

# KodifiKa. Un sistema de codificación automática de los diagnósticos generados en consultas externas

Yetano Laguna J\*, Moreno Fernández de Leceta A\*\*, Aulestia Seoane L\*\*\*, Salvador Blanco J\*, López Arbeloa G\*\*\*

\*Hospital de Galdakao. Galdakao, Vizcaya.

\*\*Instituto Iberoamericana de Innovación. Vitoria.

\*\*\*Organización Central de Osakidetza. Vitoria.

e-mail: javier.yetanolaguna@osakidetza.net

---

## Resumen

El desafío actual de los sistemas de información sanitarios españoles es el desarrollo del Conjunto Mínimo Básico de Datos (CMBD) ambulatorio para conocer qué procesos se atienden en Consultas Externas y en Urgencias hospitalarias. Reproducir el circuito de codificación de diagnósticos que hemos empleado en el CMBD de hospitalización es inviable por su laboriosidad y alto coste. El programa KodifiKa consigue codificar los diagnósticos de forma automática usando inteligencia artificial sin aumento de personal.

*Palabras clave:* Conjunto Mínimo Básico de Datos ambulatorio, CIE-9-MC, Sistema de información hospitalario.

## KodifiKa. A system of automatic codification of ambulatory care diagnoses

### Abstract

The current challenge of the spanish sanitary information systems is the development of a ambulatory Minimun Basic Data Set to know what processes are attended in ambulatory care. To reproduce the circuit of manual codification of diagnoses that has been used for hospital discharges is unviable for his laboriousness and high cost. The program KodifiKa obtains an automatic codification of the diagnoses with ICD-9-CM using artificial intelligence without increase of personnel.

*Key words:* Ambulatory Minimun Basic Data Set, ICD-9-CM, Hospital information systems.

---



## Introducción

El Conjunto Mínimo Básico de Datos (CMBD) es una excelente base de datos que proporciona información relevante sobre la actividad y calidad de la asistencia hospitalaria en el área de hospitalización de nuestro Sistema Nacional de Salud. Gracias a él, disponemos en España del registro de más de cinco millones de altas anuales codificadas desde 1990. Se consigue gracias a la codificación de los diagnósticos y procedimientos con la Clasificación Internacional de Enfermedades, novena revisión, Modificación Clínica (CIE-9-MC) realizada por codificadores expertos.

En el área de Urgencias y Consultas Externas hospitalarias no existe un CMBD oficial, aunque se han realizado experiencias interesantes<sup>1,2</sup>. El área de Consultas Externas es la hermana pobre de la asistencia especializada de los hospitales españoles en cuanto a información normalizada de los procesos que atiende. Actualmente, la información clínica normalizada de consultas es nula y no se sabe a qué tipo de pacientes atiende cada especialidad, cuáles son las complicaciones postoperatorias extrahospitalarias o cuál es la variabilidad entre diferentes hospitales del mismo nivel. Esta carencia se debe a la enorme dificultad de codificar los diagnósticos de las 75 millones de consultas de asistencia especializada que se producen cada año en la sanidad pública española (cuatro millones en el País Vasco). Si se pretendiese codificar esta actividad mediante codificadores, se puede estimar que sería necesario contratar a 1900 codificadores en España, lo que supondría unos 57 millones de euros (101 codificadores y tres millones de euros en el País Vasco), algo impensable en la actual situación de crisis económica.

Sin embargo, gracias a la progresiva implementación de la Historia Clínica Electrónica (HCE) en nuestro país y, en concreto de la HCE-Osabide Global en los hospitales de Osakidetza/Servicio Vasco de Salud, se puede llegar a codificar los diagnósticos de las consultas externas mediante la codificación automática de los mismos.

## Objetivos

El objetivo es conseguir una codificación automática de calidad de los diagnósticos generados en cada una de las visitas de las consultas externas de Osakidetza/Servicio Vasco de Salud realizadas en la HCE-Osabide Global. Para lograrlo es necesario:

- Crear una base de datos de diagnósticos ya codificados con la CIE-9-MC por codificadores expertos. Las expresiones diagnósticas son las expresadas en lenguaje natural por los médicos en la práctica diaria en Osakidetza.
- Implementar unos campos de diagnósticos en la HCE-Osabide Global para que el médico, al introducir una evolución clínica en la consulta externa, ponga en lenguaje natural el o los diagnósticos motivo de la consulta.
- Desarrollar un programa informático motor de codificación que trate la expresión diagnóstica introducida por el médico y la codifique con la CIE-9-MC. El motor de codificación debe usar inteligencia artificial y el conocimiento almacenado en la base de datos de diagnósticos ya codificados.
- Diseñar un módulo de control de la calidad de la codificación realizada



por el motor para uso de la Unidad de Codificación.

## Metodología empleada y cronograma

En 2008 se obtuvo una beca del Gobierno Vasco para el "Desarrollo del Conjunto de Datos Básicos de la Asistencia Ambulatoria Especializada (CBD-AAE) de Osakidetza".

En 2009 se hizo una propuesta de CBD-AAE al Departamento de Sanidad. Se constató la imposibilidad de la codificación manual de millones de asistencias ambulatorias y la necesidad de conseguirla mediante una codificación automática. Se formó un grupo de trabajo que diseñó el proyecto KodifiKa formado por un coordinador (médico del Hospital de Galdakao) y un investigador (del Instituto Iberoamericana de Innovación). Se presentó el proyecto a la Subdirección de Calidad de Osakidetza para su aprobación y anclaje en la HCE-Osabide Global. Se utilizó una base de datos con 120 000 expresiones diagnósticas en lenguaje natural ya codificadas por codificadores recogida previamente en el Hospital de Galdakao (de 2006 a 2009).

En 2010 se consiguió la versión inicial del programa y se probó en las evoluciones de consultas de la HCE Osabide-Global. Se consiguió una codificación automática del 81% de los casos.

En 2011 se extendió su uso a 635 111 casos. Se obtuvo el asesoramiento de la Facultad de Informática de San Sebastián para el procesamiento del lenguaje natural, con lo que se introdujeron mejoras en el programa, como la introducción de un diccionario de *stop words*. Esta herramienta elimina de la expresión diagnóstica introducida por

el médico todas aquellas palabras que no influyan en la codificación como "estudio", "discreto" o "esporádico". Con estas mejoras y con el aumento del "conocimiento" de la base de datos (a 140 000 diagnósticos distintos ya codificados), el porcentaje de codificación automática ascendió al 86%.

En 2012 se está produciendo la extensión de la HCE-Osabide Global a todos los médicos de todos los hospitales y ambulatorios de asistencia especializada de Osakidetza, con lo que el número de consultas llegará a los cuatro millones anuales. Además, se van a introducir nuevas mejoras en KodifiKa para que el motor de codificación aumente el 86% de codificación automática actual. Se logrará ampliando la fuente de información actual, solo el diagnóstico, a todo el contexto (información que está en el sistema, como el sexo y la edad del paciente, o todo el texto de la evolución escrita por el médico).

## Resultados obtenidos

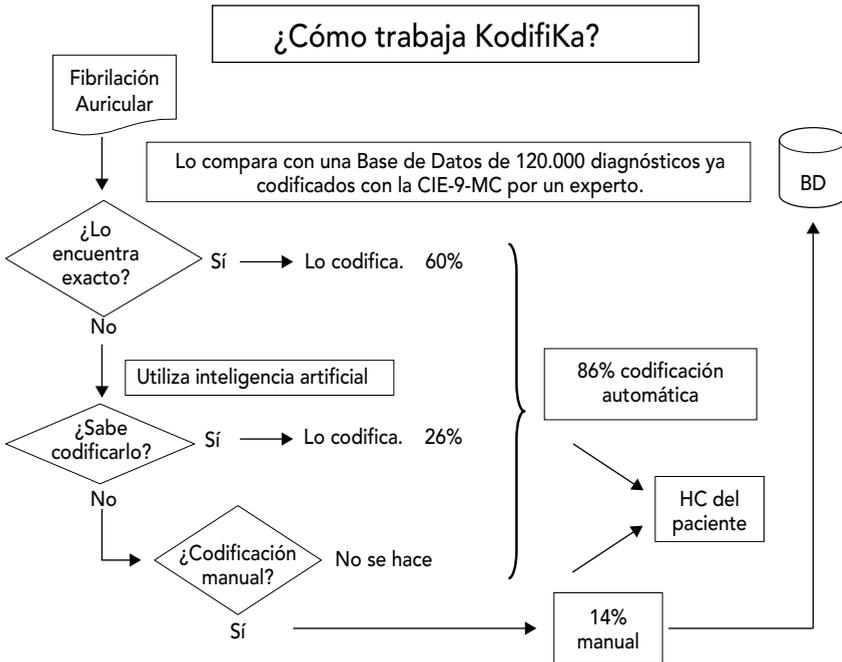
Se obtuvo el programa **KodifiKa**, que actúa como motor de codificación de los diagnósticos introducidos por los médicos en las evoluciones clínicas de cada consulta externa. A partir de la expresión diagnóstica introducida por el médico (fibrilación auricular, por ejemplo) trabaja como se muestra en la Figura 1.

El 26% que codifica con inteligencia artificial se basa en el siguiente algoritmo de selección:

- Selección de palabras relevantes.
- Indexación de las palabras por sus raíces (se agrupan conceptos por sus lemas). Por ejemplo: *Hepatitis*, *Hepático*, se transforman ambas en



Figura 1.



“Hepati”, lo cual disminuye los términos a clasificar.

- Comprensión semántica. En *Hepatitis Aguda a estudio* se asignan pesos a los términos en función de su relevancia. (*Hepatitis* tiene más peso que *Aguda* y esta más que *Estudio*).
- Filtrado y selección del mejor diagnóstico en función de la indexación y de los pesos.

Además, el programa **KodifiKa** tiene un módulo de control de calidad de la base de datos de diagnósticos literales codificados y una interfaz para que un codificador pueda codificar los casos que el programa ha sido incapaz de codificar (el 14% en 2011). En el caso de que el codificador de un hos-

pital realice manualmente la codificación de ese 14%, esos diagnósticos codificados, se añaden a la base de datos de diagnósticos, enriqueciéndola. Si el codificador no realiza la codificación de ese 14%, queda codificado automáticamente el otro 86% de los casos.

KodifiKa se instaló en 2010 en la HCE-Osabide Global implantada en Osakidetza durante 2010 y 2011. En 2011 ha conseguido codificar automáticamente el 86% de las consultas realizadas:

- Número de consultas en HCE-Osabide Global en 2011: 635 111 (100%).
- Número de consultas en HCE-Osabide Global en 2011 codificadas por KodifiKa: 546 186 (86%).



Esta información codificada puede ser explotada por clínicos y gestores y redundará en un mejor conocimiento de la asistencia prestada en el área de Consultas Externas.

nea]. Disponible en <http://sedom.es/wp-content/themes/sedom/pdf/4cbc6553a44afpm-10-2-001.pdf>

## Bibliografía

1. Gosálbez Pastor E, Soto Baca L, Albarracín Álvarez JJ. Aproximación a la codificación de urgencias hospitalarias. En: VII Congreso Nacional de Documentación Médica. Papeles Médicos. 2001;10(2):57 [en lí-
2. Romero Gutiérrez A. Conjunto Mínimo Básico de Datos en Consultas. En: 2.º Foro sobre el Sistema de Información del Sistema Nacional de Salud [en línea]. Disponible en [http://www.msps.es/estadEstudios/estadisticas/sisInfSanSNS/II-ForoSiSNS/foro\\_mesa3\\_ponencia1\\_08\\_es.pdf](http://www.msps.es/estadEstudios/estadisticas/sisInfSanSNS/II-ForoSiSNS/foro_mesa3_ponencia1_08_es.pdf)