
Del CMBD al Big Data en salud: un sistema de información hospitalaria para el siglo XXI

From MBDS to big data: a hospital information system for the 21st century

Gonzalo MARCO CUENCA, José Antonio SALVADOR OLIVÁN

Departamento de Ciencias de la Documentación e Historia de la Ciencia,
Universidad de Zaragoza. gmarco@unizar.es, jaso@unizar.es

Resumen

Se analiza la evolución del Conjunto Mínimo Básico de Datos (CMBD), un sistema de información hospitalario desde sus inicios hasta su última versión denominada Registro de Atención Sanitaria Especializada (RAE-CMBD). Su objetivo es normalizar toda la actividad a nivel especializado en hospitalización y atención ambulatoria del sector público y privado; y permite conocer el funcionamiento de los hospitales, su actividad asistencial, las patologías atendidas, su complejidad y abordaje clínico y los costes que suponen. Los registros clínico-administrativos, como es el caso del RAE-CMBD, se presentan como una gran oportunidad para la explotación de datos masivos sanitarios o *big data*, con el objetivo de poder conocer de forma anticipada la incidencia de enfermedades, el manejo de las mismas y los diferentes resultados asistenciales, tanto a nivel de conjunto del SNS, como de forma individual o comparada por hospitales o centros sanitarios. Si se cumplen las previsiones para los próximos años, el manejo de datos a gran escala podría mejorar la eficacia en la toma de decisiones para la planificación, evaluación y desarrollo de los sistemas sanitarios, convirtiéndose el RAE-CMBD en la fuente de datos del Sistema de Información Sanitaria del SNS (SISNS) de mayor relevancia e impacto para la gestión de los servicios, el conocimiento de la salud de la población y la investigación epidemiológica.

Palabras clave: Conjunto Mínimo Básico de Datos. CMBD. Big data. Sistemas de Información hospitalaria. Registro de Actividad de Atención Sanitaria Especializada. RAE-CMBD. Sistema de Información Sanitaria del SNS. SISNS. Gestión de la información sanitaria. Datos en salud.

1. Introducción

Dentro de los sistemas de información del Sistema Nacional de Salud (SNS) de España hay que destacar como una pieza clave y fundamental al *Sistema de Información Sanitaria del SNS (SISNS)*. Este sistema tiene como principal finalidad facilitar información objetiva, fiable y comparable sobre el funcionamiento del sistema sanitario, permitiendo valorar la consecución de sus objetivos, respondiendo a las necesidades de información de los agentes que lo conforman y sir-

Abstract

The Minimum Basic Data Set (MBDS) is a hospital information system of demographic type that has been developed in the National Health System (NHS) for over more than 25 years. The present study analyzes its evolution from its beginnings to latest version called the Register of Specialized Health Care (RSH-MBDS). This new model standardizing all activity at specialized level, both in hospitalization and ambulatory care in public and private sector, allowing to know about the operation of hospitals, care activity, pathologies treated, complexity and clinical approach and the costs involved. Clinical-administrative records, as is the case of RSH-MBDS, are a great opportunity for the exploitation of massive health data or big data, with the aim of being able to know in advance the incidence of diseases, their management and the different healthcare outcomes, in the entire NHS, as a separately or compared by hospitals or health centers. If the forecasts for the next years are fulfilled, large-scale data management could improve decision-making efficiency in the planning, evaluation and development of health systems, becoming the RSH-MBDS in the most relevant data source for the management of services, knowledge of the population health and epidemiological research.

Keywords: Minimum Basic Data Set. MBDS. Big Data. Hospital Information System. Specialized Health Care Activity Register. RAE-CMBD. Health Information System of the NHS. SISNS. Health Information management. Health Data.

viendo de soporte para la generación de conocimiento (Alfaro Latorre, 2012). El SISNS se caracteriza por ser un sistema que reúne información, principalmente estadística, procedente de otros sistemas sanitarios dentro del contexto del SNS (1). Estos datos que proceden de la gestión clínica y administrativa diaria que se produce en los diferentes centros y niveles asistenciales en España, tanto a nivel público, como privado, constituyen una fuente trascendental para comprender la situación actual del SNS, su contexto y su administración. Las Comunidades Autónomas, la

Administración General del Estado y las Entidades Gestoras de la Seguridad Social tienen la obligación de aportar datos al SISNS para favorecer su desarrollo y su mantenimiento. De forma paralela, las administraciones autonómicas y estatales tienen derecho a acceder y disponer de los datos que forman parte del SISNS. Es, por tanto, un sistema basado en la cooperación de las diferentes administraciones sanitarias, que busca rentabilizar el conocimiento que se genera y los recursos de información existentes en beneficio del conjunto.

El SISNS fue definido como un elemento instrumental capaz de medir la efectividad (aquello que funciona y lo que no), la eficiencia (reconocer si los recursos se están utilizando de manera óptima para obtener los resultados deseados) y dar una respuesta sobre el sistema de protección de la salud, en relación con: el nivel de acceso a las prestaciones, la medida en que el sistema sanitario responde a las necesidades de salud y el grado de satisfacción de la población. El SISNS, por tanto, se ha convertido en un recurso estratégico que busca hacer frente a los retos presentes y futuros de la realidad organizativa y de las demandas de información que plantean los tres grandes actores que integran el SNS: ciudadanos, profesionales y administraciones sanitarias (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2014).

Una de las principales fuentes de datos que integran el SISNS es el *Conjunto Mínimo Básico de Datos (CMBD)*. El CMBD es una fuente tradicional y muy arraigada en el SNS, que cuenta con más de 25 años de vida (la cumplimentación del CMBD es obligatoria desde el año 1992). El CMBD ofrece información estadística a nivel estatal relativa a los casos atendidos en los centros sanitarios de Atención Especializada y, en concreto, a los centros hospitalarios, ambulatorios y centros de especialidades dependientes de los mismos. Se trata de un registro de la actividad de la atención sanitaria especializada que reúne información poblacional y que utiliza como unidad de registro cada contacto asistencial con un paciente (2). Los registros se basan en un núcleo común de datos básicos normalizados y anonimizados que integran variables demográficas y clínico-asistenciales, permitiendo conocer la morbilidad atendida en cada hospital, ya sea público o privado. El CMBD es una fuente esencial para la investigación epidemiológica dentro del SNS y un recurso de alto valor organizacional para la planificación y evaluación de los servicios sanitarios (3).

La información como recurso es cuantiosa en un hospital y, tal y como indica Salvador Oliván (1997) “no existe un campo con una necesidad tan grande de proceso automático de datos como

el de la asistencia sanitaria”. En este sentido, el CMBD se constituye como una herramienta imprescindible para el procesamiento de datos de una forma estandarizada. Según Zapatero Gaviña (2010), el CMBD no debe interpretarse solamente como una base de datos administrativa que es utilizada para la gestión sanitaria; es, además, una herramienta muy útil para que los clínicos diseñen proyectos de investigación y, al ser una fuente de obligado cumplimiento (4), aporta abundante información que incluye registros con millones de entradas cada año (casi cuatro millones de altas anuales). El CMBD puede también considerarse como un sistema de información propiamente dicho, entendiendo el mismo como el proceso estadístico de entrada de datos y de salida de resultados en forma de indicadores que permiten medir la actividad realizada por los servicios y la morbilidad atendida en cada uno de ellos, pudiendo incluso detectar problemas potenciales de calidad e ineficiencia (Salvador Oliván, 1997).

Asimismo, es un sistema especializado que facilita el *benchmarking*, pues incluye diversos parámetros asistenciales como, por ejemplo, la duración de la estancia, los diagnósticos principales y secundarios, la continuidad asistencial, etc., que ofrecen la posibilidad de realizar comparaciones entre los distintos hospitales o entre los diversos servicios hospitalarios. Para poder analizar la información que integra el CMBD y poder ser objeto de comparación, ha sido necesaria la normalización de su estructura y la uniformidad de su contenido Carrillo y otros (1991) destacan “el tema de la uniformidad es un punto clave a la hora de tratar la información sanitaria. Además, la eficiencia, eficacia y equidad en la oferta de asistencia sanitaria no podría ser comparada entre distintos sistemas sanitarios en ausencia de una información uniforme de cada uno de ellos”.

El presente artículo mantiene como objetivo principal analizar la consolidación del CMBD como un sistema de información y un recurso estratégico para el conocimiento de la actividad asistencial especializada, a presente y futuro, en el conjunto del SNS. De igual modo, se busca exponer como su crecimiento, dada la magnitud de datos que anualmente se generan, obligará a un tratamiento basado en técnicas para el análisis de datos masivos o *big data*.

2. Origen, antecedentes y evolución del CMBD

El concepto de conjunto mínimo de datos básicos fue formulado por primera vez en 1969 en el contexto de una conferencia sobre sistemas de infor-

mación sanitarios en los Estados Unidos (Treviño, 1988; Foster y Conrick, 1998). Desde entonces hasta la actualidad el CMBD ha ido evolucionando a nivel internacional (ver tabla I), mejorando sus prestaciones, armonizando su información y normalizando sus datos y variables, tanto a nivel de datos administrativos, como clínicos.

Año	Hito cronológico
1969	Primera formulación del concepto CMBD.
1969	La OMS, a través de un Grupo de Estudio, recomienda un conjunto básico de datos con una finalidad estadística.
1972	El National Committee on Vital and Health Statistics de los Estados Unidos pone en marcha el Uniform Hospital Discharge Data Set (UHDDS), primer CMBD al Alta Hospitalaria en ese país, formado por 14 variables.
1974	El UHDDS es adoptado para los programas de salud estadounidenses: Medicare y Medicaid (Treviño, 1988).
1981	Bajo el patrocinio de la Comisión Europea, se define el European Minimum Basic Data Set (EMBDS), tomando como referencia el UHDDS. El EMBDS contenía un total de 13 variables y su cobertura debía contemplar tanto la atención hospitalaria como la ambulatoria.
2002	Primer Programa europeo de acción comunitaria en el ámbito de la salud (2001-2006). Se establece como área prioritaria el desarrollo y coordinación de un sistema de información sanitaria (Calvete, 2003).
2008	Segundo Programa europeo de acción comunitaria en el ámbito de la salud (2008-2013). Se concreta entre sus objetivos la generación y difusión de información, impulsando un sistema de información con mecanismos para recoger datos e información comparables y con indicadores estadísticos.
2008	Se publica el Reglamento (CE) 1338/2008 del Parlamento Europeo y del Consejo sobre estadísticas comunitarias de salud pública y de salud y seguridad del trabajo (5).
2014	Reglamento 282/2014 que define el Tercer Programa europeo de acción comunitaria en el ámbito de la salud (2014-2020). Como prioridad temática se encuentra la promoción de un sistema de información sobre salud que contribuya al proceso de toma de decisiones basado en datos.

Tabla I. Evolución internacional del CMBD (hitos más destacados)

2.1. Antecedentes y evolución del CMBD en España

En España, a mediados de los años 70, por medio del Real Decreto 1360/1976, se hace obligatorio el uso, por parte de los establecimientos sanitarios con régimen de internado de un libro

único de registro de ingresos y altas. Esta normativa buscaba ampliar el conocimiento, existente en esos momentos, sobre el sector hospitalario, planteando la recogida de aquellos datos estimados como fundamentales para comprender las circunstancias de las personas ingresadas. Este libro de registro se creaba como un instrumento de mejora y de apoyo al conocimiento sobre la morbilidad hospitalaria que ya se venía desarrollando, desde los años 50, por parte del Instituto Nacional de Estadística (INE) (6). Esta norma, actualmente derogada (7), se constituye como el primer antecedente legislativo de registro hospitalario en el contexto del SNS.

En 1986, la *Ley 14/1986, General de Sanidad*, en su artículo 23, dispone que las Administraciones Sanitarias, de acuerdo con sus competencias, deben crear registros y elaborar análisis de información para el conocimiento de las distintas situaciones que puedan derivar acciones de intervención de la autoridad sanitaria (Ministerio de Sanidad y Consumo, 2004). En 1987, vista la necesidad de crear una fuente de datos uniforme para cumplir con la *Ley General de Sanidad* y facilitar los procesos de gestión sanitaria, implantación de nuevos sistemas de financiación, elaboración de indicadores de rendimiento, control de la calidad asistencial y elaboración de una base de datos para investigación clínica y epidemiológica se propone, a través del *Grupo de Trabajo de la Ponencia sobre Planificación Sanitaria* del Ministerio de Sanidad, elevar una propuesta para establecer un CMBD al Alta Hospitalaria o CMBD de Hospitalización (CMBD-H) al Consejo Interterritorial del SNS (CISNS) (8). El 14 de diciembre de 1987, el Pleno del CISNS (1987), aprueba un CMBD nacional compuesto por 14 variables que toman como referencia las ya definidas en el UHDDS de los Estados Unidos y en el EMBDS europeo.

En 1988, por medio de otro Pleno del CISNS (1988) se desarrolla la estrategia de implantación del CMBD y se acuerda la creación del *Comité Técnico del Consejo Interterritorial para la Implantación y Evaluación del CMBD (Comité Técnico del CMBD)* (9) que tendrá como principales funciones: la homologación de criterios y definiciones en los datos del CMBD, la homologación de los criterios de codificación (10), la evaluación de la factibilidad y limitaciones del proceso de implantación del CMBD, el análisis de la información suministrada por el CMBD y la ampliación y/o revisión del CMBD.

A partir de esta decisión del CISNS, las Comunidades Autónomas se encargarán de ir implementando a lo largo del tiempo el CMBD, siendo Cataluña la primera Comunidad Autónoma en po-

nerlo en marcha en 1990. En el año 1992, la Secretaría General del SNS del Ministerio de Sanidad y Consumo determina que todos los centros hospitalarios del Instituto Nacional de la Salud (INSALUD) se encuentran obligados al registro y codificación del CMBD-H y su implantación deberá producirse antes del 31 de diciembre de ese mismo año. Los servicios sanitarios de las Comunidades Autónomas, en virtud de sus competencias en materia de gestión sanitaria, irán publicando normas locales para regular la implantación y recogida de datos para el CMBD-H a diferentes ritmos desde 1992 hasta el año 2011. El CMBD-H hay que considerarlo, consecuentemente, como un repositorio nacional que se abastece de una red de sistemas de información autonómicos. Estos últimos se van a encargar de suministrar la información que se registra en cada hospital de forma anual para conformar un producto único para todo el SNS.

Desde el año 1993, se viene regulando el denominado *Plan Estadístico Nacional (PEN)* (Real Decreto 136/1999) como el principal instrumento ordenador de la actividad estadística de la Administración General del Estado. El Real Decreto 1658/2012, por el que se aprueba el PEN 2013-2016, así como el Real Decreto 410/2016, por el que se aprueba el PEN 2017-2020, así como sus planes de ejecución, servirán de marco regulador de todas las operaciones estadísticas producidas con fines estatales, incluido el CMBD (11), garantizando el suministro de una información integral para el conjunto del Estado, la Unión Europea, las instituciones y los usuarios que así la requieran.

En el año 1998, se incluye en el modelo del CMBD-H una nueva variable, el tipo de financiación (12), creciendo su estructura de 14 a 15 variables.

En el año 2003, la *Ley 16/2003 de Cohesión y Calidad del SNS* dispone, a través de los artículos 53 y 55, el establecimiento de un sistema de información sanitaria para el SNS y la realización de estadísticas en materia sanitaria, reforzando aún más la necesidad de producción de un recurso como el CMBD.

Un año después, en el año 2004, se inicia la recogida de datos para el Conjunto Mínimo Básico de Datos de Atención Ambulatoria Especializada (CMBD-AAE). Desplegándose, a partir de este momento, dos productos a nivel nacional relacionados con el CMBD, el CMBD-H, definido en 1987 y obligatorio desde el año 1992, y el nuevo CMBD-AAE que nace para recoger la información clínico asistencial de los pacientes atendidos en área de hospitalización sin ingreso y hospitalización a domicilio. El modelo de datos del

CMBD-AAE será común al del CMBD-H e integrará información del paciente, del proveedor e información clínica. Esta información se integrará en el sistema único de información del CMBD pudiéndose explotar la información de forma desagregada. La actividad ambulatoria se irá integrando poco a poco, de inicio se incluye la actividad relativa a los procedimientos quirúrgicos realizados de forma ambulatoria y, posteriormente, el resto de modalidades asistenciales ambulatorias (procedimientos diagnóstico-quirúrgicos de alta complejidad, hospital de día, urgencias y consultas externas) (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2017a).

En el año 2005, de nuevo se produce un cambio en la estructura del CMBD, con la incorporación de dos nuevas variables, el código de identificación personal del paciente (CIP) y la identificación del servicio clínico responsable del alta del paciente. En estos momentos, el CMBD cuenta con 17 variables (ver tabla II).

Variable	Nombre de la variable	Año
1	Identificación del hospital	1987
2	Identificación del paciente	1987
3	Sexo	1987
4	Fecha de nacimiento	1987
5	Lugar de residencia	1987
6	CIP	2005
7	Financiación	1987
8	Tipo de financiación	1998
9	Fecha de ingreso	1987
10	Circunstancias de ingreso	1987
11	Diagnóstico principal y otros	1987
12	Procedimientos quirúrgicos y otros	1987
13	Procedimientos diagnósticos	1987
14	Fecha de alta	1987
15	Circunstancias al alta	1987
16	Identificación del médico responsables del alta	1987
17	Identificación del servicio clínico responsable del alta	2005

Tabla II. Relación de variables del CMBD por año de incorporación

2.2. El nuevo RAE-CMBD

El 10 de febrero de 2015 se publica el *Real Decreto 69/2015 por el que se regula el Registro de Actividad de Atención Sanitaria Especializada (RAE)*. La publicación de este registro fue la cul-

minación de una larga trayectoria de trabajo llevada a cabo por: *el Comité del CMBD, la Unidad Técnica de la Clasificación Internacional de Enfermedades (CIE) y la Subcomisión de Sistemas de Información del CISNS*. Esta nueva norma reguladora, que toma como base el CMBD existente y que, por este motivo, se va a reconocer por las siglas RAE-CMBD, amplía el alcance del CMBD a nivel de contenido, cobertura y finalidad, abarcando diferentes productos estadísticos y de información en un mismo instrumento, convirtiendo este nuevo registro como el principal referente y fuente de datos para la planificación, gestión sanitaria, evaluación e investigación clínica, epidemiológica y de servicios de salud en la actualidad y a futuro. El registro se implantará en soporte digital y su diseño y estructura deben facilitar el intercambio de datos, así como una explotación de los mismos por medios electrónicos (13).

El RAE-CMBD mantiene el objetivo de regular la actividad de la atención sanitaria especializada, sirviéndose del CMBD y extendiendo su estructura y contenido a un modelo homogéneo y normalizado integrado en el SISNS. Tal y como indica el Real Decreto 69/2015, en su artículo 2, el registro tiene una naturaleza administrativa y, como parte del SISNS, busca garantizar la disponibilidad de la información de la actividad sanitaria especializada, respondiendo a las necesidades de información de los diferentes agentes implicados en el SNS, así como: conocer la demanda asistencial y la morbilidad atendida, proporcionar a los registros autonómicos la información necesaria para la evaluación y control de la atención prestada y facilitar la realización de estadísticas de salud a nivel estatal, así como las que deriven de compromisos con organismos oficiales internacionales.

El RAE-CMBD amplía su modelo de datos de 17 a 31 variables, llevándose a cabo una revisión completa de su estructura y de las características de cada variable, tales como: su descripción, formato, longitud, listas de valores admitidos o de clasificaciones aceptadas, etc. Todo ello, buscando establecer los mecanismos necesarios para obtener un registro estandarizado mucho más eficiente que pueda garantizar una explotación óptima de la información. Para Sanz Sebastián (2017) *“la implantación del RAE-CMBD ha supuesto actuaciones sobre dos pilares fundamentales de la información sanitaria: por una parte, la adopción de la CIE10-ES como clasificación de referencia de codificación de la información clínica y por otra la adopción de un nuevo modelo de datos y ampliación del ámbito de aplicación de su antecesor el CMBD”*.

Hay que tener en cuenta que el 77% de las variables son de tipo administrativo e incluyen datos que hacen referencia a la identificación del paciente, proveedor de servicios y episodio asistencial, frente al 23% de las variables que incluyen información clínica. En la tabla III podemos observar el total de variables según su tipología, grupo (en función del tipo de información o datos que proporcionan) y número de variables relacionadas con cada grupo.

Tipo de variable	Grupo	Nº
Administrativa	Datos del paciente	8
	Datos Proveedor	2
	Datos del episodio asistencial	14
Clínicas	Información clínica	7
	Total	31

Tabla III. Variables del RAE-CMBD según tipología, grupos y número

En el RAE-CMBD se mantiene como unidad de registro el contacto asistencial, pudiendo ser clasificados estos contactos en función de la modalidad asistencial: hospitalización, hospitalización a domicilio, hospital de día médico, cirugía ambulatoria, procedimiento ambulatorio de especial complejidad y urgencia. En cuanto al origen de los datos, la fuente de información puede ser el informe clínico de alta o la historia clínica de cada paciente.

El proceso de implementación de dicho registro comenzó el 1 de enero de 2016 y tiene previsto extenderse hasta el 1 de enero de 2020. En los plazos en la disposición adicional cuarta del Real Decreto 69/2015, se indica que se iniciará la recogida de datos atendiendo a los siguientes criterios por tipo de contacto: hospitalización y cirugía ambulatoria en hospitales de agudos públicos y privados (1 de enero de 2016), hospitalización en otros hospitales públicos y privados (1 de enero de 2017), procedimientos ambulatorios de alta complejidad, hospital de día médico, hospitalización a domicilio y urgencias en hospitales públicos y privados (1 de enero de 2018), cirugía ambulatoria, procedimientos ambulatorios de alta complejidad y hospital de día médico en centros ambulatorios públicos especializados (1 de enero de 2019), procedimientos ambulatorios de alta complejidad, hospital de día médico, hospitalización a domicilio, urgencias y cirugía ambulatoria en hospitales de agudos y centros ambulatorios especializados privados (1 de enero de 2020). Estos plazos, según Sanz Sebastián (2017), del Ministerio de Sanidad, se están desarrollando con cierto retraso.

En julio de 2017 se encontraba a disposición la herramienta del RAE-CMBD para carga y validación por las Comunidades Autónomas. En octubre de 2017, 13 Comunidades Autónomas habían accedido a la aplicación y solamente 9 de ellas habían cargado datos. En esa misma fecha, de octubre de 2017, el nuevo sistema de información contaba ya con aproximadamente 2 millones de registros.

Con relación al sistema de clasificación de los diagnósticos clínicos para el RAE-CMBD, el CISNS fijó el 1 de enero de 2016 como fecha de entrada en vigor de la nueva CIE10-ES. La CIE10 es la décima revisión de la CIE que desde el año 1948 está a cargo de la OMS. La CIE ha sido durante muchas décadas el instrumento esencial que permitía la comparación nacional e internacional en el campo de la salud pública (Ramos Martín-Vegue, Vázquez-Barquero y Herrera Castanedo, 2002).

La transición a la nueva CIE10-ES tiene una gran complejidad, ya que resulta ser un sistema que presenta considerables cambios con respecto a la edición anterior, además de integrar diagnósticos y procedimientos (14). Uno de los principales cambios, y a su vez principal ventaja, afecta al volumen del número de entradas; se ha pasado de unos 14.000 diagnósticos aproximadamente con la CIE9-MC a más de 68.000 con el actual sistema de CIE10-ES.

Con respecto a los procedimientos el cambio ha sido aún mayor, de casi 4.000 a 72.000. También su estructura y sistema de codificación ha cambiado. La CIE10-ES incorpora nuevos capítulos, características y categorías. En cuanto al sistema de codificación pasa de ser un sistema numérico a alfanumérico (15). Todo ello completa un panorama de notables dificultades para su adaptación por los propios sistemas de información, así como para los codificadores, documentalistas, facultativos y gestores. La codificación de los diagnósticos y procedimientos es clave para la explotación de la información del CMBD y, tal y como señala Sierra (2016), *“el paso de CIE-9 a CIE-10 está lastrando el registro de información, pero la exhaustividad del nuevo código permitirá afinar mucho la gestión”*.

3. El ICMBD y el Portal Estadístico del SNS

El Ministerio de Sanidad, Servicios Sociales e Igualdad lanzó en el año 2016 la versión en línea del CMBD, denominada ICMBD (Indicadores y ejes de análisis del CMBD). Esta aplicación se encuentra disponible en la dirección de Internet: <http://icmbd.es>. La versión en línea se define como un modelo para el análisis y explotación de

los datos del registro del CMBD de una forma estructurada, fiable y sencilla (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2017b).

La aplicación ha sido desarrollada, a través de un convenio de colaboración, por el Grupo de Minería de datos de la Universidad de Cantabria (16) para el Ministerio de Sanidad. Este proyecto consistía en la implantación y soporte de un modelo de indicadores y ejes de análisis de los datos del CMBD-H desde el año 2001 hasta el año 2015.

El modelo integra una serie de cinco componentes básicos que pueden ser utilizados para la consulta, agrupación y explotación de los datos. Estos componentes son: 1) Dimensiones: constituyen los objetivos generales de conocimiento (por ejemplo, estilos de práctica clínica, eficiencia, efectividad clínica, frecuentación, resolutivez, etc.); 2) Ejes de análisis: son las líneas para el análisis de las dimensiones (por ejemplo: punto de corte, serie temporal, valor base, mejores prácticas, etc.); 3) Indicadores: son los datos o conjuntos de datos cuyo resultado informa sobre facetas concretas de una o varias dimensiones; 4) Niveles de clasificación de los indicadores y 5) Filtros.

Indicador	Descripción
1	Estancia media
2	Estancia media preoperatoria
3	Frecuentación en hospitalización
4	Hospitalizaciones potencialmente evitables
5	Procedimientos realizados en hospitalización
6	Tasa de ambulatorización quirúrgica
7	Tasa de cesáreas
8	Tasa de complicaciones
9	Tasa de infección nosocomial
10	Tasa de mortalidad
11	Tasa de reingresos
12	Informe resumen de todos los indicadores

Tabla IV. Relación de indicadores ICMBD

El modelo cuenta con un total de siete dimensiones y seis ejes de análisis diferenciados. Del mismo modo, cuenta con la posibilidad de selección de 12 indicadores para consulta y explotación de datos.

Por otra parte, el Portal Estadístico del SNS, una de las principales herramientas en línea del SISNS, incluye dentro de su consulta interactiva, desde el año 2017, la posibilidad de consulta del CMBD-H y del CMBD-AAE. Este recurso del SNS

incluye, además del CMBD, la posibilidad de consulta de una lista de los principales sistemas de información estadística del SNS (Encuesta Nacional de Salud (ENS), Barómetro Sanitario, Sistema de Información de Atención Primaria (SIAP), Sistema de Información de Atención Especializada (SIAE), Base de datos Clínicos de Atención Primaria (BDCAP), etc. (figura 1).



Figura 1. Captura de pantalla de la Consulta Interactiva del Portal Estadístico del SNS (<http://peestadistico.inteligenciadegestion.msssi.es/>)

En el caso del CMBD-H integra la información procedente de las Comunidades Autónomas en su repositorio de datos con una cobertura que va desde el año 1997 hasta el año 2015. La consulta permite cruzar datos por las variables básicas poblacionales (edad y sexo del paciente); las del episodio asistencial (tipo de ingreso y alta) y las variables clínicas (diagnósticos y procedimientos), a las que se añaden otras derivadas del uso de sistemas de clasificación de pacientes y estimaciones de costes. La información puede también obtenerse de una forma desagregada llegando al nivel de Grupo de Hospital o Comunidad Autónoma, para la consulta sobre hospitales, y hasta el nivel de Comunidad Autónoma y provincia, para la consulta sobre pacientes (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2017c).

En cuanto al CMBD-AAE se incluyen datos desde el año 2004 hasta el año 2015, integrando la práctica totalidad de la Cirugía Ambulatoria del SNS y parcialmente hospital de día médico y los hospitales privados, siendo aún muy escasa la información del resto de modalidades ambulatorias. La única variable de análisis es el número de contactos o casos y al igual que el CMBD-H

puede obtenerse información de forma desagregada hasta el nivel de Hospital o Comunidad Autónoma, para hospitales, y hasta el nivel de Comunidad Autónoma y Provincia, para pacientes (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2017d). La aplicación también facilita la reutilización de los datos proporcionando herramientas para la exportación a ficheros estructurados en diferentes formatos (xls, csv, etc.).

4. Más allá de la estadística: conocimiento útil y accesible

Los registros clínicos son la base principal para conocer la realidad asistencial. En ellos se recogen los hechos, incluyendo los datos que se van agregando y contextualizando en los sistemas para que la información se transforme en conocimiento y pueda ser aprovechado por las organizaciones, la sociedad en general y el SNS en su conjunto. Solamente con los registros del CMBD-H, desde el año 1997 hasta el año 2015 ha habido un total de 66.425.949 contactos o casos, con una media de 3.496.103 contactos anuales (ver figura 2).

Un sistema de información como el nuevo RAE-CMBD, doblando en número de variables al histórico CMBD, se transformará a muy corto plazo en un repositorio de datos masivos. Si planteamos una operación simple y multiplicamos las 31 variables del RAE-CMBD por la media de contactos anuales para el CMBD-H, considerando que estas variables sean completadas en su totalidad y que solamente se incluye un valor por variable, la cantidad de datos a registrar es de más de 108 millones de datos al año.

Estas grandes dimensiones hacen que, inevitablemente, el SNS y el Ministerio de Sanidad se planteen una estrategia sanitaria orientada al “big data” y, mucho más, teniendo en cuenta que el RAE-CMBD solamente representa una fuente de datos para el SISNS. Tal es así, que el nuevo Consejo Asesor de Sanidad (17), designado en febrero de 2018, con el cometido de establecer las bases del SNS para las próximas décadas, tenga entre sus encargos prioritarios la realización de “un análisis de las utilidades que tiene que tener el big data como elemento de información y conocimiento” para saber cómo afrontar decisiones para agilizar el diagnóstico precoz y mejorar las decisiones en cuanto a tratamientos, así como para dar una respuesta adecuada y satisfactoria al desafío que representan los retos demográficos (Huertas, 2018; Ministerio de Sanidad, Servicios Sociales e Igualdad, 2018). En este sentido, el Ministerio de Sanidad en el año 2009, fecha en la que ya se planteaba la necesidad de crear un nuevo registro para el CMBD,

comunicaba la necesidad de darle “*utilidad de verdad al CMBD*” indicando que “*una gigantesca base de datos como el CMBD debe servir para algo más que para producir estadísticas*” (Sierra, 2009).

En esa misma fecha, el Ministerio había seleccionado, dentro de la estrategia de explotación del CMBD, diez indicadores básicos con los que se pretendía generar conocimiento útil y accesible

para todos los potenciales usuarios. De igual manera, se señalaba que iba a ser necesario para poder explotar esta información utilizar “*minería de datos, generación de informes y herramientas versátiles para análisis multidimensional*” con objeto de ser capaces de generar estadísticas e informes, y modelizar indicadores de *benchmarking* que generen conocimiento mediante tecnología de *business intelligence* (Sierra, 2009).

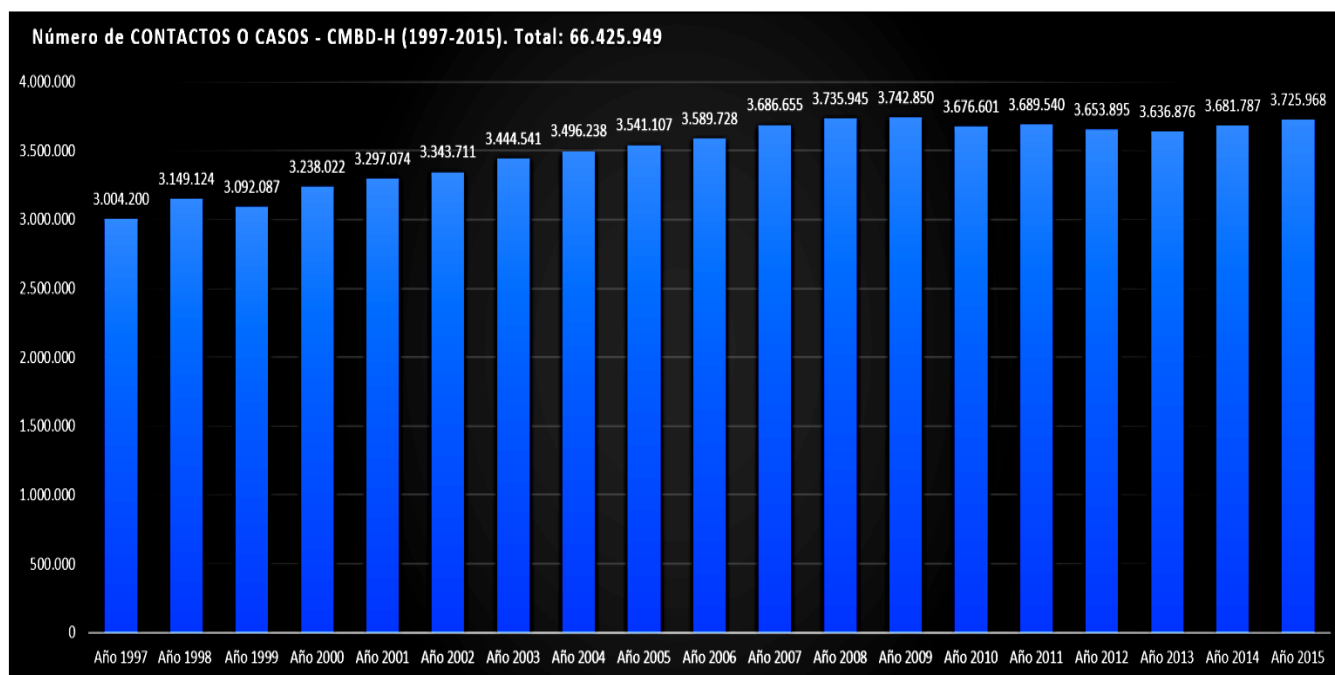


Figura 2. Registro de altas del CMBD-H. Número de contactos (1997-2015) (elaboración propia a partir de los datos del Ministerio de Sanidad, Servicios Sociales e Igualdad, <http://pestadistico.inteligenciadegestion.msssi.es>)

4.1. Una breve aproximación al Big data en salud

La Recomendación UIT-Y.3600, de la Unión Internacional de Telecomunicaciones (UIT), es la primera norma a nivel internacional donde se presentan los requisitos, las capacidades y los casos de uso de los volúmenes masivos de datos (big data). En este documento se define el concepto de *big data* como “*un paradigma para hacer posible la recopilación, el almacenamiento, la gestión, el análisis y la visualización, potencialmente en condiciones de tiempo real, de grandes conjuntos de datos con características heterogéneas*” (International Telecommunication Union, 2015). De una forma más sencilla, podríamos interpretar *big data* como el conjunto de datos e información de gran volumen y variedad, cuyo procesamiento resulta ser tan complejo que no puede realizarse por medio de bases de datos convencionales.

El resultado de toda explotación de datos masivos, más allá de los nuevos procesos, tecnologías o modelos de negocio que están surgiendo para dar soporte a su aprovechamiento, debe ser la obtención del valor o conocimiento que los propios datos encierran. En este sentido, Mayer-Schonberger y Cukier (2013) indican que “*los big data, los datos masivos, se refieren a cosas que se pueden hacer a gran escala, pero no a una escala inferior, para extraer nuevas percepciones o crear nuevas formas de valor*”. Según la Fundación COTEC para la Innovación, en su informe sobre «Generación de talento Big Data en España» (Fundación COTEC, 2017), “*la economía de los datos, implica que industrias, sectores y mercados, operarán completamente apoyados en el uso e intercambio de datos, así como en base a la “inteligencia” generada por su análisis. Como consecuencia, empresas de diversos y diferentes sectores de la industria, tendrán que organizar y orquestar múltiples aplicaciones y pro-*

cesos, para trabajar de forma integrada y eficiente, con el objetivo de optimizar entornos operativos enteros”.

Podríamos concluir, para definir este concepto, que los *big data* se caracterizan por las denominadas 6 V: *volumen, velocidad, variedad, veracidad, variabilidad y valor*. La primera característica, volumen, refleja las grandes cantidades de datos que se manejan; la segunda, velocidad, hace referencia a la frecuencia y rapidez de crecimiento de la información; la tercera, variedad, implica la diversidad de los tipos de información, estructuras y formatos que se pueden integrar; la cuarta, veracidad, expresa la autenticidad de los datos; la quinta, variabilidad, se refiere a los cambios de datos durante su procesamiento y ciclo de vida y, por último, la sexta, valor, hace mención a la calidad de los datos y el valor agregado que pueden aportar para la creación de conocimiento (Wang y Alexander, 2015).

El registro de la información y la normalización de los procesos ha provocado que organizaciones de todo tipo acumulen una ingente cantidad de datos. En el sector sanitario esta situación no ha pasado desapercibida, al contrario, monitorizar la situación de la salud de la población está aportando un volumen enorme de datos (volumen), datos que crecen a gran velocidad y de forma exponencial haciéndose necesario el uso de las nuevas tecnologías avanzadas para su captura y tratamiento con objeto de extraer el valor de los mismos de una forma eficiente (velocidad). Además, en el SNS existen una gran variedad de sistemas de información, que necesitan interrelacionarse e interoperar y que incluyen datos de salud de diferentes contextos de atención y especialización que, en muchas ocasiones, se encuentran sin normalizar, datos no estructurados, y sin codificar (variedad y variabilidad). Los datos de salud deben cumplir con la obligación de ser fidedignos y, para ello, hay que fomentar que la calidad de la información registrada en los sistemas sea, no solamente veraz, sino también útil y significativa para la seguridad y evolución de los pacientes y para la toma de decisiones de los profesionales (veracidad y valor).

Se espera, por tanto, que las aplicaciones de *big data* en el sector de la salud tengan un alto potencial para mejorar la eficiencia y la calidad en la prevención de enfermedades y la provisión de los cuidados. La medicina se está convirtiendo en una ciencia de la información y en pocos años se dispondrán de billones de datos de cada individuo, siendo un desafío el desarrollo de la tecnología de la información que pueda conducir dichos datos a hipótesis reales sobre los propios individuos (Kim, 2017).

En cuanto a la procedencia de los datos de salud, su origen puede encontrarse, entre otras fuentes, en: la historia clínica electrónica, los sistemas de prescripción, los sistemas de almacenamiento y comunicación de imágenes y pruebas de laboratorio, las bases de datos construidas con finalidades clínicas y estadísticas, los dispositivos médicos y *wearables*. En el informe “*Big data en salud digital*” (Fundación Vodafone España y Red.es, 2017) se indica que, de una forma operativa, los datos masivos deberán permitir a los centros sanitarios disponer de información práctica y precisa sobre la demanda de servicios y la disposición y calidad de los mismos, convirtiéndose en un recurso muy valioso para la gestión. Los datos masivos en salud, podrán estar disponibles en diferentes formatos (texto, audio, video u otros registros biométricos) siendo susceptibles de ser analizados para proporcionar información nueva y útil para los sistemas de salud. Del mismo modo, este documento expresa la trascendencia que el *big data* tendrá como elemento clave para el sostenimiento presente y futuro del SNS.

4.2. Analítica predictiva, descriptiva y modelos de decisión para generar conocimiento

La utilización de grandes registros de datos masivos y el empleo de herramientas analíticas, puede aportar información y evidencia para la práctica asistencial (Angus, 2017). Tal y como indica Martínez Sesmero (2015)

[...] la gran mayoría de los agentes que participan en las estructuras de los servicios de salud reconocen que el análisis del big data puede ofrecer nuevas posibilidades en la elaboración de modelos predictivos, patrones de comportamiento, el descubrimiento de nuevas necesidades, reducir riesgos, así como proveer servicios más personalizados, todo ello en tiempo real y teniendo en cuenta toda la información relevante.

Actualmente, se están utilizando tres tipos de modelos para el análisis de datos masivos: 1) los *modelos predictivos* que estiman las probabilidades, a partir de los datos, detectando tendencias y pronosticando posibles situaciones en el futuro; 2) los *modelos descriptivos* que se encargan de representar las relaciones entre individuos, grupos, poblaciones y otros factores; y 3) los *modelos de optimización o decisión*, que son un híbrido entre los dos anteriores y se centran en aportar el valor de los datos para dar soporte a la toma de decisiones. Estos modelos pretenden generar una nueva inteligencia clínica basada en el procesamiento de los datos de forma continuada y en la exploración e interpretación de los mismos, buscando extraer el conocimiento que de otro modo se nos podría haber pasado por alto o podría encontrarse oculto a nuestros ojos.

Para todos estos casos y métodos de análisis, una herramienta como el RAE-CMBD cumple con las características adecuadas para ser la fuente de información principal y mejor candidata para su tratamiento por herramientas y tecnologías *big data*. Hay que tener en cuenta que el RAE-CMBD es un producto consolidado y que presenta grandes ventajas para su análisis, ya que contiene datos normalizados, estructurados e interoperables. Además, los datos son depurados y revisados, mejorando enormemente su calidad y beneficios derivados.

Existe, asimismo, una amplia experiencia por parte de los centros sanitarios en la recolección de datos para el CMBD, en la aplicación de garantías de seguridad y privacidad, en la protección de datos personales y en la clasificación de la información clínica. De igual modo, hay una especial sensibilidad, reflejada en la legislación vigente y transmitida a lo largo del tiempo por medio del trabajo desarrollado por comités, grupos de trabajo en el SNS, así como por el propio Ministerio de Sanidad y las Comunidades Autónomas, de la utilidad e interés que esta fuente de datos tiene para la gestión y la investigación sanitaria a nivel nacional e internacional.

Todos estos motivos hacen que las posibilidades de análisis del RAE-CMBD, a todos los niveles, sean inmensas. El nuevo producto sigue integrando información demográfica del paciente, de su situación clínica, del proceso asistencial, de los procedimientos realizados, del servicio responsable y de los resultados clínicos (mortalidad, complicaciones o reingresos) y, además, mejora la información recogiendo datos de las diferentes modalidades asistencializadas, públicas y privadas. Es cierto que aún queda espacio para su evolución y desarrollo, pues acaba de comenzar una nueva etapa para el registro de datos relacionados con el CMBD y tendrá que observarse su progreso, aunque no cabe duda de que la abundancia de los datos que aportará serán una gran oportunidad para poner a prueba la tecnología y los modelos analíticos del *big data* en el SNS.

5. Conclusiones

El CMBD es una de las principales fuentes de información poblacional del SISNS. Como tal, se constituye como un recurso estratégico para el conocimiento de la actividad asistencial especializada, tanto de hospitalización como ambulatoria, pública y privada. Con la regulación en el año 2015 del RAE-CMBD se incluyen considerables ventajas a este recurso en relación a modelos anteriores, añadiendo una estructura de datos nor-

malizados y homogéneos que proporcionará información sobre la actividad asistencial, las patologías atendidas, su complejidad, su abordaje clínico y los costes que suponen para el sistema. El RAE-CMBD, además, mantiene una hoja de ruta para que su implantación obligatoria se realice de una forma progresiva atendiendo al tipo de hospital y ámbito asistencial; dando comienzo en enero del año 2016 y finalizando en enero de 2020. Se espera que el registro clínico del nuevo RAE-CMBD aporte tal magnitud de datos que generarán un repositorio de grandes dimensiones a muy corto plazo.

Para que los datos sean de utilidad deben aportar conocimiento, para ello hay que fomentar el desarrollo de modelos de negocio que sean útiles en el aprovechamiento de los datos masivos. Estos modelos deberían basarse en el consenso, la armonización y normalización de las estructuras y registro de datos, buscando facilitar la interoperabilidad y la explotación de los datos de salud en beneficio del sistema, los ciudadanos, el soporte a la toma de decisiones y la investigación científica.

Potenciar el análisis de los datos masivos o *big data* será una apuesta a largo plazo para las instituciones y autoridades sanitarias, no sin riesgo o complejidad, pues aún se desconocen todas las posibilidades que pueden ofrecer las tecnologías y técnicas alrededor de la gestión y análisis de datos a gran escala. En cualquier caso, una estrategia nacional compartida en el SNS para el buen gobierno de datos y la aplicación de *big data* sobre los recursos de información nacionales, tales como el RAE-CMBD y otras fuentes del SISNS, podría aportar valor y evitar riesgos, detectar patrones de comportamiento sobre la población y las enfermedades, ofrecer una mayor inteligencia y control sobre los procesos sanitarios y ahorrar costes al sistema y, posiblemente, aumentar la eficacia y eficiencia sobre la asistencia, la calidad de la atención y un mayor conocimiento sobre las demandas o necesidades de los diferentes agentes.

Para ello, como primer paso imprescindible, habría que asegurar la calidad de los datos desde su registro, mejorar la coordinación e intercambio de datos e información en el conjunto del SNS, de modo que, posteriormente, se puedan reutilizar los datos y establecer relaciones entre los mismos, trabajar con nuevos indicadores y métodos de análisis más potentes que puedan ofrecer respuestas a las preguntas más complejas, aportando informaciones fiables y valiosas para la interpretación y comprensión del estado de la salud, así como para pronosticar tendencias de futuro de la asistencia sanitaria.

Tal y como indica Kumholz (2014) la incorporación de grandes datos y análisis de próxima generación en la investigación, la práctica clínica y la información sobre salud demográfica requerirá no solo nuevas fuentes de datos, sino también nuevas ideas, capacitación y herramientas. Ya existen herramientas o sistemas, como por ejemplo *Hadoop* (18), que se están utilizando para almacenar, procesar y analizar grandes volúmenes de datos con cientos de terabytes, petabytes e, incluso, exabytes. Se tratan de herramientas de nueva generación que consiguen escalar desde unos pocos servidores a miles de máquinas y proponen soluciones para secuenciación que puede basarse en la información de la historia de salud electrónica de los pacientes y, de este modo, detectar patrones de riesgo relacionados con grupos poblacionales, revelar perfiles (clusters), redes de comorbilidad, eventos adversos de los medicamentos, patrones terapéuticos, etc. También parecen prometedoras las herramientas que aportan soluciones de análisis visual, con posibles aplicaciones para el desarrollo y mejoras en el campo de la salud pública (19).

El RAE-CMBD, dada sus considerables mejoras para facilitar una información estructurada e intercambiable, así como las expectativas que existen sobre su crecimiento y posibilidades de explotación para generar conocimiento, hacen que sea la fuente de información idónea para ser sometida, de forma gradual, a las técnicas y modelos de análisis para datos masivos. Sin embargo, hay que tener en cuenta que aún existen grandes incógnitas por resolver sobre lo que la tecnología *big data* en el dominio de la salud podrá ofrecer, por lo que tendrá que ser asumida por el SNS con cierta prudencia y tiempo.

La ciencia de los datos es una realidad, una oportunidad y un desafío cuyo avance y mejora dependerá de cómo se vayan transformando los sistemas de información en el futuro, facilitando datos en tiempo real, organizados, precisos y procesables. Tampoco hay que olvidar que, muy posiblemente, la mayor parte del conocimiento clínico aún se encuentre oculto, sumergido en la profundidad de las historias clínicas de los pacientes y en la gran cantidad de información conservada, no normalizada y registrada por los profesionales por medio del lenguaje natural. En este sentido, el *big data* tendrá también mucho en lo que contribuir.

Notas

(1) Tal y como indica la Ley de Cohesión y Calidad del SNS (2003), en su artículo 55, "el sistema de información sanitaria contemplará específicamente la realización de estadísticas para fines estatales en materia sanitaria, así como las de interés general supracomunitarias y las que

se deriven de compromisos con organizaciones supranacionales e internacionales, que se llevarán a cabo con arreglo a las determinaciones metodológicas y técnicas que establezca el Ministerio de Sanidad, consultado el Consejo Interterritorial del SNS".

- (2) Según el Real Decreto 69/2015, en su artículo 6, se entiende por contacto la atención sanitaria prestada bajo la misma modalidad asistencial y de forma ininterrumpida por un proveedor sanitario a un paciente.
- (3) Este registro viene a responder a las necesidades de información de los diferentes agentes del SNS, tal y como se establece en los artículos 53 y 55 de la Ley 16/2003, de 28 de mayo, de cohesión y calidad del Sistema Nacional de Salud.
- (4) El artículo 23 de la Ley 14/1986, de 25 de abril, General de Sanidad dispone que las Administraciones Sanitarias, de acuerdo con sus competencias, crearán los registros y elaborarán los análisis de información necesarios para el conocimiento de las distintas situaciones de las que puedan derivarse acciones de intervención de la autoridad sanitaria (Ministerio de Sanidad y Consumo, 2004).
- (5) El Reglamento 1338/2008, en su artículo 1.2, indica que "las estadísticas incluirán, en forma de conjunto de datos armonizados y comunes, la información que requieran la acción comunitaria en el ámbito de la salud pública, el apoyo a estrategias nacionales para el desarrollo de una asistencia sanitaria de alta calidad". Los datos estadísticos serán la fuente principal para la lista de Indicadores de Salud de la Unión Europea (ECHI) que buscan ofrecer una visión general del estado de salud, los determinantes de la salud y los sistemas sanitarios de los estados miembros. Para más información sobre ECHI se puede consultar: http://ec.europa.eu/health/indicators_data/indicators_en
- (6) El INE viene recogiendo desde el año 1951, a través de la *Encuesta de Morbilidad Hospitalaria (EMH)*, la evolución de la morbilidad hospitalaria en España. Esta encuesta, de carácter anual, ofrece información sobre las altas hospitalarias con internamiento y la estancia media de las mismas en función del diagnóstico principal asociado al alta. Se considera alta hospitalaria el procedimiento por el que un paciente ingresado en un centro o establecimiento sanitario deja de ocupar una cama de hospitalización por curación, mejoría, fallecimiento, traslado o alta voluntaria (Instituto Nacional de Estadística, 2016).
- (7) Este Real Decreto 1360/1976 se encuentra derogado por el Real Decreto 69/2015, de 6 de febrero, por el que se regula el Registro de Actividad de Atención Sanitaria Especializada.
- (8) Hay que destacar que el CISNS, como órgano de coordinación del SNS, fue creado en abril de 1987, siendo la propuesta del CMBD-H uno de los temas aprobados durante el mismo año de su creación.
- (9) Este Comité Técnico del CMBD, tal y como se recoge en el acuerdo del CISNS (1988), se constituye como una Comisión delegada del CISNS.
- (10) Dentro de las funciones de este Comité Técnico del CMBD se incluye la de asegurar la traducción completa de la Clasificación Internacional de Enfermedades 9ª edición – Modificación Clínica (CIE-9-MC). La CIE-9-MC ha sido la principal fuente para la codificación de diagnósticos en el CMBD desde sus inicios hasta el año 2016, fecha en la que se decide sustituir por la nueva versión CIE-10-ES.
- (11) En el PEN 2017-2020 figura la operación anual de explotación estadística dentro del Inventario de Operaciones

Estadísticas (IOE) la ficha 54011 – 7734. Explotación estadística del CMBD (RAE-CMBD). Demanda Hospitalaria y Morbilidad Atendida.

- (12) Esta variable, tipo de financiación, hace referencia al responsable de la financiación económica de los costes del contacto: SNS, convenios bilaterales con otros países europeos, mutualidades de regímenes especiales de la Seguridad Social de funcionarios, mutuas colaboradoras de la Seguridad Social, financiación privada directa, etc. Se diferencia de la variable financiación que indica si es pública o privada.
- (13) El fichero del RAE-CMBD se crea por Orden SSI/1885/2015, de 8 de septiembre, por la que se modifica la Orden de 21 de julio de 1994, por la que se regulan los ficheros con datos de carácter personal gestionados por el Ministerio de Sanidad y Consumo.
- (14) En el año 1991, el Ministerio de Sanidad publicó la primera edición completa en castellano de la CIE9-MC y ha realizado desde entonces, con periodicidad aproximada bianual, las sucesivas actualizaciones siguiendo la estela de las publicadas por el Centro Nacional de Estadísticas de Salud (NHSC) del Centro para el Control de las Enfermedades de los Estados Unidos (CDC) (Ministerio de Sanidad, Política Social e Igualdad, 2010).
- (15) La CIE10-ES es una clasificación alfanumérica que contiene entre 3 y 7 dígitos, siendo el primer carácter siempre una letra.
- (16) Para más información sobre este grupo consultar: <http://www.meteo.unican.es/research/datamining>
- (17) El Consejo Asesor de Sanidad es el máximo órgano consultivo y de asistencia a la persona titular del Departamento en la formulación de la Política Sanitaria, en este caso al ministro/a titular del Ministerio de Sanidad.
- (18) Hadoop es un proyecto de código abierto que incluye diferentes soluciones para datos masivos administrado por la Fundación Apache Software. Para más información se puede consultar: <http://hadoop.apache.org/>
- (19) Para más información sobre aplicaciones visuales para big data se puede consultar la revista Forbes (Marr, 2017) donde se ha publicado un listado con las 7 mejores herramientas de visualización de datos.

Referencias

- Alfaro Latorre, M (2012). Utilizando el sistema de información sanitaria del SNS. Madrid: Escuela Nacional de Sanidad. http://e-spacio.uned.es/fez/eserv/bibliuned:500569/n2.5_Utilizando_Sistema_de_Informaci__n_Sanitaria_del_SNS.pdf
- Angus, DC. (2017). Wheter to intubate during cardiopulmonary resuscitation: Conventional wisdom vs Big Data. // JAMA. 317:5, 477-478.
- Calvete, A (2003). El programa de acción de la Unión Europea en el ámbito de la salud pública. // Revista Española de Salud Pública. 77:5, 517-521.
- Carrillo, E; Juncosa, S; Bólibar, B; Parkin, D; Hutchinson, A y Fisher, P (1991). Conjuntos Mínimos de Datos en la Atención Ambulatoria. //Gaceta Sanitaria, 26 (5): 225-230.
- Consejo Interterritorial del Sistema Nacional de Salud (1987). Acuerdo nº 30. Pleno del 14 de diciembre de 1987. Acta nº 5. <https://www.msssi.gob.es/organizacion/consejoInterterri/docs/30.pdf>
- Consejo Interterritorial del Sistema Nacional de Salud (1988). Acuerdo nº 52. Pleno del 13 de julio de 1988. Acta nº 8. <https://www.msssi.gob.es/organizacion/consejoInterterri/docs/52.pdf>
- Decisión Nº 1786/2002/CE del Parlamento Europeo y del Consejo de 23 de septiembre de 2002 relativa a la adopción de un programa de acción comunitario en el ámbito de la salud pública (2003-2008). Diario Oficial de la Unión Europea. 9 de octubre de 2002. L271: 1-11.
- Decisión Nº 1350/2007/CE del Parlamento Europeo y del Consejo de 23 de octubre de 2007 por el que se establece el segundo programa de acción comunitario en el ámbito de la salud pública (2008-2013). Diario Oficial de la Unión Europea. 20 de noviembre de 2007. L301: 3-13.
- Foster, J; Conrick M. (1998). Nursing Minimun Data Sets: Historical Perspective and Australian Development. En: Saba, V; Pocklington DB; Miller KP (ed.). Nursing and Computers. An anthology, 1987-1996. New York: Springer-Verlag.
- Fundación COTEC (2017). Generación de talento Big Data en España. Madrid: Fundación COTEC para la Innovación. <http://cotec.es/media/BIG-DATA-FINAL-web.pdf>
- Fundación Vodafone España y Red.es (2017). Big data en salud digital. Madrid: Fundación Vodafone España. http://www.fundacionvodafone.es/sites/default/files/informe_big_data_en_salud_digital.pdf
- Huertas, J. (2018). Sanidad quiere nuevos roles sanitarios, medicina personalizada y "big data". // Redacción Médica, 14 de febrero de 2018.
- Instituto Nacional de Estadística (2016). Encuesta de Morbilidad Hospitalaria. Año 2015. Nota de prensa de 29 de noviembre de 2016. <http://www.ine.es/prensa/np1005.pdf>
- International Telecommunication Union (2015). Recommendation ITU-T Y.3600. Big data. Cloud computing-based requirements and capabilities. Series Y: Global Information infrastructure, Internet Protocol aspects and next-generation networks. Cloud Computing. Ginebra: ITU. <https://www.itu.int/rec/T-REC-Y.3600-201511-1>
- Kim J. (2017). Big data, health informatics, and the future of cardiovascular medicine. // Journal of the American College of Cardiology. 69:7, 899-902.
- Kumholz, HM (2014). Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system. // Health Aff (Millwood). 33:7, 1163-70.
- Ley 16/2003, de 28 de mayo, de cohesión y calidad del Sistema Nacional de Salud. // Boletín Oficial del Estado. 29 de Mayo de 2003. 128; 20567-20588.
- Marr, B. (2017). The 7 Best Data Visualization Tools in 2017. Forbes, 20 de julio. <https://www.forbes.com/sites/bernardmarr/2017/07/20/the-7-best-data-visualization-tools-in-2017/#4374e0f96c30>
- Martínez Sesmero, JM. (2015). "Big Data"; aplicación y utilidad para el sistema sanitario. // Farmacia Hospitalaria. 39:2, 69-70.
- Mayer-Schonberger, V. y Cukier, K. (2013). Big data. La revolución de los datos masivos. Madrid: Turner.
- Ministerio de Sanidad y Consumo (2004). Ley 14/1986, de 25 de abril, General de Sanidad. Colección Textos legales. Serie Sanidad; 1. Madrid : Ministerio de Sanidad y Consumo.
- Ministerio de Sanidad, Política Social e Igualdad (2010). Hacia la CIE-10. Documento Técnico 1. Madrid: Instituto de Información Sanitaria. Agencia de Calidad del SNS. https://www.msssi.gob.es/estadEstudios/estadisticas/normalizacion/CIE10/2010_HaciaCIE10.pdf
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2014). Sistema de Información Sanitaria del Sistema Nacional de Salud. <https://www.msssi.gob.es/fr/estadEstudios/estadisticas/sisInfSanSNS/pdf/SISNS.pdf>
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2017a). Registro de Altas – CMBD estatal. Atención Ambulatoria Especializada CMBD-AAE. Manual de definiciones y glo-

- sario de términos. <http://pestadistico.inteligenciadegestion.msssi.es/publicoSNS/comun/DescargaDocumento.aspx?IdNodo=6435>
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2017b). i-CMBD. Indicadores y ejes de análisis del CMBD de Hospitalización del Sistema Nacional de Salud. <http://icmbd.es/docs/modeloIndicadores.pdf>
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2017c). Portal Estadístico. Conjunto Mínimo Básico de datos – Hospitalización (CMBD-H). Acerca de. <http://pestadistico.inteligenciadegestion.msssi.es/publicoSNS/Comun/ArbolNodos.aspx?idNodo=6383>
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2017d). Portal Estadístico. Conjunto Mínimo Básico de datos – Hospitalización (CMBD-AAE). Acerca de. <http://pestadistico.inteligenciadegestion.msssi.es/publicoSNS/Comun/ArbolNodos.aspx?idNodo=6425>
- Ministerio de Sanidad, Servicios Sociales e Igualdad (2018). Notas de Prensa. Dolors Montserrat pide al Consejo Asesor que trabaje para establecer las bases de Sistema Nacional de Salud de próximas décadas (14 de febrero de 2018). <https://www.msssi.gob.es/gabinete/notas-Prensa.do?id=4297>
- Orden de 6 de septiembre de 1984 por la que se regula la obligatoriedad del informe de alta. Boletín Oficial del Estado. 14 de septiembre de 1984. 221; 26685-26686.
- Orden SSI/1885/2015, de 8 de septiembre, por la que se modifica la Orden de 21 de julio de 1994, por la que se regulan los ficheros con datos de carácter personal gestionados por el Ministerio de Sanidad y Consumo. Boletín Oficial del Estado. 17 de septiembre de 2015. 223; 81999-82001.
- Ramos Martín-Vegue, AJ, Vázquez-Barquero JL y Herrera Castanedo, S. (2002). CIE-10 (I): Introducción, historia y estructura general. // Papeles Médicos. 11 :1, 24-35.
- Real Decreto 1360/1976, de 21 de mayo, por el que se hace obligatorio el uso, por parte los establecimientos sanitarios con régimen de internado, de un libro de registro. //Boletín Oficial del Estado. 23 de junio de 1976. 150, 12200.
- Real Decreto 136/1993, de 29 de enero, por el que se aprueba el Plan Estadístico Nacional 1993-1996. // Boletín Oficial del Estado. 5 de marzo de 1993. 55; 7012-7013.
- Real Decreto 1658/2012, de 7 de diciembre, por el que se aprueba el Plan Estadístico Nacional 2013-2016. // Boletín Oficial del Estado. 8 de diciembre de 2012. 295; 84141-84322.
- Real Decreto 69/2015, de 6 de febrero, por el que se regula el Registro de Actividad de Atención Sanitaria Especializada. // Boletín Oficial del Estado. 10 de febrero de 2015. 35; 10789-10809.
- Real Decreto 410/2016, de 31 de octubre, por el que se aprueba el Plan Estadístico Nacional 2017-2020. // Boletín Oficial del Estado. 18 de noviembre de 2016. 279; 80830-80995.
- Reglamento (CE) N° 1338/2008 del Parlamento Europeo y del Consejo de 16 de diciembre de 2008 sobre estadísticas comunitaria// s de salud pública y de salud y seguridad en el trabajo. Diario Oficial de la Unión Europea. 31 de diciembre de 2008. L354: 70-81.
- Reglamento (CE) N° 282/2014 del Parlamento Europeo y del Consejo de 11 de marzo de 2014 relativo a la creación de un tercer programa de acción de la Unión Europea en el ámbito de la salud para el período 2014-2020 y por el que se deroga la Decisión N° 1350/2007/CE. // Diario Oficial de la Unión Europea. 21 de marzo de 2014. L86: 1-13.
- Salvador Oliván, J.A. (1997). Sistemas de información hospitalarios: el C.M.B.D. Scire. 3:2, 115-130.
- Sanz Sebastián, C. (2017). El Registro de Actividad de Atención Sanitaria Especializada (RAE-CMBD) la experiencia de implantación. // 8º Foro sobre el Sistema de Información del Sistema Nacional de Salud. Madrid, Ministerio de Sanidad, Servicios Sociales e Igualdad. 18 de octubre de 2017.
- Sierra, R. (2009). El Ministerio debe darle utilidad de verdad al CMBD. // Diario Médico, 3 de noviembre de 2009.
- Sierra, R. (2016). Un CMBD de valor incalculable... pero no este año. // Diario Médico, 10 de octubre de 2016.
- Treviño FM (1988). Uniform minimum data sets: in search of demographic comparability. // American Journal of Public Health. 78:2, 126-127.
- Wand, L.; Alexander, CA. (2015). Big data in medical applications and health care. // Current Research in Medicine. 6:1, 1-8.
- Zapatero Gaviria, A (2010). ¿Por qué es importante el informe médico de alta? // Revista Clínica Española. 210:7, 355-358.

Enviado: 2018-05-28. Segunda versión: 2018-06-07.
Aceptado: 2018-06-17.
