



## Metodología de investigación y lectura crítica de estudios

# Estándares en informática médica: generalidades y aplicaciones

Fernando Suárez-Obando<sup>a,b,\*</sup> y Jhon Camacho Sánchez<sup>b,c</sup>

<sup>a</sup>Instituto de Genética Humana, Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá, D.C., Colombia

<sup>b</sup>Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania, Estados Unidos

<sup>c</sup>Departamento de Epidemiología y Bioestadística, Facultad de Medicina, Pontificia Universidad Javeriana, Bogotá, D.C., Colombia

### INFORMACIÓN DEL ARTÍCULO

#### Historia del artículo:

Recibido el 5 de noviembre de 2012

Aceptado el 23 de enero de 2013

#### Palabras clave:

Informática médica

Aplicaciones de informática médica

Ciencia de la información

#### Keywords:

Medical informatics

Medical informatics applications

Information science

### R E S U M E N

La aplicación de la informática a la práctica médica ha permitido el desarrollo de novedosas formas de comunicación en la atención de la salud. La optimización de los procesos comunicativos se alcanza con el uso de estándares que armonicen el intercambio de información y provean un lenguaje común para todos los agentes involucrados. En el presente artículo, se describe el concepto de estándar aplicado a la informática médica y su importancia en el desarrollo de diversas aplicaciones tales como la representación computacional del conocimiento médico, la codificación diagnóstica, la búsqueda de literatura médica y la integración de las ciencias biológicas a las aplicaciones clínicas.

© 2012 Asociación Colombiana de Psiquiatría. Publicado por Elsevier España, S.L.

Todos los derechos reservados.

### Standards in Medical Informatics: Fundamentals and Applications

#### A B S T R A C T

The use of computers in medical practice has enabled novel forms of communication to be developed in health care. The optimization of communication processes is achieved through the use of standards to harmonize the exchange of information and provide a common language for all those involved. This article describes the concept of a standard applied to medical informatics and its importance in the development of various applications, such as computational representation of medical knowledge, disease classification and coding systems, medical literature searches and integration of biological and clinical sciences.

© 2012 Asociación Colombiana de Psiquiatría. Published by Elsevier España, S.L.

All rights reserved.

\*Autor para correspondencia.

Correo electrónico: fernando.suarez@javeriana.edu.co (F. Suárez-Obando).

## Introducción

La informática médica (IM) es la disciplina que aplica las ciencias de la información al contexto de la medicina<sup>1</sup>. Su principal objetivo es optimizar la adquisición, el almacenamiento, la recuperación y el uso de la información en salud. Sus herramientas incluyen computadores, *software* especializado, sistemas de soporte a decisiones clínicas, terminología médica formal y sistemas de información y comunicación<sup>2</sup>, su ámbito se extiende desde el ambiente clínico hasta diversos aspectos de la atención de la salud, como odontología, enfermería, farmacia, salud pública e investigación médica<sup>3</sup>.

Entre las tareas fundamentales de la IM, se encuentra mejorar la comunicación entre las diversas áreas involucradas en la atención de la salud. Para avanzar en este propósito, la informática médica propone el desarrollo y la aplicación de estándares que armonicen la comunicación y provean un lenguaje común para todos los involucrados<sup>4</sup>.

Se puede apreciar las ventajas de la aplicación de estándares en ambientes en los que la ambigüedad del lenguaje genera ineficiencias o interfiere con la efectividad. Tal puede ser el caso de la distribución de tareas en un hospital, donde diferentes unidades atienden a los pacientes sin que exista una apropiada comunicación entre ellas. En tales circunstancias, el uso de sistemas de codificación heterogéneos puede desembocar en fragmentación de la atención. El ambiente hospitalario está dividido en especialidades clínicas que frecuentemente tratan a los pacientes sin considerar el manejo de otros equipos médicos, mientras que los procesos de cobro y facturación se valen de abstracciones fragmentadas de la historia clínica para generar listas de códigos de diagnóstico que no necesariamente se relacionan con la información médica del paciente<sup>5,6</sup>.

Un estándar comprende un grupo de reglas y definiciones que especifican cómo llevar a cabo un proceso<sup>2</sup>. Los estándares son útiles para la IM porque ayudan a armonizar los métodos de gestión y análisis de información. Estos métodos se basan principalmente en el empleo de un lenguaje común y el uso de terminología biomédica específica. El uso de un lenguaje estándar permite el intercambio de información entre sistemas de historias clínicas electrónicas (HCE), favorece la extracción eficiente de información de las bases de datos, contribuye al desarrollo de sistemas de soporte diagnóstico, habilita la minería de datos y facilita la evaluación estadística<sup>7</sup>. La implementación de un lenguaje estándar en el ámbito clínico es fundamental para mantener una comunicación efectiva entre las diversas unidades que atienden al paciente, el desarrollo adecuado de la investigación básica y clínica y el uso eficiente de la medicina basada en la evidencia (MBE).

Con el propósito de destacar la importancia de la comunicación armonizada, se procede a describir la noción de estándar en el campo de la IM, a partir de la definición de sus conceptos básicos y la exposición de ejemplos que demuestran el uso de las terminologías en sistemas de HCE, sistemas inteligentes aplicados al diagnóstico clínico, búsqueda de literatura científica especializada e intercambio de datos. Los estándares aquí expuestos hacen parte de sistemas informáticos utilizados en centros de atención clínica para dar

apoyo a diversos aspectos de la atención, como la HCE y la digitalización de imágenes radiológicas.

## La necesidad de estándares en informática médica

Las fuentes de información que respaldan el ejercicio médico actual se pueden clasificar en cuatro categorías: repositorios de datos epidemiológicos y de salud pública, repositorios de datos bioinformáticos (p. ej., datos genómicos, proteómicos, metabolómicos, etc.), bases de datos de literatura científica y sistemas de registro de información clínica. La disponibilidad y la calidad de estas fuentes de información tienen un profundo impacto tanto en los medios y los métodos de atención clínica como en los diseños de la investigación biomédica.

Los repositorios y las bases de datos se valen de códigos y terminologías para describir conceptos clínicos en un lenguaje entendible para un computador. Sin embargo, el uso de estas terminologías sin criterios adecuados puede generar imprecisiones; por ejemplo, si se utilizan los códigos de la Clasificación Internacional de Enfermedades (CIE-10) para el registro de vigilancia de tuberculosis en salud pública, se puede especificar incluso el nivel histopatológico de la enfermedad: A15 es el código que define la «Tuberculosis respiratoria, confirmada bacteriológica e histológicamente», mientras que el código A15.0, por su parte, define la enfermedad como «Tuberculosis del pulmón, confirmada por hallazgo microscópico del bacilo tuberculoso en esputo, con o sin cultivo». El uso de uno u otro código marcaría una diferencia en el estado diagnóstico de la enfermedad de acuerdo con criterios de alta complejidad técnica, como el cultivo del bacilo. Este ejemplo muestra cómo la especificidad de los códigos permite una mayor exactitud en la representación del diagnóstico, pero su empleo sin criterios adecuados puede generar imprecisiones.

Un fenómeno relacionado se da en la descripción de conceptos biológicos: por ejemplo, los fenotipos humanos se definen de forma heterogénea en diversas bases de datos médicas<sup>8</sup> y las estructuras biológicas subcelulares tienen notaciones diferentes según la institución donde se desarrollen los métodos de búsqueda y predicción<sup>9,10</sup>. Los diagnósticos registrados en la historia clínica difieren entre los médicos tratantes, y su interpretación contrasta con la interpretación del paciente aun cuando todos estos conceptos se refieren a distintos aspectos de la misma enfermedad. En ausencia de procesos de armonización, no existe un mecanismo unificado para correlacionarlos<sup>11</sup>.

A pesar de la diversidad de los datos clínicos y los conceptos biomédicos, estos tienen aspectos técnicos y biológicos en común que facilitan los procesos de armonización. Los aspectos técnicos incluyen: métodos de codificación para conceptos no numéricos (p. ej., secuencias de genes y proteínas, nombres de órganos y sus relaciones en sistemas fisiológicos, descripción de hallazgos clínicos, nombres de enfermedades, etc.), métodos de organización de los datos para ubicarlos, extraerlos e interpretarlos fácilmente y métodos de representación jerárquica o agrupamiento para datos que, *a priori*, tienen relación entre sí (p. ej., medición de niveles de expresión genética o un set de análisis sanguíneos relacionados).

Los aspectos biológicos incluyen: métodos de representación de fenómenos de nivel superior (fenotipos, enfermedades y terapias) en relación con fenómenos de nivel inferior (genotipos, procesos celulares y estructuras corporales), así como la asignación sistemática y coherente de nombres y clasificaciones a todas las entidades, procesos y relaciones<sup>12</sup>.

Apoyándose en estos aspectos técnicos y biológicos, la aplicación de estándares (incluidas la codificación, la organización, la representación y la asignación coherente de conceptos a los datos biomédicos) facilita la comprensión de la complejidad de los diversos niveles de atención médica. El primer paso en este propósito es definir las formas básicas del estándar, a través del análisis de sus conceptos fundamentales.

---

### Vocabularios, códigos y ontologías

La codificación de la información médica es una función básica de los sistemas de información clínicos. Los estándares utilizados para esta codificación cumplen dos funciones principales: en primer lugar, ahorran tiempo en el desarrollo de sistemas; por ejemplo, si una aplicación tiene como propósito la compilación de listas de problemas de los pacientes, los desarrolladores no tendrán que definir el compendio de todos los posibles problemas, y en cambio podrían utilizar una terminología estándar que contenga dicha definición. En segundo lugar, el uso de terminología estándar facilitaría el intercambio de información entre sistemas: si una base de datos centralizada recibe datos de diversas fuentes, el trabajo se simplifica si cada fuente utiliza el mismo esquema de codificación<sup>2</sup>.

El estándar a su vez define el uso adecuado de los términos. Esta definición contribuye a limitar el uso indistinto de las palabras, lo que es especialmente importante cuando las diferencias conceptuales son relevantes en los procesos de comunicación.

Una jerarquía, no exhaustiva, de los componentes necesarios para definir un estándar en informática médica, puede comenzar con la definición general de vocabulario o conjunto de palabras de un idioma, noción que a su vez incluye la terminología, que constituye el conjunto de términos propios de una profesión o ciencia. Un vocabulario controlado es un subconjunto de términos cuidadosamente seleccionados, en el que uno solo de los términos describe cada concepto de un dominio y, a su vez, cada término describe un solo concepto, lo que evita la redundancia y la ambigüedad.

En otras palabras, una terminología es un conjunto de términos que se usan para nombrar cosas (todo lo que tiene entidad) en un campo propio del saber; cuando la terminología se selecciona con criterios que eviten la redundancia y la ambigüedad, se constituye en un vocabulario controlado; por ejemplo, el tesoro es un vocabulario controlado empleado en indexación. Por su parte, una taxonomía es un vocabulario controlado que existe como una ordenación jerarquizada y sistemática.

Un código es un sistema de signos y reglas que permite formular y comprender un mensaje (p. ej., el código genético, el código Morse, el código de barras, etc.). La abstracción es el proceso en que se examinan los datos, y se rotulan utili-

zando ítems de una terminología. Por ejemplo, se realiza una abstracción cuando los registros de la hospitalización de un paciente se rotulan con el código CIE-10 A15: «Tuberculosis respiratoria, confirmada bacteriológica e histológicamente». Por otra parte, la representación es el proceso a través del cual se codifica la mayor cantidad de detalles. En la representación del paciente con tuberculosis, se incluirían códigos para cada uno de los hallazgos físicos, resultados de laboratorio y medicamentos<sup>2</sup>.

Un nivel de ordenación conceptual más complejo es la ontología. Una ontología constituye la organización formal, rigurosa y exhaustiva de un conocimiento en la forma de un conjunto de conceptos y sus relaciones dentro de un dominio<sup>13</sup>. En la organización de una ontología se incluyen códigos, vocabularios controlados y taxonomías. En términos generales, una ontología es la descripción de la estructura de los conceptos de un dominio. La ontología confiere significado a los datos en bruto de una base de datos; el significado de las relaciones entre los datos puede ser inferido por un computador o por un ser humano; sin embargo, la creciente complejidad y la cantidad de datos hacen casi obligatorio que los procesos de inferencia basados en ontologías sean llevados a cabo por *software* especializado<sup>14</sup>.

La importancia de estos conceptos se puede apreciar analizando casos específicos de terminologías y estándares de uso común en IM. Su desarrollo y su aplicación obedecen a la necesidad de organizar la información y promover su uso eficiente en el entorno de la atención al paciente. A continuación se describen ejemplos específicos de la aplicación de terminologías y estándares en el ámbito clínico.

---

### Casos específicos

La descripción que se expone a continuación hace referencia a la importancia que cada terminología tiene en la IM la motivación para desarrollarla. Las terminologías se actualizan periódicamente, por lo que es necesario consultar habitualmente versiones actualizadas o nuevas terminologías. Se exponen ejemplos de estándares aplicados al diagnóstico clínico, la búsqueda de literatura científica, la correlación genotipo y fenotipo y el intercambio de datos.

#### Diagnóstico clínico

*Clasificación Internacional de enfermedades (CIE)* o *International Classification of Diseases (ICD)*

La CIE es la más conocida de las terminologías médicas, se publicó por primera vez en 1893 y desde entonces ha sido revisada aproximadamente cada 10 años. Su décima edición (CIE-10) se publicó en 1992. La CIE es la clasificación de diagnósticos de uso obligatorio en los reportes epidemiológicos y en las estadísticas de mortalidad de la Organización Mundial de la Salud (OMS)<sup>15</sup>, su uso se extiende desde ámbito clínico hasta las actividades de administración y facturación en los hospitales<sup>16</sup>.

La CIE consiste de un núcleo de códigos de tres dígitos, que es el mínimo de dígitos para reportar estadísticas de mor-

talidad a la OMS. La CIE utiliza un código alfanumérico con una letra en la primera posición y números en la segunda y la tercera posición. El cuarto carácter se ubica después de un punto, y los códigos posibles se extienden de A00.0 a Z99.9.

La cuarta cifra provee un nivel adicional de detalle; usualmente es un número de 0 a 7 y se utiliza para definir formas más específicas de los códigos del núcleo, 8 define el grupo «otros» y 9 es «inespecífico». Los términos están organizados en una estructura jerárquica basada en los números del código. Por ejemplo, los códigos A00-B99 comprenden: Ciertas enfermedades infecciosas y parasitarias. A00 específica para el cólera. A00.0 es el código para el cólera debido a *Vibrio cholerae*, biovariedad cólera o cólera clásico. El código A00.1 especifica: cólera debido a biovariedad el Tor o cólera el Tor. Finalmente el código A00.9 define: Cólera, inespecífico.

La CIE-10 se encuentra en proceso de implementación en los países miembros de la OMS desde 1994. Se puede descargar de la página de la OMS (<http://www.who.int/classifications/icd/en/#>) y se encuentra traducida al español. La OMS también provee herramientas de entrenamiento *online* tanto para autoaprendizaje como para enseñanza en un salón de clase (<http://apps.who.int/classifications/apps/icd/icd10training/>).

¿Cuándo se utiliza? La CIE-10 se puede utilizar en un sistema de HCE para armonizar la información clínica y administrativa, así como el reporte de salud pública. En tal sistema, el médico selecciona el diagnóstico de una lista basada en CIE-10, y el sistema de facturación utiliza el mismo estándar para la abstracción de los diagnósticos y la representación de los procedimientos. Finalmente, el sistema de reporte utiliza la CIE-10 para construir las consultas a la base de datos y definir diversos indicadores.

#### Grupos de diagnóstico relacionados (GDR) o Diagnosis Related Groups (DRG)

Los GDR se desarrollaron en la Universidad de Yale para el sistema Medicare (programa de seguro social, administrado por el gobierno de Estados Unidos, que ofrece cobertura de salud a personas mayores de 65 años o que cumplan otros criterios especiales). Esta terminología constituye una doble abstracción, ya que los códigos GDR agrupan códigos de la CIE, que en su aplicación son abstracciones de los datos de la historia clínica. El propósito de los GDR es proveer un número reducido de códigos que permita clasificar las hospitalizaciones según la gravedad de la enfermedad<sup>17</sup>. Los GDR agrupan a los pacientes según su complejidad o gravedad, lo que facilita la estimación del tiempo de estancia y los recursos que consumirán.

¿Cuándo se utiliza? Aunque la aplicación de los GDR es principalmente administrativa, los resultados de su aplicación se reflejan en la práctica clínica. Por ejemplo, la agrupación de los pacientes de un hospital según su complejidad y su consumo de recursos permite a la institución comparar servicios similares con los de otras instituciones, determinar el comportamiento real del costo por cada paciente y comparar el costo específico de cada morbilidad y el costo por protocolo de atención clínica. Este estándar codificado permite a los equipos administrativos y clínicos analizar los casos clínicos usando un mismo lenguaje, lo que permite la comparación

con otros centros de atención. El análisis de los casos con base en su complejidad permite asignar recursos según la relación de costo y beneficio<sup>18</sup>.

#### Systematized Nomenclature of Medicine-Clinical Terms (SNOMED-CT)

A diferencia de los sistemas de codificación CIE y GDR, SNOMED-CT es una recopilación de terminología médica organizada sistemáticamente para el procesamiento computarizado. El propósito de esta nomenclatura es facilitar el desarrollo de sistemas de HCE que permitan el intercambio de información entre diferentes sistemas<sup>19</sup>.

En SNOMED-CT, los conceptos están organizados en una jerarquía del tipo «es\_un(a)» (en inglés, *IS-A hierarchy*), por ejemplo: «Neumonía bacteriana» es\_una «enfermedad del pulmón». Un concepto también puede tener varias relaciones es\_un(a); por ejemplo, «Neumonía bacteriana» es\_una «infección bacteriana» y también «Neumonía bacteriana» es\_una «neumonía infecciosa».

SNOMED-CT también permite agregar términos para enriquecer la descripción del concepto; por ejemplo, se puede añadir términos que describan el agente causal («agente bacteriano») o el proceso patológico («inflamación»).

También se puede incluir atributos como «inicio» o «gravedad», lo que permitiría calificar la neumonía como «de inicio agudo» y asignarle un grado de gravedad. El uso de diversos términos en combinación con el concepto principal («Neumonía bacteriana») se conoce como proceso de poscoordinación<sup>20</sup>. En un lenguaje precoordinado, los términos se combinan en el momento de la descripción, mientras que en un lenguaje poscoordinado los términos se coordinan en el momento de la recuperación de un término.

¿Cuándo se utiliza? Un investigador clínico que busque conocer la prevalencia de una enfermedad crónica y sus complicaciones (p. ej., diabetes mellitus) en varios centros hospitalarios necesitaría un marco de referencia que le indique las diversas formas en que puede codificarse la enfermedad: SNOMED-CT proporciona dicho marco.

Si los códigos diagnósticos de los centros de atención tienen como referencia las relaciones conceptuales del sistema SNOMED-CT, los datos de prevalencia se pueden agregar de manera exacta y precisa. La capacidad descriptiva del sistema también proveería otras ventajas; por ejemplo, el investigador podría partir de la generalidad —«diabetes mellitus»— y gradualmente alcanzar mayor especificidad diagnóstica usando combinaciones de términos, como «diabetes mellitus asociada con cetoacidosis» o «diabetes mellitus asociada con manifestaciones neurológicas». El uso del estándar también permitiría el uso de métodos de búsqueda computarizados más precisos y rápidos, mientras que su ausencia dificultaría la detección de las diversas formas diagnósticas de la enfermedad<sup>21</sup>.

#### Unified Medical Language System (UMLS)

Con el mismo propósito que caracteriza el SNOMED-CT, el UMLS se diseñó para la implementación y el desarrollo de la HCE, buscando facilitar el desarrollo de sistemas que «comprendan» el significado del lenguaje médico<sup>22</sup>. La principal característica del UMLS es la integración de tesauros y ontologías desarrolladas independientemente, y su principal herra-

mienta es el metatesauro: una base de datos multilingüaje que contiene información sobre conceptos biomédicos, sus nombres y sus relaciones.

El metatesauro está construido a partir de las versiones electrónicas de diferentes tesauros médicos, clasificaciones, códigos y vocabularios controlados, utilizados en el cuidado de pacientes, gerencia hospitalaria, estadísticas de salud pública, indexación de literatura científica, e investigación básica y clínica<sup>22</sup> y agrupa: códigos CPT (*Current Procedural Terminology*) o códigos de procedimientos médicos desarrollados por AMA (*American Medical Association*), CIE-10, LOINC (*Logical Observation Identifiers Names and Codes*, o código universal para la identificación de observaciones clínicas y de laboratorio), MeSH (*Medical Subject Headings*), RxNorm (nombre normalizados de medicamentos) y SNOMED-CT.

¿Cuándo se utiliza? Una facultad de medicina interesada en identificar las dificultades que sus residentes encuentran al realizar entrevistas clínicas a los pacientes puede desarrollar una investigación en la que se analicen grabaciones de entrevistas clínicas.

Tal análisis puede identificar en el lenguaje empleado por los residentes conceptos clave asociados al diagnóstico y el tratamiento, así como datos específicos del paciente y su familia. Un reto en la realización de este trabajo es que los residentes pueden usar diferentes términos para referirse al mismo concepto. Para enfrentar este reto, los investigadores pueden emplear el método de integración semántica, que consiste en la interrelación de información proveniente de diversas fuentes<sup>23</sup>. En el caso de la investigación de la facultad, la integración semántica emplea dos fuentes de información: las grabaciones y el metatesauro del UMLS. Al identificar en las grabaciones los conceptos definidos en el UMLS, el análisis puede concentrarse en aspectos tales como la comprensión del concepto por el residente, sus habilidades para la comunicación efectiva y su capacidad para llevar a cabo el interrogatorio<sup>24</sup>.

### Búsqueda de literatura científica

Medical Subject Headings (MeSH)-Descriptor en Ciencias de la Salud (DeCS)

MeSH es el tesoro utilizado por la *National Library of Medicine* (NLM) para indexar las referencias bibliográficas de PubMed<sup>25</sup>. Por su parte, el DeCS lo estableció la Biblioteca Regional de Medicina (BIREME) para la indexación de documentos disponibles en la Biblioteca Virtual en Salud (BVS). El DeCS se desarrolló a partir del MeSH con el objetivo de utilizar terminología común para la búsqueda de literatura científica en tres idiomas (español, portugués e inglés)<sup>26</sup>.

Los términos en MeSH y en DeCS están organizados tanto alfabéticamente como en estructura jerárquica. Estas estructuras parten de términos generales de amplio significado (p. ej., en MeSH, *Anatomy* y en DeCS, Anatomía), y conforme se desciende en la jerarquía, se encuentran términos cada vez más específicos (p. ej., en MeSH, *Sense Organs* y en DeCS, Órganos de los sentidos; MeSH, *Eye* y DeCS, Ojo).

Cada término cuenta con un encabezado general (*Heading* en MeSH y Clase en DeCS) y subencabezados más especí-

ficos (p. ej., en MeSH el término *Eye* tiene un *heading Eye* y *sub-headings: secretion, microbiology, metabolism*, etc. En DeCS este término correspondería a clase Ojo, subclases: secreción, microbiología, metabolismo, etc.).

MeSH y DeCS fueron creados con el propósito de facilitar el acceso a la literatura científica en el campo de la biomedicina. Además de los términos originales del MeSH, DeCS incluye conceptos de salud pública, homeopatía y vigilancia sanitaria. DeCS también provee los términos en portugués y español para UMLS.

¿Cuándo se utiliza? Un médico que desee estudiar la evidencia científica acerca de las características operativas de una prueba diagnóstica puede utilizar los términos MeSH y DeCS para buscar artículos relevantes en las bases de datos correspondientes (PubMed y BIREME). Por otro lado, el hecho de que las palabras clave incluidas en un artículo se correspondan con términos MeSH o DeCS sirve de guía para la indexación y facilita la búsqueda de literatura.

### Correlación entre Genotipo y Fenotipo

Gene Ontology (GO)

*Gen Ontology* fue desarrollada por el Consorcio GO, y tiene por objetivo proveer un vocabulario estructurado y controlado que describa la localización, los productos y las funciones de los genes de diferentes especies<sup>27</sup>. Originalmente GO incluía las bases de datos genómicos de ratón (*Mouse Genome Database [MGD]*), mosca (*Drosophila [FlyBase]*) y levadura (*Saccharomyces Genome Database [SGD]*). Actualmente GO agrupa información de 16 fuentes: la comunidad para la anotación de *Escherichia coli* (*the Community annotation for Escherichia coli [EcoliWiki]*) y la base de datos genómicos de *Rattus norvegicus* (*Rat Genome Database [RGD]*). GO también forma parte del metatesauro del UMLS.

La descripción de un producto génico se denomina anotación. El uso más relevante de GO es la predicción de las funciones de un gen basada en los patrones de anotaciones. Si varias anotaciones para dos atributos tienden a aparecer juntas en la base de datos, un gen que tenga uno de los atributos probablemente también posea el otro. Las predicciones funcionales pueden realizarse aplicando conocimiento previo para inferir la función de una nueva entidad, sea esta un gen o una proteína<sup>28</sup>.

GO está formado por tres ontologías, cada una de las cuales organiza y describe los productos génicos en diferente manera: Función molecular (FM), que describe la actividad bioquímica del producto génico; Proceso biológico (PB), que describe a qué evento biológico contribuye el producto génico, y Componente celular (CC), que describe la localización celular donde el producto génico ejerce su actividad biológica<sup>29</sup> (fig.).

¿Cuándo se utiliza? Un investigador en busca de objetivos celulares para los mecanismos de acción de un antibiótico necesita conocer los productos génicos involucrados en la síntesis proteica bacteriana, poniendo énfasis en las diferencias entre las secuencias de los genes o las estructuras proteicas de las células humanas y las bacterianas. Si una base de datos describe que dichas moléculas están involucradas en la «traducción proteica» mientras que otras bases de datos utili-

zan el término «Síntesis proteica», encontrar la equivalencia entre los términos es muy difícil aun para una computadora<sup>30</sup>. GO provee al investigador una descripción coherente de los productos génicos utilizando un vocabulario controlado que describe con exactitud los procesos biológicos con que están relacionados. GO también ayuda al investigador a identificar las similitudes entre los procesos de ambas especies y las diferencias específicas requeridas para que el efecto terapéutico se ejerza contra la bacteria, para minimizar el daño en las células humanas.

#### Foundational Model of Anatomy Ontology (FMA)

El FMA es un modelo formal de la anatomía humana que detalla sus estructuras hasta las escalas celular y subcelular, con sus propiedades y relaciones<sup>31</sup>. El FMA comprende cuatro componentes que representan diferentes aspectos de la anatomía humana. El primer componente es la anatomía taxonómica (*Anatomical Taxonomy [AT]*), que organiza jerárquicamente los componentes anatómicos. A través del segundo componente, la abstracción estructural anatómica (*Anatomical Structural Abstraction [ASA]*), el FMA expresa las relaciones entre los componentes anatómicos empleando expresiones como: *partes constituyentes*, *contenido\_en* y *conectividad*.

El FMA representa la anatomía en forma canónica, es decir, las relaciones expresadas son ciertas para la estructura anatómica humana, y no como un conjunto de atributos de un ser humano en particular. El tercer componente, la abstracción de la transformación anatómica (*Anatomical Transformation Abstraction [ATA]*), permite al FMA representar estructuras que cambian con el tiempo (p. ej., desarrollo embrionario). Por último, el FMA provee un metalenguaje consistente en un conjunto de reglas y principios (axiomas anatómicos) que los demás componentes de la estructura deben cumplir<sup>32</sup>.

¿Cuándo se utiliza? El FMA facilita el desarrollo de programas para analizar imágenes obtenidas por resonancia magnética. La ontología permite la anotación de las imágenes (proceso por el cual se asignan etiquetas o palabras clave a una imagen digital) según las relaciones anatómicas utilizando vocabulario estandarizado. El FMA provee a los desarrolladores un marco de referencia para analizar las imágenes respetando la realidad anatómica y proveyendo información más allá de la definición conceptual<sup>33</sup>. Por ejemplo, una parte de la imagen de un riñón puede anotarse como «segmento anterosuperior del riñón derecho», y usando la FMA se puede añadir relaciones como: «hace parte del segmento superior del riñón», «hace parte del riñón derecho», «hacer parte de sistema urinario», «tiene suplencia arterial de la rama anterosuperior de la arteria renal derecha», etc. La ontología permite representar las relaciones entre las anotaciones en términos de relaciones entre las entidades anatómicas (p. ej., suplencia arterial, drenaje venoso, drenaje linfático, partes constituyentes, relación espacial, etc.)<sup>34</sup>.

#### Intercambio de datos

##### Health Level Seven (HL7)

HL7 es una organización sin ánimo de lucro dedicada a proveer estándares para intercambio, integración y recuperación

de información electrónica relacionada con la provisión y la evaluación de servicios de salud<sup>35</sup>.

Usualmente se requiere más de un sistema de información para sustentar los procesos clínicos y administrativos de un hospital. La calidad del intercambio de información entre esos sistemas tiene un impacto considerable en la efectividad y la eficiencia de los procesos. HL7 provee una serie de estándares flexibles, guías y metodologías a través de las cuales varios sistemas hospitalarios pueden comunicarse entre sí. HL7 proporciona un lenguaje común para intercambiar mensajes entre los sistemas de información e interpretar su significado considerando el contexto y el contenido de los datos<sup>36</sup>.

¿Cuándo se utiliza? En el caso del cribado neonatal de enfermedades metabólicas, se requiere la participación de diversos grupos de atención clínica: médicos tratantes, personal de enfermería, laboratorio clínico y farmacia, además de las áreas responsables del reporte epidemiológico y de salud pública. Debido a la naturaleza de estas enfermedades, la comunicación entre estos grupos debe ser coordinada, rápida y efectiva. Una forma de mejorar la comunicación entre los sistemas involucrados es implementar mensajes que usen el estándar HL7 y contengan conjuntos de códigos derivados de SNOMED-CT junto con el resultado de las pruebas diagnósticas realizadas expresado unidades de medición estandarizadas. De esta forma, todos los grupos involucrados reciben un mensaje estandarizado con la información necesaria para actuar rápidamente cuando detecten un caso<sup>37</sup>. En ausencia de un estándar para el intercambio de información, las diversas interpretaciones de los datos pueden generar errores.

---

#### El principal objetivo: el paciente

La aplicación de estándares de información a los sistemas de atención en salud han recibido el impulso de iniciativas gubernamentales y privadas que tienen por objetivo disminuir los errores médicos, evitar o detectar rápidamente eventos adversos e incrementar la seguridad del paciente en el medio hospitalario<sup>38</sup>. Las recomendaciones sobre la implementación de estos estándares se ha basado principalmente en estudios relacionados con errores médicos<sup>39</sup>, la optimización del uso de recursos según perfiles epidemiológicos de morbilidad y mortalidad y el mejoramiento de la calidad del cuidado<sup>40</sup>.

En Colombia, el Ministerio de la Protección Social expidió en 2008 la Política de Seguridad del Paciente. Dicha política indica que, para reducir los errores en el proceso de medicación, las órdenes médicas deben realizarse por computador y se debe implementar la HCE y establecer mecanismos en los sistemas informáticos que generen alarmas sobre eventos adversos, alergias, interacciones o indicios de atención insegura. La política también resalta la importancia de definir planes de contingencia para mitigar los riesgos relacionados con el uso conjunto de HCE e historia física<sup>41</sup>.

La implementación de estas recomendaciones obliga a las instituciones a implementar estándares de los sistemas informáticos con el objetivo de mejorar la atención al paciente y garantizar su seguridad durante la atención clínica. Sin embargo, es importante anotar que ni estándares ni computa-

doras son garantía suficiente para alcanzar este objetivo. Los estándares no están aceptados universalmente, cada estándar tiene sus propias limitaciones y su correcta implementación depende de los recursos invertidos en el desarrollo del *software* y su implementación en el hospital<sup>42</sup>.

El esfuerzo económico y los retos tecnológicos también son limitaciones en el proceso de alcanzar la atención óptima, y aún más importante, se debe tener en cuenta que implementar nuevas tecnologías genera nuevos problemas en la práctica, las enfermedades del progreso médico han estado presentes desde la irrupción de nuevas tecnologías en medicina.

El uso de estándares en IM tiene el potencial de generar desarrollos notables en la atención clínica, que permitan que la historia clínica del paciente sea consultada por sus médicos tratantes aun cuando este se desplace entre instituciones integrando los resultados históricos de pruebas de laboratorio, alertando al médico sobre potenciales eventos adversos e interacciones entre medicamentos, poniendo a disposición del médico la literatura científica que respalde la mejor atención de cada caso y apoyando el desarrollo de nuevas formas de atención clínica como la telemedicina y la prescripción de medicamentos *online*. Estos desarrollos deben tener como propósito mejorar la atención del paciente, velar por su seguridad y mantener la confidencialidad de su información.

## Financiación

Este trabajo fue posible gracias a la colaboración entre la Pontificia Universidad Javeriana y The Department of Biomedical Informatics, University of Pittsburgh. NLM grant HHSN276201000030C. Parcialmente patrocinado por Fogarty International Center NIH Grant No. 1 D43 TW008443-01.

## Conflicto de intereses

Ninguno.

## BIBLIOGRAFÍA

- Bernstam EV, Smith JW, Johnson TR. What is biomedical informatics? *J Biomed Inform.* 2010;43:104-10.
- Shortliffe E, editor. *Biomedical informatics*. 3.<sup>a</sup> ed. New York: Springer; 2006. p. 412-3.
- Patel VL, Yoskowitz NA, Arocha JF, Shortliffe EH. Cognitive and learning sciences in biomedical and health instructional design: A review with lessons for biomedical informatics education. *J Biomed Inform.* 2009;42:176-97.
- Edwards A, Hollin I, Barry J, Kachnowski S. Barriers to cross—institutional health information exchange: a literature review. *J Healthc Inf Manag.* 2010;24:22-34.
- Jollis JG, Ancukiewicz M, DeLong ER, Pryor DB, Muhlbaier LH, Mark DB. Discordance of databases designed for claims payment versus clinical information systems. Implications for outcomes research. *Ann Intern Med.* 1993;119:844-50.
- Pine M, Jordan HS, Elixhauser A, Fry DE, Hoaglin DC, Jones B, et al. Modifying ICD-9-CM coding of secondary diagnoses to improve risk-adjustment of inpatient mortality rates. *Med Decis Making.* 2009;29:69-81.
- Ingenerf J, Poppl SJ. Biomedical vocabularies—the demand for differentiation. *Stud Health Technol Inform.* 2007;129(Pt 1):610-5.
- Robinson P, Mundlos S. The human phenotype ontology. *Clin Genet.* 2010;77:525-34.
- Hagen MS, Lee EK. BIOSPIDA: A relational database translator for NCBI. *AMIA Annu Symp Proc.* 2010;2010:422-6.
- Shen HB, Yang J, Chou KC. Euk-PLoc: an ensemble classifier for large-scale eukaryotic protein subcellular location prediction. *Amino Acids.* 2007;33:57-67.
- Patrick TB, Monga HK, Sievert ME, Houston Hall J, Longo D. Evaluation of controlled vocabulary resources for development of a consumer entry vocabulary for diabetes. *J Med Internet Res.* 2001;3:E24.
- Rosse C, Kumar A, Mejino JL Jr, Cook DL, Detwiler LT, Smith B. A strategy for improving and integrating biomedical ontologies. *AMIA Annu Symp Proc.* 2005:639-43.
- Gruber TR. A translation approach to portable ontology specifications. *Knowl Acquis.* 1993;5:199-220.
- Baclawski K, Niu T. Introduction to ontologies. En: *Ontologies for bioinformatics*. Cambridge: MIT Press books; 2006. p. 5-10.
- WHO. *Classifications: International Classification of Diseases (ICD)*. 2011 [citado 20 Mar 2012]. Disponible en: <http://www.who.int/classifications/icd/en/#>
- Rogoski RR. If you're not coding, you're not billing. The world of codes and HIM alike represents a fast-moving landscape. *Health Manag Technol.* 2004;25:14,16,18.
- Müller ML, Forschner A, Wenke A, Luger TA, Rompel R, Roeder N, et al. G-DRG Version 2009: new developments. *J Dtsch Dermatol Ges.* 2009;7:318-27.
- Elkin PL, Liebow M, Bauer BA, Chaliki S, Wahner-Roedler D, Bundrick J, et al. The introduction of a diagnostic decision support system (DXplain) into the workflow of a teaching hospital service can decrease the cost of service for diagnostically challenging Diagnostic Related Groups (DRGs). *Int J Med Inform.* 2010;79:772-7.
- Roma-Ferri MT, Palomar M. [Analysis of health terminologies for use as ontologies in healthcare information systems]. *Gac Sanit.* 2008;22:421-33.
- Lee D, Cornet R, Lau F. Implications of SNOMED CT versioning. *Int J Med Inform.* 2011;80:442-53.
- Wasserman H, Wang J. An applied evaluation of SNOMED CT as a clinical vocabulary for the computerized diagnosis and problem list. *AMIA Annu Symp Proc.* 2003;2003:699-703.
- NLM. *Fact Sheet: Unified Medical Language System*. 2006 [citado 12 Abr 2012]. Disponible en: <http://www.nlm.nih.gov/pubs/factsheets/umls.html>
- Pesquita C, Faria D, Falcão AO, Lord P, Couto FM. Semantic similarity in biomedical ontologies. *PLoS Comput Biol.* 2009;5:e1000443.
- Chase HS, Kaufman DR, Johnson SB, Mendonca EA. Voice capture of medical residents' clinical information needs during an inpatient rotation. *J Am Med Inform Assoc.* 2009;16:387-94.
- Mary V, Marquet G, Le Beux P. MeSH and specialized terminologies: coverage in the field of molecular biology. *Stud Health Technol Inform.* 2004;107(Pt 1):530-4.
- BVS. *DeCS Descriptores en Ciencias de la Salud 2011* [citado 29 Abr 2012]. Disponible en: <http://decs.bvs.br/E/decsweb2011.htm>
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, et al. Gene ontology: tool for the unification

- of biology. The Gene Ontology Consortium. *Nat Genet.* 2000;25:25-9.
28. King OD, Foulger RE, Dwight SS, White JV, Roth FP. Predicting gene function from patterns of annotation. *Genome Res.* 2003;13:896-904.
  29. Yang X, Li J, Lee Y, Lussier YA. GO-Module: functional synthesis and improved interpretation of Gene Ontology patterns. *Bioinformatics.* 2011;27:1444-6.
  30. GO. Gene Ontology Consortium Wiki: GO FAQ. 2010 [citado 23 Abr 2012]. Disponible en: [http://wiki.geneontology.org/index.php/GO\\_FAQ#Why\\_do\\_we\\_need\\_GO.3F](http://wiki.geneontology.org/index.php/GO_FAQ#Why_do_we_need_GO.3F)
  31. Kalet I. Modeling biological structure. En: *Biomedical informatics.* London: Academic Press; 2009. p. 304-5.
  32. Rosse C, Mejino JL Jr. A reference ontology for biomedical informatics: the Foundational Model of Anatomy. *J Biomed Inform.* 2003;36:478-500.
  33. Mejino JL, Rubin DL, Brinkley JF. FMA-RadLex: An application ontology of radiological anatomy derived from the foundational model of anatomy reference ontology. *AMIA Annu Symp Proc.* 2008;2008:465-9.
  34. Kalet IJ, Mejino JL, Wang V, Whipple M, Brinkley JF. Content-specific auditing of a large scale anatomy ontology. *J Biomed Inform.* 2009;42:540-9.
  35. HL7. About HL7. 2011 [citado 3 May 2011]. Disponible en: <http://www.hl7.org/about/index.cfm?ref=common>
  36. Ohmann C, Kuchinke W. Future developments of medical informatics from the viewpoint of networked clinical research. Interoperability and integration. *Methods Inf Med.* 2009;48:45-54.
  37. Abhyankar S, Lloyd-Puryear MA, Goodwin R, Copeland S, Eichwald J, Therrell BL, et al. Standardizing newborn screening results for health information exchange. *AMIA Annu Symp Proc.* 2010;2010:1-5.
  38. Strom BL, Schinnar R. Evaluating health information technology's clinical effects. *LDI Issue Brief.* 2011;16:1-4.
  39. Kohn L, Corrigan J, Donaldson M, editores. *To err is human: Building a safer health system.* 1.ª ed. Washington: National Academies Press; 2000.
  40. Kilbridge PM, Welebob EM, Classen DC. Development of the Leapfrog methodology for evaluating hospital implemented inpatient computerized physician order entry systems. *Qual Saf Health Care.* 2006;15:81-4.
  41. Consejo Técnico, Unidad Sectorial de Normalización en Salud. *Guía Técnica "Buenas Prácticas" para la seguridad del paciente en la atención en salud.* Bogotá: Ministerio de la Protección Social, USN-Unidad Sectorial de Normalización en Salud; 2010. Goodman KW, et al. Challenges in ethics, safety, best practices, and oversight regarding HIT vendors, their customers, and patients: a report of an AMIA special task force. *J Am Med Inform Assoc.* 2011;18:77-81.
  42. Moser RH. Diseases of medical progress. *N Engl J Med.* 1956;255:606-14.