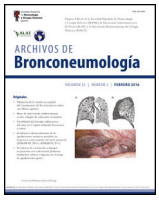
Arch Bronconeumol. 2020;56(5):267–268

Sociedad Española de Neumología y Cirugía Torácica SEPAR

ARCHIVOS DE Bronconeumología

www.archbronconeumol.org



Editorial

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

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

Javier de Miguel Díez,[a,*] Ana López de Andrés,[b] 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, Spain*

[b] *Departamento de Medicina Preventiva y Salud Pública, Faculta de Ciencias de la Salud, Universidad Rey Juan Carlos, Alcorcón, Madrid, Spain*

Respiratory diseases are a major cause of disability and death, and are associated with a significant healthcare burden.[1] Moreover, the prevalence of these diseases continues to increase, spurred on by factors such as aging of the population, exposure to tobacco smoke, climatic and environmental conditions, and increasing obesity.[2] In this context, the collection of objective data on the burden of care imposed by these diseases has become a priority for our National Health System (SNS, in its Spanish acronym), and one of the best sources of this information is hospital admissions.

In 1987, the Spanish Interterritorial SNS Council approved the implementation of the Minimum Basic Data Set (MBDS), an administrative database that hospitals in our health system are mandatorily required to complete. Since 2005, it also has been extended to the private sector, so the MBDS now contains more than 90% of discharges from acute care hospitals in Spain.[3,4] It is therefore a very useful source for epidemiological research into respiratory diseases, and a resource with high organizational value for the planning and evaluation of health services.[5]

The MBDS database is completed by specially trained coding staff working in the admissions departments of Spanish hospitals. These professionals receive periodic refresher courses sponsored by their autonomous communities to ensure that they are properly trained.[6] The coders use the medical discharge report and information from the hospital database to populate the MBDS database fields. In this way, they collect data such as age, sex, comorbidities, principal and secondary diagnoses, procedures conducted (both diagnostic and therapeutic), complications, mean length of stay, in-hospital mortality, destination at discharge, and re-admission within 30 days. It is important to note that each MBDS episode is assigned to a Diagnostic Related Group (DRG), a system that classifies hospital patients in homogeneous groups in terms of resource consumption.[7] In order to ensure the quality of the coding, the autonomous communities also perform periodic audits.[6] The data are sent annually to the Ministry of Health and Consumer Affairs, which is responsible for the management of the database.[7] Some authors have studied the validity of the MBDS for different diseases and have corroborated its high sensitivity and specificity.[8–10]

The MBDS is, therefore, the largest administrative and clinical database available in the Spanish health system. Given its large sample size, analysis of this data could be considered as research known as big data. It is therefore a valuable tool for evaluating the epidemiology of respiratory diseases (morbidity), the management of health resources (in-hospital mortality, mean length of stay, readmissions, transfer to other social welfare centers) and patient safety (adverse reactions, complications associated with diagnostic and therapeutic procedures). Because the information is collected on a yearly basis, it can be analyzed to estimate the course over time of variables such as age, comorbidity, and diseases that have been treated over a specified time period.[7]

The results of the RECALAR project (Resources and Quality in Pathology of the Respiratory System) were recently published.[11] By analyzing the MBDS, the authors showed that in 2015 respiratory diseases represented the second most frequent diagnosis, accounting for 12.1% of total discharges. Mean length of stay for respiratory diseases was 8.3 days. Respiratory diseases also played an important role in terms of mortality by causing 19.6% of the total deaths in that year.

In relation to epidemiological changes over time, analysis of the MBDS has revealed a decrease in recent years in the incidence of hospital admissions for some diseases, such as COPD[12] and asthma.[13] In contrast, there has been an increase in the incidence of hospitalizations for other conditions, such as respiratory infections,[14] pulmonary fibrosis,[15] and pulmonary thromboembolism.[16] This increase, however, has occurred in parallel with a reduction in the mean length of stay[16] and in-hospital mortality associated with some of these processes[12,14–16], suggesting that their management in Spain has improved over time. As

☆ Please cite this article as: de Miguel Díez J, López de Andrés A, Jiménez García R. El Conjunto Mínimo Básico de Datos (CMBD), nuestro *big data* para para la investigación epidemiológica de la patología respiratoria. Arch Bronconeumol. 2020;56:267–268.

* Corresponding author.
E-mail address: javier.miguel@salud.madrid.org (J. de Miguel Díez).

regards malignancies, the rate of lung cancer has fallen among men and risen among women.[17]

The use of the MBDS for clinical research has several advantages. It is faster and cheaper than collecting primary data, and easy to obtain.[18] In addition, the sample size is high, conferring a great statistical power to the analysis of some variables, such as in-hospital mortality. It also useful for performing subgroup analyses by sex, age, comorbidities, and for studying diseases or procedures with very low prevalence. The database methodology has also remained stable over many years and its external validity is high, since it includes more than 95% of admissions to public and private hospitals.[7] Nevertheless, its quality depends on accuracy of diagnosis and procedure coding, a process usually conducted on the basis of discharge reports.[18] Another limitation of this database is the lack of significant clinical information, such as disease severity and duration, laboratory results, and treatments administered. Nor does it follow up patients after discharge. Finally, the MBDS encodes admissions, so the same patient can be "repeated" in the database if they were admitted several times during the same year.[7]

Despite these limitations, the use of administrative databases to evaluate health service outcomes has been validated in comparison with clinical records.[19] The MBDS may therefore be seen as a valuable tool for epidemiological research into respiratory diseases, with their inherent social and healthcare importance. In any case, we must not forget that a large part of the clinical knowledge is still hidden, submerged in the clinical records of patients and non-standardized information gathered by health professionals in a natural language, an issue that big data will have to address in the near future.[5]

References

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. Available at: http://panelfenin.es/uploads/plataforma/documents/24_documentacion_FENIN_TerapiasResp_ok.pdf [accessed 31.01.19].
3. Secretaria 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 – C.M.B.D. – Registro de altas. Informe resumen 2013 [Internet]. Madrid: Ministerio de Sanidad, Servicios Sociales e Igualdad; 2015. Available at: www.msssi.gob.es/estadEstudios/estadisticas/cmbdhome.htm [accessed 31.01.19].
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. Available at: http://www.juntadeandalucia.es/servicioandaluzdesalud/contenidos/publicaciones/datos/348/pdf/ManualauditoCMBD_hospSSPA.pdf [accessed 31.01.19].
7. Méndez Bailón M. El conjunto mínimo básico de datos (CMBD) como fuente para la investigación en Medicina Interna. Available at: https://medicinainternaaltovalor.fesemi.org/tag/dr-mendez-bailon [accessed 31.01.19].
8. Márquez Cid M, Valera Niñirola I, Chirlaque 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, Chirlaque 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-Diez J, Villegas Fernández FR, Segura Fragoso 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.