



Editorial

El Conjunto Mínimo Básico de Datos (CMBD), nuestro *big data* para para la investigación epidemiológica de la patología respiratoria



The Minimum Basic Data Set (MBDS), Our Big Data for the Epidemiological Investigation of Respiratory Diseases

 Javier de Miguel Díez^{a,*}, Ana López de Andrés^b y Rodrigo Jiménez García^b
^a Servicio de Neumología, Hospital General Universitario Gregorio Marañón, Facultad de Medicina, Universidad Complutense de Madrid (UCM), Instituto de Investigación Sanitaria Gregorio Marañón (IISGM), Madrid, España

^b Departamento de Medicina Preventiva y Salud Pública, Facultad de Ciencias de la Salud, Universidad Rey Juan Carlos, Alcorcón, Madrid, España

Las enfermedades respiratorias son una causa importante de discapacidad y muerte, imponiendo una elevada carga sanitaria¹. Además, su incremento es constante, a lo que contribuyen factores como el envejecimiento de la población, la exposición al humo del tabaco, las condiciones climatológicas y ambientales y el aumento de la prevalencia de obesidad². En este contexto, la posibilidad de disponer de información objetiva sobre la carga asistencial de estas enfermedades se ha convertido en una prioridad de nuestro Sistema Nacional de Salud (SNS), siendo los ingresos hospitalarios una de las principales fuentes de las que disponemos para ello.

El Consejo Interterritorial del SNS aprobó en España, en el año 1987, el Conjunto Mínimo de Datos al alta hospitalaria (CMBD). Es una base de datos administrativa de obligado cumplimiento para los hospitales de nuestro sistema sanitario. Además, desde el año 2005 se ha ampliado al sector privado, por lo que el CMBD contiene en la actualidad más del 90% de las altas registradas en los hospitales de agudos de España^{3,4}. Se trata, por tanto, de una fuente muy útil para la investigación epidemiológica de las enfermedades respiratorias, siendo además un recurso de elevado valor organizativo para la planificación y evaluación de los servicios sanitarios⁵.

La base de datos del CMBD es completada por personal especialmente formado en codificación, perteneciente a los servicios de admisión de los hospitales españoles. Estos profesionales reciben cursos de actualización periódicos por parte de las comunidades autónomas para garantizar su correcta formación⁶. Los codificadores utilizan el informe de alta del médico así como información de la base de datos del hospital para rellenar los campos de la base de datos del CMBD. De esta forma recogen datos como la edad, el sexo, las comorbilidades, el diagnóstico principal y los secundarios, los procedimientos efectuados (tanto diagnósticos como terapéuticos), las complicaciones, la estancia media, la mortali-

dad intrahospitalaria, el destino al alta y el reingreso en menos de 30 días. Es importante tener presente que cada episodio del CMBD es asignado a un grupo relacionado con el diagnóstico (GRD), un sistema que clasifica a los pacientes hospitalarios en grupos homogéneos en cuanto al consumo de recursos⁷. Además, para garantizar una correcta codificación, las comunidades autónomas realizan auditorías periódicas⁶. La información es enviada anualmente al Ministerio de Sanidad y Consumo, que se encarga de la gestión de la base de datos⁷. Algunos autores han estudiado la validez del CMBD para diferentes patologías demostrando su elevada sensibilidad y especificidad⁸⁻¹⁰.

El CMBD es, por lo tanto, la base de datos clínico-administrativa más amplia de la que disponemos en nuestro SNS. Debido a su elevado tamaño muestral, su análisis podría entrar dentro de la investigación conocida como *big data*. Se trata por tanto de una herramienta de alto valor para evaluar la epidemiología de las enfermedades respiratorias (morbilidad), la gestión de los recursos sanitarios (mortalidad intrahospitalaria, estancia media, reingresos, traslado a centros sociosanitarios) y la seguridad del paciente (reacciones adversas, complicaciones asociadas a procedimientos diagnósticos y terapéuticos). Por otra parte, dado que los datos se recogen anualmente, a través de su análisis podemos estimar la evolución a lo largo del tiempo de variables como la edad, la comorbilidad y las enfermedades que se han atendido durante un periodo determinado⁷.

Recientemente se han publicado los resultados del proyecto RECALAR (Recursos y Calidad en Patología del Aparato Respiratorio)¹¹. A través del análisis del CMBD, los autores han puesto de manifiesto que las enfermedades del aparato respiratorio representaron en el año 2015 el segundo diagnóstico más frecuente, alcanzando el 12,1% del total de altas. La duración media de la hospitalización por procesos respiratorios fue de 8,3 días. Con relación a la mortalidad, la patología respiratoria jugó también un papel importante, llegando a ser la causa responsable del 19,6% del total de fallecidos en ese año.

* Autor para correspondencia.

 Correo electrónico: javier.miguel@salud.madrid.org (J. de Miguel Díez).

En relación con los cambios epidemiológicos a lo largo del tiempo, el análisis del CMBD ha evidenciado que, en los últimos años, se ha producido una disminución de la incidencia de ingresos hospitalarios de algunas enfermedades como la EPOC¹² y el asma¹³. Por el contrario, se ha constatado un aumento de la incidencia de hospitalizaciones por otros procesos como las infecciones respiratorias¹⁴, la fibrosis pulmonar¹⁵ y el tromboembolismo pulmonar¹⁶. A pesar de ello, se ha objetivado una reducción concomitante de la estancia media¹⁶ y de la mortalidad intrahospitalaria de algunos de estos procesos^{12,14-16}, lo que sugiere que su manejo ha ido mejorando en España con el paso del tiempo. En relación con las neoplasias, se ha observado una disminución en la tasa de ingresos por cáncer de pulmón en los hombres y un aumento en las mujeres¹⁷.

El uso del CMBD para la investigación clínica tiene diversas ventajas. Es rápido y barato, en comparación con el hecho de tener que recabar datos primarios, y también fácil de conseguir¹⁸. Además, su tamaño muestral es elevado, lo que le otorga una gran potencia estadística a la hora de analizar algunas variables como la mortalidad intrahospitalaria. También permite realizar análisis de subgrupos por sexo, edad, comorbilidades, o estudiar enfermedades o procedimientos de muy baja prevalencia. Por otra parte, su metodología se ha mantenido estable a lo largo de muchos años y tiene una elevada validez externa, dado que cubre más del 95% de los ingresos en hospitales públicos y privados⁷. Sin embargo, su calidad depende de la precisión con la que se codifican los diagnósticos y los procedimientos, algo que se realiza generalmente a partir de los informes de alta¹⁸. Otra limitación de esta base de datos es la falta de información clínica relevante, como la gravedad de la enfermedad, su duración, los resultados de las pruebas de laboratorio o el tratamiento realizado. Tampoco permite el seguimiento de los pacientes tras el alta. Por último, el CMBD codifica ingresos y puede ocurrir que un mismo paciente puede estar «repetido» en la base si ha presentado varias hospitalizaciones durante el mismo año⁷.

A pesar de estas limitaciones, el uso de bases de datos administrativas para evaluar los resultados de los servicios de salud ha sido validado en comparación con los registros clínicos¹⁹. Debido a ello, el CMBD podría considerarse como una herramienta muy valiosa para la investigación epidemiológica de las enfermedades respiratorias, cuyo interés social y sanitario es más que relevante. En cualquier caso, no hay que olvidar que una gran parte del conocimiento clínico está aún oculto, sumergido en las historias clínicas de los pacientes y en la información no normalizada recogida por los profesionales sanitarios en un lenguaje natural, algo a lo que el *big data* tendrá que dar respuesta en un futuro próximo⁵.

Bibliografía

1. Foro de las Sociedades Respiratorias Internacionales. El impacto global de la enfermedad respiratoria – Segunda edición. México: Asociación Latinoamericana de Tórax; 2017.
2. Federación Española de Empresas de Tecnología Sanitaria (FENIN). Estudio sobre la eficiencia y los beneficios de las terapias respiratorias domiciliarias. Madrid: Comuniland SL; 2011 [consultado 31 Ene 2019]. Disponible en: http://panelfenin.es/uploads/plataforma/documents/24.documentacion.FENIN_TerapiasResp.ok.pdf.
3. Secretaría General para el Sistema Nacional de Salud. Resolución 1/92 sobre el establecimiento de un CMBD al alta hospitalaria. Madrid: Ministerio de Sanidad y Consumo; 1992.
4. Informe de hospitalización - CMBD - Registro de altas. Informe resumen 2013 [Internet]. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad; 2015 [consultado 31 Ene 2019]. Disponible en: www.msssi.gob.es/estadEstudios/estadisticas/cmbdhome.htm.
5. Marco Cuenca G, Salvador Oliván JA. Del CMBD al Big Data en salud: un sistema de información hospitalaria para el siglo XXI. *Scire*. 2018;24:77–89.
6. Manual de auditoría. Conjunto mínimo básico de datos. Junta de Andalucía [consultado 31 Ene 2019]. Disponible en: http://www.juntadeandalucia.es/servicioandaluzdesalud/contenidos/publicaciones/datos/348/pdf/ManualauditoriaCMBD_hospSSPA.pdf.
7. Méndez Bailón M. El conjunto mínimo básico de datos (CMBD) como fuente para la investigación en Medicina Interna [consultado 31 Ene 2019]. Disponible en: <https://medicinainternaaltovalor.fesemi.org/tag/dr-mendez-bailon>.
8. Márquez Cid M, Valera Niñirola I, Chirilaque López MD, Tortosa Martínez J, Párraga Sánchez E, Navarro Sánchez C. Validación de los códigos diagnósticos de cáncer de colon y recto del conjunto mínimo básico de datos. *Gac Sanit*. 2006;20:266–72.
9. Bernal-Delgado EE, Martos C, Martínez N, Chirilaque MD, Márquez M, Navarro C, et al. Is hospital discharge administrative data an appropriate source of information for cancer registries purposes? Some insights from four Spanish registries. *BMC Health Serv Res*. 2010;10:9.
10. Ramalle-Gomara E, Ruiz E, Serrano M, Bartulos M, Gonzalez MA, Matute B. Validity of discharge diagnoses in the surveillance of stroke. *Neuroepidemiology*. 2013;41:185–8.
11. Alfageme I, Fernández Villar A, Soriano JB. Las enfermedades respiratorias en España a la luz del CMBD de RECALAR. *Monogr Arch Bronconeumol*. 2018;5, 00153.
12. De Miguel-Díez J, Jiménez-García R, Hernández-Barrera V, Puente-Maestu L, Rodríguez-Rodríguez P, López de Andrés A, et al. Trends in hospital admissions for acute exacerbation of COPD in Spain from 2006 to 2010. *Respir Med*. 2013;107:717–23.
13. De Miguel-Díez J, Jiménez-García R, Hernández-Barrera V, López de Andrés A, Villa-Asensi JR, Plaza V, et al. National trends in hospital admissions for asthma exacerbations among pediatric and young adult population in Spain (2002–2010). *Respir Med*. 2014;108:983–91.
14. De Miguel-Díez J, Jiménez-García R, Hernández-Barrera V, Jiménez-Trujillo I, de Miguel-Yanes JM, Méndez-Bailón M, et al. Trends in hospitalizations for community-acquired pneumonia in Spain: 2004 to 2013. *Eur J Intern Med*. 2017;40:64–71.
15. Pedraza-Serrano F, López de Andrés A, Jiménez-García R, Jiménez-Trujillo I, Hernández-Barrera V, Sánchez-Muñoz G, et al. Retrospective observational study of trends in hospital admissions for idiopathic pulmonary fibrosis in Spain (2004–2013) using administrative data. *BMJ Open*. 2017;7:e013156.
16. De Miguel-Díez J, Jiménez-García R, Jiménez D, Monreal M, Guijarro R, Otero R, et al. Trends in hospital admissions for pulmonary embolism in Spain from 2002 to 2011. *Eur Respir J*. 2014;44:942–50.
17. Palacio Nebreda MM, de Miguel-Díez J, Villegas Fernández FR, Segura Frago A, Rodríguez Calderón JL, Martínez Hernández D. Tendencias en la incidencia de hospitalizaciones por cáncer de pulmón en España entre 2001 y 2011. *Arch Bronconeumol*. 2016;52:411–9.
18. Bernal JL, Barrabés JA, Íñiguez A, Fernández-Ortiz A, Fernández-Pérez C, Bardají A, et al. Datos clínicos y administrativos en la investigación de resultados del síndrome coronario agudo en España. Validez del Conjunto Mínimo Básico de Datos. *Rev Esp Cardiol (Engl Ed)*. 2019;72:56–62.
19. Horwitz LI, Lin Z, Herrin J, Bernheim S, Drye EE, Krumholz HM, et al. Association of hospital volume with readmission rates: a retrospective cross-sectional study. *BMJ*. 2015;350:h447.