

Aplicación de la minería de datos al Sistema Cubano de Farmacovigilancia

Application of data mining to the Cuban Pharmacovigilance System

Omar Calzadilla Fernández de Castro^I; Giset Jiménez López^{II}; Blanca Estela González Delgado^{III}; Jenny Ávila Pérez^{IV}

^I Ingeniero Informático. Especialista en Sistemas Informáticos. Centro para el Desarrollo de la Farmacoepidemiología.

^{II} Especialista de II Grado en Farmacología. Unidad Nacional Coordinadora de Farmacovigilancia.

^{III} Especialista de I Grado en Medicina General Integral. Master en Economía de la Salud. Unidad Nacional Coordinadora de Farmacovigilancia.

^{IV} Técnico en Farmacia. Unidad Nacional Coordinadora de Farmacovigilancia.

RESUMEN

En el año 1999, se crea la Unidad Nacional Coordinadora de Farmacovigilancia, la cual funciona como el órgano técnico-científico que desarrolla la política de vigilancia de seguridad de medicamentos del Ministerio de Salud Pública e integra las actividades de los centros provinciales y de otros programas concertados de farmacovigilancia en un sistema único. Entre sus funciones más relevantes se encuentran la de definir, diseñar y desarrollar los sistemas de información; administrar la base de datos nacional; depurar y validar la información contenida en ella; y realizar y coordinar estudios científicos sobre la seguridad de los medicamentos; así como elaborar informes para las autoridades sanitarias y el Sistema Nacional de Salud. En este sistema existe la necesidad de utilizar herramientas de análisis, por lo que se trazó el objetivo de definir, diseñar y desarrollar los sistemas de tratamiento de la información y administrar la base de datos nacional "VigiBaseCuba". Aplicando una serie de transformaciones, validaciones y la adecuación de la metodología CRISP-DM para la elaboración de proyectos de minería de datos, se conformó la base de datos nacional, en un sistema de gestión de bases de datos relacional con los registros de las notificaciones de sospechas de reacciones adversas a los medicamentos y un proceso de descubrimiento de conocimiento que permite gestionar eficazmente la

seguridad de los medicamentos, así como desarrollar aplicaciones para la visualización de las señales de reacciones adversas y su evolución.

Palabras clave: Farmacovigilancia, minería de datos, data mining, reacciones adversas a medicamentos.

SUMMARY

In 1999, The National Coordinating Unit of Pharmacovigilance was created. It works as a scientific and technical agency that develops the drug safety surveillance policy of the Ministry of Public Health, and integrates the activities of the provincial centres and other pharmacovigilance programs in a unique system. Some of its most important functions are to define, design and develop the information systems; to manage the national database; to deuplicate and validate the information contained in it; and to conduct and coordinate scientific studies on drug safety; as well as to report the health authorities and the National Health System. As in this system it is necessary to use analysis tools, the objective of defining, designing and developing the information treatment systems, and of managing the "VigiBaseCuba" national database, was established. By applying a series of transformations, validations and the adequacy of the CRISP-DM methodology to make data mining projects, it was organized the national database, in a database management system related to the registries of notifications of suspicions of adverse reactions to drugs and to a process of knowledge discovery that allows to efficiently manage drug safety, and to develop applications for the visualization of the signs of adverse reactions and their evolution.

Key words: Pharamcovigilance, data mining, adverse reactions to drugs.

INTRODUCCIÓN

El Centro para el Desarrollo de la Farmacoepidemiología (CDF), creado en junio de 1996 ha realizado la estrategia de la Farmacoepidemiología, con el objetivo de velar por la salud pública mediante la vigilancia de la eficacia y la seguridad de los medicamentos que circulan en el mercado cubano; su estructura descansa en una red nacional y entre sus principales líneas de trabajo se encuentra la farmacovigilancia.

En el año 1999, se crea la Unidad Nacional Coordinadora de Farmacovigilancia (UNCFv), la cual funciona como el órgano técnico-científico que desarrolla la política de vigilancia de seguridad de medicamentos del Ministerio de Salud Pública e integra las actividades de los centros provinciales y de otros programas concertados de farmacovigilancia en un sistema único.¹

Entre sus funciones más relevantes se encuentran la de definir, diseñar y desarrollar los sistemas de información, administrar la base de datos nacional, depurar y validar la información contenida en ella y realizar y coordinar estudios científicos sobre la seguridad de los medicamentos, así como elaborar informes

para las autoridades sanitarias y el Sistema Nacional de Salud (Normas y procedimientos del Sistema Cubano de Farmacovigilancia. Unidad Coordinadora Nacional, Centro para el Desarrollo de la Farmacoepidemiología. MINSAP, 2003).

En el sistema, la tendencia es positiva al aumento de la calidad de las notificaciones y los reportes de sospechas de reacciones adversas a medicamentos (RAM) son revisados y clasificados por primera vez a nivel municipal, ya en el nivel provincial ocurre una segunda revisión de la calidad del dato y la clasificación y se introducen las notificaciones en una base de datos, la cual es enviada con frecuencia quincenal a la UCNFv, donde está la base de datos central, en la cual se realiza una tercera revisión, se homogeneizan los datos y practican los análisis que se incluyen en los informes e investigaciones.

MÉTODOS

Se utilizó para programar la base de datos Microsoft Excel, por ser un software de uso extendido en la red y de fácil manejo, que necesita poco entrenamiento. La base de datos nacional tiene su entrada de datos en una hoja de cálculo configurada adecuadamente con los campos que se reflejan en el modelo de notificación oficialmente establecido, por ejemplo: datos del paciente, del fármaco, de la reacción y de "cierre de información" como otros fármacos, mes de entrada a la base, grupo de edad, nivel de atención, etc. y "clasificación" como causalidad, severidad, grupo farmacológico, sistema de órgano afectado entre otros.

En nuestro sistema existe la necesidad de utilizar herramientas de análisis, por lo que nos trazamos el objetivo de definir, diseñar y desarrollar los sistemas de tratamiento de la información y administrar la base de datos nacional "VigiBaseCuba".

Características de la base de datos nacional

- Contiene las sospechas de RAM reportadas por la red nacional de farmacovigilancia, con promedio entre 7 000 a 10 000 reportes anuales.
- 70 campos o variables de análisis
- Actualización quincenal
- Análisis de información con periodicidad mensual, trimestral, semestral, anual y comparación con períodos anteriores, seguimiento de variables acumuladas, descubrimientos de tendencias ejemplo: obtención de series de casos esperados, componente de información² y estudio, aplicación y desarrollo de nuevos modelos matemáticos (Marín Moran M., Utilización de la teoría de conjuntos aproximados en bases de datos médicas. Tesis de Grado, Licenciatura en Ciencias de la Computación Minería de Datos. Universidad de La Habana, 1997).

RESULTADOS

Aplicando una serie de transformaciones y validaciones, se conformó la base de datos nacional, en un sistema de gestión de bases de datos relacional con los registros de las notificaciones de sospechas de reacciones adversas a los medicamentos y su clasificación que permita gestionar eficazmente la seguridad de los medicamentos, así como desarrollar aplicaciones para la visualización de las señales de reacciones adversas y su evolución, de modo que al detectar una señal se pueda actuar sobre ella, es decir, que además del registro se pudieran visualizar

resultados, dado que la detección de una señal es solo el primer paso de una posterior evaluación crítica.

Todas estas transformaciones de datos, desarrollo y utilización de herramientas de análisis se hicieron sobre base de la metodología CRISP-DM (CRoss-Industry Standard Process for Data Mining) para la elaboración de proyectos de minería de datos ([fig. 1](#)).³

DISCUSIÓN

CRISP-DM, es una metodología para el desarrollo de proyectos de minería de datos que se ha convertido en un estándar de facto. El consorcio CRISP-DM, responsable de esta metodología, está integrado por importantes empresas europeas y estadounidenses que poseen una amplia experiencia en proyectos de análisis de datos relacionados con muy diversos campos de la industria.

La Minería de Datos es una disciplina de la informática, que surge de la necesidad de desarrollar herramientas y modelos de análisis para grandes y complejos volúmenes de datos con el objetivo de extraer conocimiento, y forma parte de un proceso llamado descubrimiento de conocimiento en bases de datos Knowledge Discovery in Databases (KDD, por sus siglas en inglés) definido como: todo el proceso no trivial de identificar patrones en datos que sean válidos, novedosos, potencialmente útiles y humanamente comprensibles", con el objeto de predecir de forma automatizada tendencias y comportamientos y descubrir modelos previamente desconocidos.⁴

El [cuadro](#) muestra la aplicación de la Minería de Datos al proceso de manejo, evaluación y análisis de las notificaciones de sospechas de RAM en el sistema.

En el sistema cubano de farmacovigilancia, la metodología CRISP-DM 1.0 ha sido adaptada a las condiciones concretas de la infraestructura de la red de farmacovigilancia ilustrada de la forma que muestra la [figura 2](#).

En conclusión, el desarrollo de un proceso de descubrimiento de conocimiento en bases de datos con las adaptaciones a las condiciones concretas del sistema cubano de farmacovigilancia y su infraestructura con bajos recursos económicos, ha permitido crear nuestra propia metodología, que va desde procesos no automatizados hasta la creación de una base de datos que permite aplicar los actuales modelos de análisis como es la minería de datos y reportar al Centro de Monitoreo Internacional de Uppsala, Suecia.

REFERENCIAS BIBLIOGRÁFICAS

1. Debesa F, Jiménez G, Pérez J, Ávila J. La estrategia de Farmacoepidemiología en Cuba. Características y papel de la Unidad Coordinadora de Farmacovigilancia en Cuba. Boletín Fármacos. [serie en Internet]. 2001 [citado Ene 2006]. 4(2). Disponible en: <http://www.boletinfarmacos.org/042001/c25.htm>
2. Andrew B. The Use of a Bayesian Confidence Propagation Neural Network in Pharmacovigilance, Umeå, Tesis Doctoral, Uppsala. 2003. [citado Mar 2006]. Disponible en: http://www.diva-portal.org/diva/getDocument?urn_nbn_se_umu_diva-83-1_fulltext.pdf

3. CRISP-DM consortium: [citado Mar 2006]. Disponible en: <http://www.crisp-dm.org/>

4. Fayyad UM, Piatetsky G, Smyth P. From Data Mining to Knowledge Discovery: an overview. Advances Discovery and Data Mining. California: AAAI Press/MIT Press; 1996.

Recibido: 30 de mayo de 2007.

Aprobado: 2 de julio de 2007.

Ing. *Omar Calzadilla Fernández de Castro*. Centro para el Desarrollo de la Farmacoepidemiología. Calle 44 esq. 5ta Ave No. 502. Miramar, Playa, CP 11300, La Habana, Cuba. Correo electrónico: omar@mcdf.sld.cu
Centro para el Desarrollo de la Farmacoepidemiología
Unidad Coordinadora Nacional de Farmacovigilancia