos contextuales generales observados a nivel de centro. Para ello, calculamos el cambio proporcional de la varianza (*proportional change in variance*, o PCV por sus siglas en inglés), que se define como la proporción de la varianza entre centros residenciales (que se corresponde con la varianza contextual total) en el modelo del paso 2, explicada al añadir información contextual específica (es decir, al añadir las variables ‘tamaño’, ‘titularidad’ y ‘territorio histórico’) en el modelo del paso 3.

$$PVC = \frac{\text{Varianza contextual total}_{\text{paso 2}} - \text{Varianza contextual total}_{\text{paso 1}}}{\text{Varianza contextual total}_{\text{paso 2}}} \times 100 \quad (11)$$

Expresamos el PCV en forma de porcentaje. Un PCV grande sugeriría que el efecto contextual general está sustancialmente mediado por esas variables específicas.

Obsérvese que en el paso 2 calculamos el AUC como una forma de cuantificar el efecto contextual general de los centros más allá de las características del individuo que reside en ellos. En el paso 3, incluimos las características contextuales específicas (es decir, las variables tamaño, titularidad y territorio histórico). Sin embargo, añadir estas variables contextuales específicas no puede aumentar el AUC obtenido en el modelo del paso 2, ya que este modelo da el máximo AUC que puede obtenerse combinando la información individual disponible y la identidad del centro residencial. Esta identidad capta la totalidad de los factores potencialmente observables, pero también inobservables, relacionados con el centro. La inclusión de variables específicas del centro como efectos fijos explicará parte de la varianza contextual y, por ejemplo, disminuirá el tamaño absoluto medio de los estimadores u_j , u_k , u_l , reduciendo así el papel que juegan los efectos aleatorios en la predicción. Sin embargo, este cambio en la especificación del modelo mejora simultáneamente la predicción mediante la incorporación de los coeficientes de regresión para el territorio histórico, la titularidad o el tamaño. En otras palabras, el efecto contextual general del paso 2 se descompone en el paso 3 en un efecto contextual específico y un efecto general residual que, juntos, dan los mismos valores de AUC que en el paso 2.

3.2.2. Modelos de mortalidad

En el caso de la mortalidad, el análisis inicial de componentes de la varianza indicó que sólo era necesario incluir a las personas residentes (nivel 1) agrupadas en centros residenciales (nivel 2), ya que los niveles 3 y 4 (zona de salud y OSI) capturaban

una parte muy pequeña de la varianza total en el riesgo de mortalidad, y no resultan relevantes (véase la sección de resultados). Los modelos de mortalidad fueron muy similares a los elaborados para la variable respuesta ‘test SARS-CoV-2 positivo’, excepto por algunas peculiaridades.

En el paso 1, además de las variables sexo y edad utilizadas anteriormente, incluimos la variable ‘resultado en el test de SARS-CoV-2’ como efecto fijo en el modelo 3:

$$\text{logit}(\pi_i) = \beta_0 + \beta_1 \text{edad}_i + \beta_2 \text{sexo}_i + \beta_3 \text{covid}_i \quad (12)$$

donde π_i denota la probabilidad de fallecimiento de un individuo, dadas sus características individuales edad, sexo, y COVID.

En el paso 2 (modelo 4) incluimos dos efectos aleatorios: uno para los centros residenciales (u_{0j}) y otro para la pendiente de la asociación entre la variable ‘resultado en el test de SARS-CoV-2’ y la mortalidad (u_{3j}). Es decir, abandonamos el supuesto de que la pendiente de esta asociación es la misma en todas las residencias. En lugar de ello, permitimos que sea diferente en las distintas residencias:

$$\text{logit}(\pi_{ij}) = \beta_0 + \beta_1 \text{edad}_{ij} + \beta_2 \text{sexo}_{ij} + \beta_3 \text{covid}_{ij} + u_{0j} + u_{3j} + u_{1j} \beta_3 \text{covid}_{ij} \quad (13)$$

$$\begin{bmatrix} u_{0j} \\ u_{3j} \end{bmatrix} \sim \text{BND}(0, \Omega_u), \Omega_u = \begin{bmatrix} \sigma_{u0}^2 & \sigma_{u03}^2 \\ \sigma_{u03}^2 & \sigma_{u3}^2 \end{bmatrix} \quad (14)$$

donde los residuos del intercepto (u_{0j}) y la pendiente (u_{3j}) siguen una distribución normal bivalente (BND) con media 0 y una matriz de covarianza de los efectos aleatorios Ω_u .

Por lo tanto, la varianza entre centros residenciales se convierte en una función de la variable ‘resultado en el test de SARS-CoV-2’, de manera que esta varianza puede ser diferente para los individuos con un resultado positivo en el test de COVID-19 (COVID = 1), en comparación con los individuos que han tenido una prueba con resultado negativo (COVID = 0):

$$\text{Varianza (COVID = 1)} = \sigma_{u0}^2 + 2\sigma_{u03}^2 \mathcal{X}_{ij} + \sigma_{u3}^2 \mathcal{X}_{ij}^2 \quad (15)$$

$$\text{Varianza (COVID = 0)} = \sigma_{u0}^2 \quad (16)$$

En consecuencia, también hay dos VPC y dos AUC, uno para los individuos con resultado COVID-19 positivo y otro para los individuos con resultado COVID-19 negativo.

3.2.3. Estimación de los modelos

Los modelos se estimaron utilizando métodos de Monte Carlo con cadena de Markov (MCMC), tal y como se implementa en el software de modelización multinivel MLwiN (Browne, 2019). Especificamos distribuciones previas difusas (vagas, planas o mínimamente informativas) para

todos los parámetros. Utilizamos la estimación de cuasi verosimilitud para proporcionar buenos valores de partida para todos los parámetros. Para cada modelo, especificamos una longitud de *burn-in* de 5.000 iteraciones y una longitud de cadena de seguimiento de 10.000 iteraciones. Las evaluaciones visuales de las cadenas de parámetros y los diagnósticos de convergencia MCMC estándar sugieren que las longitudes de estos periodos son suficientes. El criterio de información de desviación bayesiano (DIC) se utilizó como medida de la corrección del ajuste de nuestros modelos (Spiegelhalter *et al.*, 2002). El DIC considera tanto la desviación del modelo como su complejidad. Los modelos con un DIC menor se prefieren a los modelos con un DIC mayor, y las diferencias de cinco o más se consideran sustanciales (Lunn *et al.*, 2012).

Obsérvese que ajustamos los modelos mediante un enfoque totalmente bayesiano, utilizando métodos MCMC. Sin embargo, dado que los resultados son muy similares, nos limitamos a predecir los efectos aleatorios mediante una estimación de máxima verosimilitud (es decir, la predicción empírica de Bayes) (Browne, 2019).

4. Resultados

La base de datos obtenida de Osakidetza contenía inicialmente información relativa a 22.773 personas que figuraban como usuarias de centros residenciales de Euskadi entre el 1 de marzo y el 31 de octubre de 2020. La información relativa al primer centro de residencia de cada persona en la base de datos de Osakidetza se cruzó con el registro de centros aportado por las diputaciones forales al Ararteko para el estudio. Se eliminaron los datos relativos a 2.193 personas que habían estado alojadas en centros residenciales que no figuraban en dicho registro, por tratarse de centros dirigidos a otros colectivos (fundamentalmente, personas con discapacidad y con enfermedad mental). De las 20.580 personas restantes, identificadas como usuarias de servicios residenciales para personas mayores, se incluyeron en el análisis 20.186 casos, tras eliminar los datos relativos a los establecimientos que funcionaron como centros de referencia en COVID-19 durante los meses de pandemia. Estos casos se excluyeron del análisis por considerar que, en ellos, no resultaba posible diferenciar adecuadamente la información relativa a las personas que residían en esos centros con anterioridad al mes de marzo y las que fueron trasladadas desde domicilios particulares para realizar convalecencias o aislamientos.

En consecuencia, la cohorte analizada está formada por 20.186 personas, 13.765 mujeres (68,2%) y 6.421 hombres (31,8%), la mayoría de ellas (60,1%) mayores de 85 años y un 7,7% menores de 65 años. El 94,5% de estas personas vivieron en un único centro residencial entre marzo y finales de octubre de 2020, mientras que un 5,5% pasaron por dos o

más. De acuerdo con la información recabada de Osakidetza, 4.121 de las 18.634 personas mayores (22,1%) y 250 de las 1.552 menores de 65 años (16,1%) que vivían en centros residenciales de Euskadi tuvieron al menos una prueba positiva para SARS-CoV-2 durante el periodo de estudio (del 1 de marzo al 31 de octubre de 2020). La tasa media de positividad en los establecimientos con al menos un caso positivo fue del 23,3% (desviación estándar = 23,09), con una gran variabilidad entre centros (mediana = 12,5; rango intercuartil = 32,9).

Hasta el 10 de enero de 2021, 910 de las 4.121 personas usuarias mayores infectadas (22,1%) y 14 de las 250 menores de 65 años infectadas (5,6%) fallecieron en los dos meses posteriores al diagnóstico. La tasa de letalidad media fue del 17,5% (desviación estándar = 19,3) para el conjunto de la muestra analizada, y del 28,7% (desviación estándar = 18,7) para los mayores de 65 años. En total, 4.507 personas usuarias mayores y 91 usuarias menores de 65 años fallecieron en el período de estudio, lo que resultó en una tasa de mortalidad por todas las causas del 24,2% y el 5,8%, respectivamente.

Del análisis preliminar de componentes de la varianza sobre modelos multinivel vacíos, que incluían únicamente efectos aleatorios para los

cuatro niveles de agrupación de los datos planteados inicialmente, se obtuvieron resultados significativos para los tres niveles contextuales (centro, zona de salud y OSI) al modelar la variable ‘test positivo’, pero únicamente para el centro residencial al modelar la variable ‘fallecimiento por cualquier causa’. En efecto, la OSI tan solo captura un 0,23% de la varianza total en la mortalidad, y la zona de salud, apenas un 0,15% (Tabla 5). En consecuencia, a la hora de definir los modelos para la variable dependiente ‘prueba positiva para SARS-CoV-2’ se incluyeron los cuatro niveles —(1) persona, (2) centro residencial, (3) zona de salud y (4) OSI—, mientras que al modelar la variable ‘fallecimiento por cualquier causa’ se utilizaron únicamente dos —(1) persona usuaria y (2) centro—.

La Tabla 6 recoge los resultados de los modelos de regresión estimados para la variable dependiente ‘prueba positiva para SARS-CoV-2’. Lo primero que cabe destacar de los resultados obtenidos es la importancia de los efectos contextuales generales —con un coeficiente de correlación intraclase (CCI) del 55,1% para los cuatro niveles considerados— y fundamentalmente de los efectos de nivel de centro, con un coeficiente de partición de la varianza (VPC) del 49,9%. El área bajo la curva COR (AUC) mejora sustancialmente entre el modelo 2 y el modelo

Tabla 4. Características de la cohorte incluida en el análisis de regresión logística multinivel*

		N	% vertical
Sexo	Mujer	13.765	68,2
	Hombre	6.421	31,8
Edad (años)	< 65	1.552	7,7
	65-74	1.799	8,9
	75-84	4.695	23,3
	≥ 85	12.140	60,1
Nº de centros en los que ha vivido	1	19.076	94,5
	≥ 2	1.110	5,5
Test SARS-CoV-2 positivo	No	15.815	78,3
	Sí	4.371	21,7
Fallecido	No	15.588	77,2
	Sí	4.598	22,8
Total		20.186	100,0

* Personas que vivían en centros residenciales de Euskadi entre el 1 de marzo y el 31 de octubre de 2020, excluidos los casos correspondientes a los centros de referencia en COVID-19.

Fuente: Informe de indicadores de la COVID-19 del sistema Osakidetza Business Intelligence-Oracle Analytic Server (OBI-OAS).

Tabla 5. Componentes de la varianza en un modelo de regresión logística multinivel “vacío” en cuatro niveles para la variable dependiente ‘fallecimiento por cualquier causa’

	OSI (n = 13)	Zona de salud (n = 110)	Centro (n = 259)
Varianza	0,01 (0,00-0,03)	0,01 (0,00-0,02)	0,15 (0,11-0,20)
CCI (%)	0,23	0,38	4,71
VPC	0,23	0,15	4,33

OSI: organización sanitaria integrada. CCI: coeficiente de correlación intraclase. VPC: coeficiente de partición de la varianza.

Fuente: Osakidetza Business Intelligence-Oracle Analytic Server (OBI-OAS).

3, al tomar en consideración la agrupación de las observaciones a distintos niveles (centro, zona de salud y OSI), pasando de 0,52 a 0,85, lo que indica que la capacidad predictiva del modelo mejora sustancialmente.

No se aprecian diferencias significativas en el riesgo de contagio por sexo, ni tampoco por edad, entre los mayores de 65 años. Las personas menores de 65 años tuvieron entre un 20% y un 37% menos de riesgo de tener una prueba de SARS-CoV-2 positiva en comparación con las personas de entre 65 y 74 años en los mismos centros (OR = 0,65; IC95% 0,53-0,80).

Los efectos contextuales específicos se pueden evaluar comparando los resultados en los modelos 3, 4 y 5. No se aprecian diferencias significativas en el riesgo de contagio para las personas residentes en función de la titularidad de los centros, pero sí en función del territorio y del tamaño de aquellos. En

relación con el territorio, se aprecia un mayor riesgo en Álava (OR = 6,57; IC95% 2,3-18,3) y Bizkaia (OR = 4,1; IC95% 2,4-6,9) en comparación con Gipuzkoa. En relación con el tamaño, las personas alojadas en centros con un mayor número de plazas tuvieron un riesgo mayor de contagio: entre 3 y 12 veces más las residentes en centros de 26 a 70 plazas que las residentes en centros de 25 plazas o menos (OR = 6,3; IC95% 3,1-12,4), y entre 5 y 21 veces más las residentes en centros de más de 70 plazas. La varianza a nivel de centro disminuye muy ligeramente cuando se introducen en el modelo las variables contextuales 'territorio histórico' y 'tamaño', pero sin embargo la varianza a nivel de zona de salud y OSI se ve incrementada. Estos resultados sugieren que las características de centro incluidas en el modelo, a pesar de estar fuertemente asociadas con el riesgo de contagio —tal y como indican las elevadas *odds ratios* obtenidas—, explican, sin embargo, una parte muy pequeña de la varianza inicialmente

Tabla 6. Resultados de un modelo de regresión logística multinivel en cuatro niveles (OSI, zona de salud, centro residencial, persona residente) para la variable 'test positivo para SARS-CoV-2'. Población residente en centros residenciales para personas mayores de Euskadi entre el 1 de marzo y el 31 de octubre de 2020 (n = 20.186)

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Odds ratios (IC95%)						
Edad (años)	< 65	0,68 (0,57-0,81)	0,68 (0,57-0,81)	0,65 (0,52-0,79)	0,66 (0,53-0,81)	0,65 (0,53-0,80)
	65-74	Referencia	Referencia	Referencia	Referencia	Referencia
	75-84	1,09 (0,96-1,25)	1,09 (0,95-1,24)	1,04 (0,88-1,21)	1,04 (0,89-1,21)	1,04 (0,89-1,19)
	≥ 85	1,13 (1,00-1,27)	1,12 (0,99-1,27)	1,12 (0,97-1,30)	1,12 (0,97-1,30)	1,12 (0,97-1,27)
Sexo	Mujer		Referencia	Referencia	Referencia	Referencia
	Hombre		0,98 (0,91-1,06)	0,93 (0,86-1,01)	0,93 (0,85-1,02)	0,93 (0,85-1,02)
Titularidad	Pública				1,39 (0,72-2,39)	0,89 (0,46-1,67)
	Privada mercantil				Referencia	
	Privada sin fin lucrativo				1,12 (0,46-2,31)	0,65 (0,33-1,09)
Territorio histórico	Álava				3,98 (1,94-6,48)	6,57 (2,3-18,83)
	Bizkaia				2,36 (1,56-4,03)	4,10 (2,36-6,92)
	Gipuzkoa				Referencia	Referencia
Tamaño del centro (nº plazas)	≤ 25					Referencia
	26-70					6,32 (3,1-12,4)
	> 70					11,26 (5,2-21,1)
Componentes de la varianza						
OSI (n _i = 13)	Varianza			0,19 (0,00-0,77)	0,09 (0,00-0,45)	0,24 (0,00-0,99)
	VPC (%)			2,51	1,26	3,26
	CCI (%)			2,51	1,26	3,26
Zona de salud (n _k = 110)	Varianza			0,12 (0,00-0,51)	0,02 (0,02-0,12)	0,15 (0,00-0,44)
	VPC (%)			1,59	0,29	2,00
	CCI (%)			4,10	1,55	5,26
Centro residencial (n _j = 259)	Varianza			3,98 (3,09-5,13)	4,08 (3,24-5,07)	3,66 (2,83-4,66)
	VPC (%)			52,5	54,47	49,88
	CCI (%)			56,6	56,02	55,14
Área bajo la curva COR (AUC)		0,52 (0,51-0,53)	0,52(0,51-0,53)	0,85 (0,84-0,85)	0,85(0,84-0,85)	0,85 (0,84-0,85)

OSI: organización sanitaria integrada. VPC: coeficiente de partición de la varianza. CCI: coeficiente de correlación intraclase. Modelos 1 y 2: regresión logística multivariante en un solo nivel. Modelos 3 a 5: regresión logística multinivel (cuatro niveles).

Fuente: Osakidetza Business Intelligence-Oracle Analytic Server (OBI-OAS).

grande que se da a nivel de centro residencial. En definitiva, la mayor parte de las diferencias en el riesgo de contagio que se deben al centro residencial permanecen inexplicadas en el modelo final estimado, tras incluir las variables ‘territorio histórico’, ‘tamaño’ y ‘titularidad de los centros’.

Pasando a analizar los resultados obtenidos en los modelos de mortalidad en la Tabla 7, el modelo de efectos contextuales generales (modelo 4) indica que la parte de la varianza que puede ser explicada por el centro residencial es, en este caso, mucho menor que al analizar la variable de contagios. Entre un 3% y un 5% de la varianza en el riesgo de mortalidad depende del centro residencial (frente a un 49% en el riesgo de contagio), siendo la importancia del centro menor para las personas que resultaron contagiadas por COVID-19 que para aquellas que no se infectaron. Para las personas contagiadas, en torno a un 3,5% de las diferencias en el riesgo de mortalidad se deben a factores de nivel de centro, mientras que para las no contagiadas la relevancia del centro es algo mayor, del 5,1%. Estos resultados sugieren que, al contrario de lo que ocurría con los contagios por coronavirus, el riesgo de mortalidad depende fundamentalmente de características individuales. A pesar de que la varianza a nivel de centro es reducida en comparación con la que se obtenía en los modelos para la variable de contagios, el incremento del área bajo la curva COR (AUC) entre el modelo 3 y el modelo 4 indica una mejora sustancial de la capacidad predictiva al tener en cuenta el agrupamiento de residentes en centros, ya que permite pasar de una tasa de acierto del 62% cuando se realiza la predicción basándose únicamente en el sexo, la edad y el estatus frente a la infección de las personas, a una del 70% en los modelos que incluyen la información sobre el centro residencial en el que vivió la persona.

Por lo que se refiere a las características individuales analizadas, los resultados indican un mayor riesgo de mortalidad para los hombres que para las mujeres, y también para las personas de más edad. En el modelo final estimado (modelo 6), a igual edad y estatus frente a la infección, los residentes hombres tendrían entre un 52% y un 77% más riesgo de fallecer que las mujeres en centros con características similares; las personas con edades comprendidas entre los 75 y 84 años tendrían entre un 48% y un 98% más de riesgo que las de 65 a 74 años; y para las de más de 85 años, el riesgo sería tres veces mayor (OR = 2,9; IC95% 2,5-3,35).

El modelo también ofrece información sobre el efecto que haber resultado infectado por SARS-CoV-2 tiene en el riesgo de mortalidad. Por una parte, los efectos fijos para la variable que recoge el resultado de la prueba para cada persona, en el modelo final estimado, indican que el riesgo de mortalidad fue entre un 21% y un 57% mayor para las personas infectadas que para las no infectadas (OR = 1,39; IC95% 1,21%-1,57%). No obstante, el incremento del área bajo la curva COR (AUC) entre

el modelo 2 y 3 es muy modesto, de tan solo 2 puntos porcentuales, lo que indica que la capacidad predictiva del modelo no mejora sustancialmente al introducir la variable que recoge el resultado frente a la infección, respecto a un modelo ajustado únicamente por sexo y edad. En definitiva, los resultados sugieren que la variabilidad individual en el riesgo de mortalidad se debe fundamentalmente a características personales que permanecen inobservadas en el modelo.

Por último, en lo referente a los efectos contextuales específicos, los resultados de los modelos 5 y 6 indican que ninguna de las características de nivel de centro analizadas tiene un efecto muy relevante en la mortalidad. Se obtienen asociaciones significativas para las variables ‘titularidad’ y ‘tamaño’, indicando que las personas alojadas en centros pertenecientes a entidades sin fin de lucro tendrían entre un 3% y un 30% menos de riesgo que las que vivían en centros pertenecientes a entidades mercantiles, y las usuarias de centros pequeños (menos de 25 plazas), entre un 4% y un 42% menos de riesgo que las de centros de entre 26 y 70 plazas. Sin embargo, por una parte, estos efectos contextuales específicos deben interpretarse en relación con los efectos contextuales generales, y, como se ha dicho, estos indican que el riesgo de mortalidad depende fundamentalmente de características personales, siendo la variabilidad debida al centro residencial del 5% como máximo. Por otra parte, si bien tener en cuenta esta variabilidad debida al centro mejoraría sensiblemente la capacidad predictiva del modelo (incrementa el AUC de 0,63 a 0,7), la introducción de las variables específicas ‘titularidad’, ‘tamaño’ y ‘territorio histórico’ en los modelos 5 y 6 no reduce la varianza correspondiente al centro residencial respecto a la que se obtenía en el modelo de efectos contextuales generales (modelo 4), lo que indica que las características analizadas no consiguen explicar las diferencias en mortalidad que se deben al centro residencial.

5. Discusión

En el último año y medio, se ha escrito mucho sobre los factores de riesgo —individuales y de nivel de centro— que han podido intervenir en la rápida expansión y la elevada mortalidad por COVID-19 entre las personas mayores usuarias de centros residenciales, en todo el mundo. El tamaño de los centros (Abrams *et al.*, 2020; Burton *et al.*, 2020; Emmerson *et al.*, 2020; Fresno *et al.*, 2020), la titularidad (Stall *et al.*, 2020; Abrams *et al.*, 2020; Burton *et al.*, 2020; Fresno *et al.*, 2020; Junta de Castilla y León, 2020), las ratios de atención (Li *et al.*, 2020) o la contratación de personal eventual y la rotación del personal (Tinsley, 2020; Chen, Chevalier y Long, 2020; McMichael *et al.*, 2020) son algunos de los muchos factores que se han analizado y para los que diversos estudios han mostrado la existencia de asociaciones significativas con los contagios o la mortalidad.

Tabla 7. Resultados de un modelo de regresión logística multinivel en dos niveles (centro residencial, persona residente) para la variable dependiente ‘mortalidad por cualquier causa’. Población residente en centros residenciales para personas mayores de Euskadi entre el 1 de marzo y el 31 de octubre de 2020 (n = 20.186)

		Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6			
Odds ratios										
Edad (años)	< 65	0,38 (0,30-0,49)	0,40 (0,31-0,50)	0,40 (0,32-0,51)	0,40 (0,31-0,51)	0,39 (0,31-0,49)	0,39 (0,31-0,50)			
	65-74	Referencia	Referencia	Referencia	Referencia	Referencia	Referencia			
	75-84	1,52 (1,31-1,76)	1,67 (1,44-1,93)	1,66 (1,43-1,93)	1,73 (1,50-2,00)	1,71 (1,48-1,98)	1,72 (1,47-2,00)			
	≥ 85	2,27 (1,98-2,60)	2,71 (2,35-3,11)	2,70 (2,35-3,10)	2,92 (2,54-3,36)	2,87 (2,51-3,29)	2,90 (2,50-3,35)			
Sexo	Mujer		Referencia		Referencia	Referencia	Referencia			
	Hombre		1,63 (1,52-1,76)	1,64 (1,53-1,76)	1,64 (1,52-1,77)	1,64 (1,52-1,77)	1,64 (1,52-1,76)			
Test SARS-CoV-2 positivo	No				Referencia	Referencia	Referencia			
	Sí			1,42 (1,31-1,53)	1,45 (1,28-1,63)	1,44 (1,27-1,61)	1,39 (1,21-1,57)			
Titularidad	Pública					0,97 (0,82-1,16)	0,98 (0,86-1,12)			
	Mercantil					Referencia	Referencia			
	Privada sin fin lucrativo					0,86 (0,73-1,02)	0,83 (0,70-0,97)			
Tamaño (plazas)	≤ 25						0,74 (0,58-0,96)			
	26-70						Referencia			
	> 70						1,08 (0,95-1,23)			
Territorio histórico	Álava					0,98 (0,81-1,18)	1,02 (0,84-1,24)			
	Bizkaia					1,03 (0,89-1,18)	1,06 (0,91-1,22)			
	Gipuzkoa					Referencia	Referencia			
Componentes de la varianza										
Varianza (centro residencial)					0,18 (0,13-0,24)	0,18 (0,13-0,24)	0,18 (0,13-0,24)			
Varianza (test SARS-CoV-2 positivo)					0,26 (0,16-0,39)	0,25 (0,16-0,38)	0,26 (0,16-0,39)			
Covarianza (centro, test positivo)					0,16 (0,23-0,10)	0,16 (0,23-0,10)	0,16 (0,23-0,10)			
VPC (%)				5,20	No infec. 3,30	Infec. 5,20	No infec. 3,30	Infec. 5,10	No infec. 3,50	Infec. 3,50
Área bajo la curva COR (AUC)		0,59 (0,58-0,60)	0,61 (0,61-0,62)	0,63 (0,62-0,63)	0,69 (0,69-0,69)	0,70 (0,70-0,70)	0,70 (0,70-0,70)			

Modelos 1 a 3: regresión logística multivariante en un solo nivel. Modelos 4 a 6: regresión logística multinivel (dos niveles).

VPC: coeficiente de partición de la varianza. No infec.: resultado negativo en los test diagnósticos para SARS-CoV-2. Infec.: resultado positivo en al menos un test diagnóstico para SARS-CoV-2.

Fuente: Osakidetza Business Intelligence-Oracle Analytic Server (OBI-OAS).

Por lo general, este tipo de trabajos se centran en analizar las diferencias medias entre grupos. Por ejemplo, el riesgo “medio” de contagiarse para personas que viven en centros públicos se compara con el riesgo “medio” para personas que viven en centros privados, y este efecto se comunica como una *odds ratio* media. En este tipo de análisis, la heterogeneidad individual dentro de cada grupo se ignora, utilizándola únicamente para calcular la incertidumbre estadística en torno a la estimación de las diferencias entre las medias de los grupos. Sin embargo, algunos expertos del ámbito de la epidemiología social han mostrado la necesidad de relacionar ambos tipos de información —la referida a las asociaciones medias entre grupos con la variabilidad individual dentro de cada grupo— a la hora de buscar factores realmente

relevantes en la clínica (Duncan, Jones y Moon, 1998; Merlo, 2014 y 2003; Merlo *et al.*, 2009). Lo que plantean estos autores es que, en situaciones en las que la heterogeneidad individual intragrupal es grande en comparación con las diferencias entre grupos, conocer el comportamiento “medio” de los grupos tiene escasa relevancia para predecir el comportamiento individual (Rose, 2001).

En definitiva, lo que argumentan es que, sin información previa sobre la relevancia de determinado contexto (por ejemplo, el centro residencial) en la condición de salud que se desea investigar, la búsqueda de factores de riesgo se hace a ciegas, pudiendo encontrar asociaciones fuertes a nivel grupal que, sin embargo, son escasamente relevantes en la clínica. Esta noción está relativamente clara en otras

ramas de la epidemiología, como la que se ocupa de la identificación de nuevos factores de riesgo y biomarcadores, o de la evaluación de test diagnósticos y de cribado. En esos ámbitos de investigación, es bien sabido que las medidas de asociación media como las *odds ratios* proporcionan información muy limitada para evaluar la capacidad diagnóstica o pronóstica de un test o biomarcador. En consecuencia, la regla en estos ámbitos de investigación es que las medidas de asociación deben interpretarse junto con información sobre la capacidad discriminatoria como, por ejemplo, el área bajo la curva COR (AUC), que se elabora a partir de información sobre la sensibilidad (capacidad para detectar verdaderos positivos) y especificidad (capacidad para descartar falsos positivos) del test o marcador evaluado.

La principal aportación del estudio realizado por el SIIS para el Ararteko, con la colaboración de la Unidad de Epidemiología Social de la Universidad de Lünd, ha sido, precisamente, el incorporar este análisis previo de la heterogeneidad individual y de la capacidad discriminatoria del contexto residencial, a la hora de buscar factores que han podido incrementar el riesgo de contagio y de mortalidad por COVID-19 para la población mayor que vive en esos centros. El análisis se ha realizado utilizando técnicas de regresión logística multinivel, que permiten descomponer la varianza total en sus componentes individual y grupal, utilizando los estadísticos coeficiente de correlación intraclase (CCI) y componentes de partición de la varianza (VPC), que ofrecen una medida de la relevancia que ha tenido el contexto (en este caso, el centro residencial, la zona de salud y la OSI) en el riesgo de contagio y de mortalidad. De manera complementaria, las probabilidades estimadas por los modelos se han utilizado para calcular el área bajo la curva COR, lo que permite comparar la capacidad predictiva⁹ de modelos basados únicamente en información individual de las personas residentes, con la de modelos que incorporan información sobre la agrupación de estas personas en centros residenciales, zonas de salud y OSI.

El hallazgo más relevante del estudio es que el riesgo de contagio por SARS-CoV-2 para las personas mayores que han vivido en centros residenciales de Euskadi entre marzo y octubre de 2020 ha dependido en gran medida del propio centro residencial. Los resultados indican que un 55,1% de las diferencias individuales en el riesgo de contagio se deben al contexto (49,9% al centro residencial, 2% a la zona de salud y un 3,3% a la OSI en la que se ubica el centro) y solamente un 44,9% a las características personales.

Estos resultados coinciden con los obtenidos por Mehta y colaboradores (2021) en el que sea, probablemente, el estudio más completo publicado

⁹ La capacidad predictiva de los modelos se refiere a la capacidad para identificar correctamente a personas residentes contagiadas o que han fallecido a partir de los valores de las variables predictoras.

hasta el momento sobre los factores asociados con la infección y la mortalidad por COVID-19 en centros residenciales. Utilizando una metodología multinivel similar a la empleada en el presente estudio, estos autores estiman que el riesgo de infección para las más de 480.000 personas que vivían en centros residenciales de los Estados Unidos entre abril y septiembre de 2020 se asociaba fundamentalmente con el centro residencial y el condado de residencia. En su estudio, el centro residencial captura en torno al 37% de la variabilidad individual en el riesgo de infección y el condado, un 23%, por lo que la relevancia del centro es algo menor que en nuestro estudio y, en cambio, la del condado sería mayor que la que hemos estimado para las zonas de salud en Euskadi. Con todo, los resultados apuntan en el mismo sentido, indicando que las variables y factores contextuales (relativos al centro residencial o la zona) podrían explicar gran parte de las diferencias en el riesgo de contagio.

El centro residencial se confirma, por lo tanto, en Euskadi, como un contexto altamente relevante en la prevención de los contagios por COVID-19, lo que da sentido a la búsqueda de factores —características del centro, medidas preventivas— que podrían contribuir a reducir el riesgo de contagio. Para ilustrar esta relevancia del contexto, se ha estimado que, al intentar predecir si una persona usaría tomada al azar ha resultado contagiada basándose únicamente en su edad y sexo, la tasa de acierto sería del 52%, es decir, sería equivalente a realizar una predicción prácticamente al azar; si, además, se incorporara a la predicción información sobre el centro residencial concreto en el que ha vivido la persona, la tasa de acierto mejoraría hasta el 85%.

La confirmación de la relevancia del centro residencial en el riesgo de contagio es importante, porque indica que intervenir de forma preventiva a ese nivel puede resultar efectivo. Con todo, no implica que conozcamos los factores específicos del centro que permitirían explicar las diferencias en el riesgo. Una limitación importante del estudio realizado para el Ararteko es que no se pudo acceder a información suficientemente completa sobre las características de los centros y las medidas preventivas aplicadas. En consecuencia, se sabe poco acerca de los factores de centro que pueden ser determinantes.

Una conclusión relativamente clara que se ha podido obtener es que, en Euskadi, la titularidad de los centros no se asocia con el riesgo de contagio. Estudios anteriores, provenientes fundamentalmente de los Estados Unidos, habían establecido una asociación entre el ánimo de lucro o la pertenencia a cadenas y un mayor riesgo de contagio (Stall *et al.*, 2020; Abrams *et al.*, 2020). La presencia relativamente reducida de grandes entidades mercantiles entre las titulares de centros residenciales en Euskadi y, en general, el carácter más mercantilista del sector residencial en los Estados Unidos podrían ser parte de la explicación

de los distintos resultados obtenidos en relación con la relevancia de la titularidad de los centros. Por otra parte, cabría analizar en el futuro si otros indicadores relacionados —como el beneficio empresarial de las entidades titulares o el número de centros residenciales gestionados por cada entidad— tienen mayor capacidad predictiva.

Por otra parte, sí se ha encontrado una asociación relativamente fuerte entre las otras dos variables de centro analizadas —el territorio histórico y el tamaño— y el riesgo de contagio. Los resultados obtenidos no dejan lugar a dudas de que el riesgo de contagio ha sido más elevado en Álava y en Bizkaia que en Gipuzkoa (al menos hasta octubre de 2020), y entre las personas que vivían en centros con un mayor número de plazas, en relación con las que vivían en centros de menor tamaño. La magnitud de la influencia que ejercen esas variables no se ha podido estimar, sin embargo, de forma muy precisa; los intervalos de confianza obtenidos para las *odds ratios* son muy amplios en ambos casos, lo que indica que existe incertidumbre en la estimación. Por otra parte, la varianza a nivel de centro apenas se reduce cuando se introducen estas variables en el modelo, lo que indica que explican una parte pequeña de la varianza inicialmente grande que se da a nivel de centro residencial. Estos resultados sugieren que las grandes diferencias que se observan en el riesgo de contagio por tamaño y territorio histórico están probablemente mediadas por otros factores que, al estar a su vez relacionados con estos, explicarían las diferencias observadas.

De acuerdo con la literatura revisada y la información aportada por las personas expertas que participaron en el grupo de discusión para el estudio del Ararteko, algunas variables clave para explicar las diferencias por tamaño podrían ser las relacionadas con las características estructurales de los centros —espacio disponible por persona, número de salas o espacios compartidos, características de los sistemas de ventilación, disponibilidad de espacios para realizar aislamientos— y los flujos de personas en el centro —rotación del personal, número de proveedores de productos y servicios— que podrían variar mucho entre centros de distinto tamaño. En relación con el territorio histórico, factores relevantes podrían ser las tasas de incidencia del virus en el entorno comunitario de los centros y la temporalidad en la aplicación de las medidas preventivas, dado que su implantación se ha realizado en gran medida a escala territorial.

Una cuestión a la que la literatura analizada ha otorgado mucha relevancia —y que apenas se ha podido analizar en nuestro estudio— es la relativa al papel que han tenido en la transmisión del virus las y los trabajadores contratados por varios centros residenciales o sanitarios de manera simultánea o intermitente. Aunque las personas expertas consultadas para el estudio consideraron, en general, que este no había sido un problema importante en Euskadi, se trata de una cuestión que

no se ha podido analizar en detalle por el momento. En cualquier caso, dado que el presente estudio confirma que los factores relativos al centro pueden explicar gran parte de las diferencias en el riesgo de contagio, futuras investigaciones deberían tratar de profundizar en el conocimiento de estos aspectos.

Por lo que se refiere a la mortalidad, el análisis realizado revela que, junto con la edad y el sexo (se confirma un mayor riesgo de mortalidad para los hombres, a igual edad), haber tenido un resultado positivo en una prueba diagnóstica para SARS-CoV-2 incrementa el riesgo de mortalidad entre un 21% y un 57%. En este caso, la relevancia del centro residencial es mucho menor: la variabilidad a nivel de centro representa un 5% de la variabilidad total en el riesgo de fallecimiento, de lo que se deduce que los factores más importantes respecto al riesgo de mortalidad son las características individuales de las personas residentes. Respecto a estas, en futuras investigaciones se debería tratar de incorporar al análisis información relativa a las condiciones de salud preexistentes o al grado de dependencia de las personas usuarias, ya que varios estudios han señalado que hay una fuerte asociación entre estas variables y la mortalidad (Mehta *et al.*, 2021; Panagiotou *et al.*, 2021; Mas Romero *et al.*, 2020; Causa *et al.*, 2021; Suñer *et al.*, 2021). Ello requerirá, por una parte, mejorar los sistemas de información en el ámbito de los servicios sociales y su interoperabilidad. Sería asimismo conveniente, dadas las dificultades que han tenido las administraciones para monitorizar la situación en los centros durante la pandemia, avanzar en la definición de un conjunto mínimo de datos que aporte información individual y comparable sobre las personas atendidas por los servicios sociales de responsabilidad pública, independientemente de que sean prestados por las administraciones directamente o por entidades privadas.

Por último, cabe señalar algunas limitaciones del estudio llevado a cabo y, concretamente, del análisis multinivel en cuyos resultados se centra el presente artículo. En primer lugar, ha de tenerse en cuenta que, en el análisis realizado, se consideran “contagiadas” las personas usuarias que tienen al menos un resultado positivo en una prueba diagnóstica para SARS-CoV-2 registrada en la base de datos de Osakidetza. La veracidad de esta equiparación depende, lógicamente, de que la mayoría de los contagios de personas usuarias se hayan podido verificar por test, y de que los resultados hayan sido transmitidos a Osakidetza. Si, por ejemplo, los centros privados hubieran recurrido con frecuencia a medios diagnósticos propios, sin transmitir la información sobre los positivos a Osakidetza, el número de positivos confirmados en nuestros datos estarían subestimando la población realmente infectada. Si bien no es posible conocer en qué medida los resultados pueden estar sesgados por la realización de pruebas diagnósticas fuera de la red pública, la opinión de las personas expertas que participaron en el estudio y los datos de

prevalencia de positivos que se obtienen de los datos analizados hacen pensar que el número de casos infradetectados no debe ser muy importante.

En segundo lugar, tal y como se ha indicado, la atribución de usuarios/as positivos a los centros se ha realizado para el primer centro de residencia de cada persona, con lo que dicha atribución resulta incierta para las personas que residieron en más de un centro en el periodo analizado. Esta limitación, que afecta como máximo al 5,5 % de los casos analizados, es la principal fuente de incertidumbre en relación con los resultados obtenidos en este estudio. Futuros análisis a partir de los datos de Osakidetza deberían tratar de incorporar información sobre las fechas de traslados de cada persona, lo que permitiría, al ser cotejadas con las fechas de los test positivos, una atribución más precisa de los contagios a los centros.

Finalmente, a pesar de que las distintas administraciones públicas de Euskadi disponen de cantidad de información sobre los centros

residenciales que hubiera permitido un conocimiento más preciso de los factores involucrados en el riesgo de contagio, a lo largo del estudio se encontraron dificultades para acceder a esos datos, relacionadas en parte con reservas, por parte de las administraciones, respecto a posibles conflictos con el deber de proteger dicha información. No obstante, la investigación realizada ha puesto de manifiesto que los análisis basados en fuentes secundarias —y concretamente los realizados con registros administrativos procedentes de información recogida de forma rutinaria en los procesos de atención— resultan valiosos para orientar la práctica en el ámbito de los servicios sociales y que, sin ir más lejos, es posible obtener datos válidos para analizar a la población que vive en centros residenciales a partir de datos del Sistema de Salud. A partir de ahora, habría que tratar de analizar cuáles son los impedimentos —técnicos, normativos o de otro tipo— existentes para obtener y cruzar bases de datos que están en manos de diferentes administraciones con fines estadísticos y de investigación, y tratar de resolverlos.

Referencias bibliográficas

- ABRAMS, H.R.; LOOMER, L.; GANDHI, A. y GRABOWSKI, D.C. (2020): "Characteristics of US nursing homes with COVID-19 cases", *Journal of the American Geriatrics Society*, vol. 68, nº 8, págs. 1.653-1.656.
- ANDREU, I. y ABELA, J. (2011): "El análisis multinivel una revisión actualizada en el ámbito sociológico", *Metodología de Encuestas*, vol. 13, nº 1, págs. 161-176, <<http://casus.usal.es/pkp/index.php/MdE/article/viewFile/1017/958>>.
- BROWNE, W.J. (2019): *MCMC Estimation in MLwiN (Version 3.03)*, University of Bristol, Centre for Multilevel Modelling, <<https://www.bristol.ac.uk/cmm/media/software/mlwin/downloads/manuals/3-03/mcmc-web.pdf>>.
- BURTON, J.K.; BAYNE, G.; EVANS, C.; GARBE, F.; GORMAN, D.; HONHOLD, N.; MCCORMICK, D.; OTHIENO, R.; STEVENSON, J.; SWIETLIK, S. *et al.* (2020): "Evolution and effects of COVID-19 outbreaks in care homes: A population analysis in 189 care homes in one geographical region of the UK", *The Lancet Healthy Longevity*, vol. 1, nº 1, págs. e21-e31, <[https://doi.org/10.1016/S2666-7568\(20\)30012-X](https://doi.org/10.1016/S2666-7568(20)30012-X)>.
- CAUSA, R.; NIEVAS, D.A. y TAMAYO, C.B. (2021): "COVID-19 y dependencia funcional: análisis de un brote en un centro sociosanitario de personas mayores", *Revista Española de Salud Pública*, vol. 95, e202103045.
- CHEN, M.K.; CHEVALIER, J.A. y LONG, E.F. (2020): "Nursing home staff networks and COVID-19", *Proceedings of the National Academy of Sciences*, vol. 118, nº 1, <<https://doi.org/10.1073/pnas.2015455118>>.
- COMAS-HERRERA, A.; ZALAKÁIN, J.; LEMMON, E.; HENDERSON, D.; LITWIN, C.; HSU, A.T.; SCHMIDT, A.E.; ARLING, G.; KRUSE, F. y FERNÁNDEZ, J.L. (2021): *Mortality Associated with COVID-19 in Care Homes: International Evidence. Last Updated 1st February 2021*, International Long-Term Care Policy Network, <https://ltccovid.org/wp-content/uploads/2021/02/LTC_COVID_19_international_report_January-1-February-1-2.pdf>.
- COMAS-HERRERA, A.; ZALAKÁIN, J.; LEMMON, E.; LITWIN, C.; HSU, A.T.; SCHMIDT, A.E.; ARLING, G.; KRUSE, F. y FERNÁNDEZ, J.L. (2020): *Mortality Associated with COVID-19 in Care Homes: International Evidence. Last Updated 14 October 2020*, International Long-Term Care Policy Network, <<https://ltccovid.org/wp-content/uploads/2021/02/Mortality-associated-with-COVID-among-people-living-in-care-homes-14-October-2020.pdf>>.
- DIEZ ROUX, A.V. (2008): "La necesidad de un enfoque multinivel en epidemiología", *Región y Sociedad*, vol. 20, nº especial 2, págs. 77-91, <<https://doi.org/10.22198/rys.2008.2.a528>>.
- DUNCAN, C.; JONES, K. y MOON, G. (1998): "Context, composition and heterogeneity: Using multilevel models in health research", *Social Science & Medicine*, vol. 46, nº 1, págs. 97-117, <[https://doi.org/10.1016/S0277-9536\(97\)00148-2](https://doi.org/10.1016/S0277-9536(97)00148-2)>.
- EMMERSON, C.; ADAMSON, J.P.; TURNER, D.; GRAVENOR, M.B.; SALMON, J.; COTTRELL, S.; MIDDLETON, V.; THOMAS, B.; MASON, B. y WILLIAMS, C.J. (2020): "Risk factors for outbreaks of COVID-19 in care homes following hospital discharge: A national cohort analysis", *Influenza and Other Respiratory Viruses*, vol. 15, nº 3, págs. 371-380, <<https://doi.org/10.1111/irv.12831>>.

- FRESNO, J.M.; RUIZ VILLAFRANCA, R.; ÁLVAREZ PUERTA, F. y HENAR, L. (2020): *Auditoría de los centros residenciales de Navarra ante la crisis de la COVID-19*, Gobierno de Navarra, Observatorio de la Realidad Social.
- GOLDSTEIN, H.; BROWNE, W. y RASBASH, J. (2002): "Partitioning variation in multilevel models", *Understanding Statistics*, vol. 1, nº 4, págs. 223-231, <https://doi.org/10.1207/S15328031US0104_02>.
- JUNTA DE CASTILLA Y LEÓN (2020): *El impacto del COVID19 en las residencias de personas mayores de Castilla y León y medidas adoptadas*, Servicios Sociales de Castilla y León, <https://serviciossociales.jcyl.es/web/jcyl/binarios/941/542/DEF%20INFORME%20IMPACTO%20COVID19%20%20RESIDENCIAS%20CYL%20Firmado.pdf>.
- LARSEN, K. y MERLO, J. (2005): "Appropriate assessment of neighborhood effects on individual health: Integrating random and fixed effects in multilevel logistic regression", *American Journal of Epidemiology*, vol. 161, nº 1, págs. 81-88, <https://doi.org/10.1093/aje/kwio17>.
- LARSEN, K.; PETERSEN, J.H.; BUDTZ-JØRGENSEN, E. y ENDAHL, L. (2000): "Interpreting parameters in the logistic regression model with random effects", *Biometrics*, vol. 56, nº 3, págs. 909-914, <https://doi.org/10.1111/j.0006-341x.2000.00909.x>.
- LECKIE, G. y GOLDSTEIN, H. (2011a): "A note on 'The limitations of school league tables to inform school choice'", *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 174, nº 3, págs. 833-836, <http://www.jstor.org/stable/23013524>.
- (2011b): "Understanding uncertainty in school league tables", *Fiscal Studies*, vol. 32, nº 2, págs. 207-224, <http://www.jstor.org/stable/24440202>.
- LI, Y.; TEMKIN-GREENER, H.; SHAN, G. y CAI, X. (2020): "COVID-19 infections and deaths among Connecticut nursing home residents: Facility correlates", *Journal of the American Geriatrics Society*, vol. 68, nº 9, págs. 1.899-1.906, <https://doi.org/10.1111/jgs.16689>.
- LECKIE, G. y GOLDSTEIN, H. (2009): "The limitations of using school league tables to inform school choice", *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 172, nº 4, págs. 835-851, <http://www.jstor.org/stable/20622553>.
- LUNN, D.; JACKSON, C.; BEST, N.; THOMAS, A. y SPIEGELHALTER, D. (2012): *The BUGS Book. A Practical Introduction to Bayesian Analysis*, Nueva York, Chapman and Hall/CRC, <https://doi.org/10.1201/b13613>.
- MAS ROMERO, M.; AVENDAÑO CÉSPEDAS, A.; TABERNERO SAHUQUILLO, M.T.; CORTÉS ZAMORA, E.B.; GÓMEZ BALLESTEROS, C.; SÁNCHEZ-FLOR ALFARO, V.; LÓPEZ BRU, R.; LÓPEZ UTIEL, M.; CELAYA CIFUENTES, S.; PEÑA LONGOBARDO, L.M. et al. (2020): "COVID-19 outbreak in long-term care facilities from Spain. Many lessons to learn", *PLoS One*, vol. 15, nº 10, e0241030.
- MCMICHAEL, T.M.; CURRIE, D.W.; CLARK, S.; POGOSJANS, S.; KAY, M.; SCHWARTZ, N.G.; LEWIS, J.; BAER, A.; KAWAKAMI, V.; LUKOFF, M.D. et al. (2020): "Epidemiology of Covid-19 in a long-term care facility in King County, Washington", *New England Journal of Medicine*, vol. 382, nº 21, págs. 2.005-2.011, <https://doi.org/10.1056/NEJMoa2005412>.
- MÉDICOS SIN FRONTERAS (2020): *Too Little, Too Late: The Unacceptable Neglect of the Elderly in Care Homes During the COVID-19 Epidemic in Spain*, Médicos Sin Fronteras, <https://www.msf.es/file/164581/download?token=WWiimuB->.
- MEHTA, H.B.; LI, S. y GOODWIN, J.S. (2021): "Risk factors associated with SARS-CoV-2 infections, hospitalization, and mortality among US nursing home residents", *JAMA Network Open*, vol. 4, nº 3, e216315-e216315.
- MERLO, J. (2003): "Multilevel analytical approaches in social epidemiology: Measures of health variation compared with traditional measures of association", *Journal of Epidemiology and Community Health*, vol. 57, nº 8, págs. 550-552, <https://doi.org/10.1136/jech.57.8.550>.
- (2014): "Invited commentary: Multilevel analysis of individual heterogeneity—a fundamental critique of the current probabilistic risk factor epidemiology", *American Journal of Epidemiology*, vol. 180, nº 2, págs. 208-214, <https://doi.org/10.1093/aje/kwu108>.
- MERLO, J.; CHAIX, B.; OHLSSON, H.; BECKMAN, A.; JOHNNELL, K.; HJERPE, P.; RÅSTAM, L. y LARSEN, K. (2006): "A brief conceptual tutorial of multilevel analysis in social epidemiology: Using measures of clustering in multilevel logistic regression to investigate contextual phenomena", *Journal of Epidemiology and Community Health*, vol. 60, nº 4, págs. 290-297, <https://doi.org/10.1136/jech.2004.029454>.
- MERLO, J.; CHAIX, B.; YANG, M.; LYNCH, J. y RÅSTAM, L. (2005): "A brief conceptual tutorial of multilevel analysis in social epidemiology: Linking the statistical concept of clustering to the idea of contextual phenomenon", *Journal of Epidemiology and Community Health*, vol. 59, nº 6, págs. 443-449, <https://doi.org/10.1136/jech.2004.023473>.
- MERLO, J.; OHLSSON, H.; LYNCH, K.F.; CHAIX, B. y SUBRAMANIAN, S.V. (2009): "Individual and collective bodies: Using measures of variance and association in contextual epidemiology", *Journal of Epidemiology and Community Health*, England, vol. 63, nº 12, págs. 1.043-1.048, <https://doi.org/10.1136/jech.2009.088310>.
- MERLO, J.; VICIANA-FERNÁNDEZ, F.J. y RAMIRO-FARIÑAS, D. (2012): "Bringing the individual back to small-area variation studies: A multilevel analysis of all-cause mortality in Andalusia, Spain", *Social Science & Medicine (1982)*, vol. 75, nº 8, págs. 1.477-1.487, <https://doi.org/10.1016/j.socscimed.2012.06.004>.
- MERLO, J.; WAGNER, P.; AUSTIN, P.C.; SUBRAMANIAN, S.V. y LECKIE, G. (2018): "General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: A conceptual tutorial", *SSM - Population Health*, vol. 5, págs. 33-37, <https://doi.org/10.1016/j.ssmph.2018.05.006>.

- MERLO, J.; WAGNER, P.; GHITH, N. y LECKIE, G. (2016): "An original stepwise multilevel logistic regression analysis of discriminatory accuracy: The case of neighbourhoods and health", *PLoS One*, vol. 11, nº 4, e0153778, <<https://doi.org/10.1371/journal.pone.0153778>>.
- PANAGIOTOU, O.A.; KOSAR, C.M.; WHITE, E.M.; BANTIS, L.E.; YANG, X.; SANTOSTEFANO, C.M.; FEIFER, R.A.; BLACKMAN, C.; RUDOLPH, J.L.; GRAVENSTEIN, S. *et al.* (2021): "Risk factors associated with all-cause 30-day mortality in nursing home residents with COVID-19", *JAMA Internal Medicine*, vol. 181, nº 4, págs. 439-448, <<https://doi.org/10.1001/jamainternmed.2020.7968>>.
- PEPE, M.S.; JANES, H.; LONGTON, G.; LEISENRING, W. y NEWCOMB, P. (2004): "Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker", *American Journal of Epidemiology*, vol. 159, nº 9, págs. 882-90.
- ROSE, G. (2001): "Sick individuals and sick populations", *International Journal of Epidemiology*, vol. 30, nº 3, págs. 427-432, <<https://doi.org/10.1093/ije/30.3.427>>.
- SIIS (2021): *El impacto de la pandemia del COVID-19 en los centros residenciales para personas mayores en Euskadi*, Vitoria-Gasteiz, Ararteko, <<https://www.siiis.net/es/documentacion/catalogo/Record/568683>>.
- SPIEGELHALTER, D.J.; BEST, N.G.; CARLIN, B.P. y VAN DER LINDE, A. (2002): "Bayesian measures of model complexity and fit", *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 64, n 4, págs. 583-639, <<https://doi.org/10.1111/1467-9868.00353>>.
- STALL, N.M.; JONES, A.; BROWN, K.A.; ROCHON, P.A. y COSTA, A.P. (2020): "For-profit long-term care homes and the risk of COVID-19 outbreaks and resident deaths", *Canadian Medical Association Journal*, vol. 192, nº 33, págs. E946-E955, <<https://doi.org/10.1503/cmaj.201197>>.
- SUÑER, C.; OUCHI, D.; MAS, M.À.; ALARCON, R.L.; MESQUIDA, M.M.; PRAT, N.; BONET-SIMÓ, J.M.; IZQUIERDO, M.E.; SÁNCHEZ, I.G.; NOGUEROLA, S.R. *et al.* (2021): "A retrospective cohort study of risk factors for mortality among nursing homes exposed to COVID-19 in Spain", *Nature Aging*, vol. 1, nº 7, págs. 579-584, <<https://doi.org/10.1038/s43587-021-00079-7>>.
- TINSLEY, B. (2020): *Impact of Coronavirus in Care Homes in England: 26 May to 19 June 2020*, Office for National Statistics, <<https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/impactofcoronavirusincarehomesinenglandivaldi/latest>>.