

Medición de la calidad de los registros codificados en una historia clínica electrónica de internación

Pollán J.^a, Arbelbide J.^a, Pedernera F.^b, Borbolla D.^b, Achilli F.^a, Victoria V.^c; Gonzalez Bernaldo de Quiros L.^c; Gomez A.^c, Luna D.^c, Gonzalez Bernaldo de Quiros F.^c

^a Gerencia Médica Plan de Salud del Hospital Italiano de Buenos Aires, Argentina

^b Residencia de informática Médica, Hospital Italiano de Buenos Aires, Argentina

^c Departamento de Información Hospitalaria, Hospital Italiano de Buenos Aires, Argentina

Resumen

El Hospital Italiano de Buenos Aires desde 1998 desarrolló e implementó un programa para la captura, procesamiento y codificación de reportes de internación. Se utilizó la CIE-9-CM y los GRD para tal fin. El presente trabajo tiene como finalidad medir la calidad de la base de datos luego de este período inicial de implementación. Métodos: una muestra randomizada de pacientes internados fue tomada para crear una nueva base de datos gold standard. Una clasificación de errores fue utilizada para la evaluación. Se reportaron la proporción de registros correctos de la base original. Una comparación entre los GRD correctos e incorrectos fue creada mediante el test de Chi Cuadrado. Resultados: 107 reportes de internación fueron reprocesados para la construcción de la base gold standard. La asignación incorrecta de GRD fue de 26/107 (24.3%), 8 de los cuales no coincidían con una Categoría Diagnóstica Mayor, errores en el diagnóstico principal: 24/107 (22.4%), incorrecto procedimiento principal 16/89 (18%), errores en comorbilidades 28/107 (26.2%), menos del 1% de errores en complicaciones. La frecuencia de errores en correcto vs. Incorrecto GRD fue evaluado para diferentes campos: la proporción de diagnósticos principales correctos fue de 0.90 correctos GRD vs. 0.38 incorrectos ($\chi^2 p=0,0000$), para comorbilidades 0.22 y 0.38 para correctos e incorrectos GRD respectivamente ($\chi^2 p= 0.10$), para procedimiento principal: 0.87 y 0.61 para correcto e incorrecto GRD respectivamente ($\chi^2 p= 0.0097$). Discusión: la proporción de errores en códigos GRD fue similar a reportes publicados. La mayoría de los errores en los GRD fueron mala asignación de códigos en una misma Categoría Diagnóstica Mayor. La codificación incorrecta de diagnósticos principales y procedimientos principales son causados por falta de especificidad.

Summary

The Electronic Medical Record has become a central tool for many medical activities in the last years. We decided to implement it 5 years ago through a self developed software and a process of capture and codification of the medical data. We use the ICD9-CM and DRG codes for it. We wanted to know the data quality after this initial period of implementation. Methods: a random sampling of inpatients was taken in order to create a gold standard database. An error classification was used for the evaluation. We reported the proportion of right registries of the original database. A comparison between right and wrong DRG was made with chi square. Results: 107 abstracts of inpatient medical record were reabstracted for building the gold standard database. Wrong DRG assignment in 26/107 (24.3%), 8 of them outside from a Major diagnostic Category, errors in the principal diagnosis: 24/107 (22.4%), incorrect principal procedure: 16/89 (18%), any errors in comorbidities 28/107 (26.2%), less than 1% of errors in complications. The frequency of errors in wrong vs. right DRG was evaluated for the different fields: the proportion of right principal diagnosis is 0.90 in correct DRG vs. 0.38 in the incorrect DRG ($\chi^2 p=0,0000$), for comorbidities 0.22 and 0.38 for correct and incorrect DRG respectively ($\chi^2 p= 0.10$), for principal procedure 0.87 and 0.61 for correct and incorrect DRG, respectively ($\chi^2 p= 0.0097$). Discussion: The proportion of errors in DRG codes is similar to some reports of the literature. The majority of errors in DRG are miss codification inside of Mayor diagnostic Category. The wrong coding in principal diagnosis and principal procedure are caused by lack of specificity.

Palabras Clave

Base de datos, control de calidad, historia clínica electrónica, medidas de case mix, sistemas de información hospitalarios

Introducción

La informatización de los registros médicos permite el almacenamiento de datos para su utilización en diversas actividades en torno a la práctica médica [1]. En nuestro país se utiliza para el pago de prácticas y prestaciones un Nomenclador Nacional de Prestaciones que no representa adecuadamente el universo de las mismas ni toma en cuenta datos relevantes como son la edad, factores comórbidos y complicaciones de los pacientes que permita el agrupamiento en grupos homogéneos en cuanto al uso de recursos. Por lo antedicho en el año 1997 en el Plan de Salud del Hospital Italiano de Buenos Aires se decidió la implementación de un registro electrónico de pacientes que posibilitara la utilización de sistemas de codificación con la finalidad de conocer la casuística de los pacientes del prepago que sufrían internaciones. Los objetivos perseguidos fueron: poder contar con herramientas para la construcción de indicadores de gestión hospitalaria en un escenario de aumento de cápitales financiadas a riesgo compartido y que permitiera la realización de análisis clínico epidemiológicos así como trabajar en actividades de mejora de la calidad.

Se decidió codificar dichos registros con categorías que utilizan vocabularios estándares. La Clasificación Internacional de Enfermedades en su novena revisión con su Modificación Clínica (CIE-9-CM) [2] para los diagnósticos de internación y los Grupos Relacionados de Diagnósticos (GRD) [3] para contar con una herramienta que de cuenta del case mix hospitalario.

Una de las primeras medidas, fue la implementación de un aplicativo el cual permitía la transcripción a una base de datos, de un formulario en papel del reporte de internación de cada paciente. Este modelo de epicrisis estructurada contiene todos los puntos necesarios para la construcción del GRD. Para informatizar estos datos médicos, se capacitó a un grupo de estudiantes avanzados de la carrera de medicina en la codificación de la historia clínica mediante la CIE-9-CM y GRD. Dicho enfoque también fue utilizado en un proceso de codificación secundaria y centralizada para una lista de problemas en los registros ambulatorios con probada calidad [4, 5].

Luego de la implementación de la Historia Clínica Electrónica (HCE) para los episodios de internación [6], los procesos de la mejora en la calidad del dato consistieron en, A) mejoramiento de las herramientas del sistema, actualizando el software; B) capacitación en el uso de la HCE a los médicos; y C) la educación y supervisión continua de los codificadores, realizada por un médico con experiencia en codificación.

Hasta el momento no se habían realizado mediciones para evaluar la calidad de los datos de nuestra base de internación. La medición y los ciclos de mejora han demostrado utilidad en el incremento de precisión de los datos médicos registrados [7]. Por lo tanto creímos adecuado hacer una primera medición de la calidad de la base de datos que nos permita cuantificar la correctibilidad y completitud de los diferentes campos, con el propósito de implementar futuras estrategias de mejoría de calidad en la codificación, focalizadas en los tipos de errores más relevantes.

Hemos encontrado numerosos reportes que hablan de la calidad en las bases de datos médicas [8], muchos de ellos miden la repercusión económica que conllevan las alteraciones en la misma [3, 9, 10]. La construcción y mantenimiento de una base de datos no tiene por el momento funciones directas relacionadas con retribuciones de prestaciones médicas. Nuestra idea de implementar mejoras en la calidad de los datos médicos obedecen fundamentalmente a acercarse a la problemática médica de nuestros pacientes utilizando herramientas modernas de gestión y datos confiables [11].

Elementos del trabajo y metodología

El Hospital Italiano de Buenos Aires es un hospital universitario de 450 camas. Dentro de la población atendida existe un plan de medicina prepaga de 110 mil afiliados cuyo único ámbito de atención es el hospital y centros asociados. La información obtenida de todos los procesos de atención se halla centralizada en una historia clínica informatizada en la cual figuran los eventos de atención ambulatoria y de internación. Desde hace unos 4 años se registran los egresos hospitalarios en una base de datos y se codifican los mismos con la CIE-9-CM y se asigna un GRD. La recolección se lleva a cabo mediante una epicrisis estructurada (ver Tabla 1).

<i>Modelo de Epicrisis</i>	
Motivo de Internación	Situación que genera la necesidad de internación
Diagnóstico Principal	Afección fisiopatológica y/o anatomoclínica causante de la internación, definido al alta del paciente
Diagnósticos Secundarios (Otros Diagnósticos)	Diagnósticos diferentes al principal, no relacionados con episodios anteriores, que coexisten en el momento del ingreso o se desarrollan durante la internación.
Antecedentes	Enfermedades pre-existentes ya diagnosticadas e inactivas que no generan actos médicos, estudios o tratamiento.
Comorbilidades	Enfermedades pre-existentes ya diagnosticadas y en actividad que generan actos médicos, estudios o tratamientos, sin relación con el diagnóstico principal.
Complicaciones	Situaciones esperadas o no que ocurren durante la internación, se relacionan con el Diagnóstico principal, y efectos adversos de medidas diagnósticas o terapéuticas.
Procedimientos Primarios	Procesos invasivos especializados diagnósticos o terapéuticos relacionados con el diagnóstico principal.
Procedimientos Secundarios	Procesos invasivos especializados diagnósticos o terapéuticos no relacionados con el diagnóstico principal.
Evolución: texto libre	

Tabla 1. Modelo de epicrisis estructurada

Se construyó una base de datos gold estándar (GS), mediante la realización de una nueva epicrisis, para el análisis comparativo de la base de datos de internación.

Base de datos de pacientes internados:

La creación de un registro de internación comienza con la apertura del episodio por parte del médico responsable o residente. Éste genera la historia clínica en papel y completa los campos mencionados de la historia clínica electrónica. Una vez que el episodio se cierra, tanto la historia clínica en papel como el registro electrónico son revisados por un transcriptor que completa los datos de este último y reasigna campos mal asignados en base a la historia clínica. Posteriormente se procede a la codificación de los episodios mediante ICD-9-CM y DRG.

Construcción de la base gold estándar:

Se tomó una muestra de pacientes internados del hospital mediante un muestreo simple aleatorizado del censo diario de pacientes internados en el Hospital Italiano de Buenos Aires entre junio y julio de 2003. Diariamente dos médicos residentes de informática médica y un fellow del programa de gerenciamiento hospitalario recolectaron información clínica para la realización de la nueva epicrisis [12]. Partiendo de la premisa que el gold estándar ideal para una base de datos es el paciente y su realidad [8, 13], la información se obtuvo mediante revisión de historia clínica de internación, interrogatorio directo al paciente y/o familiar y revisión de la historia clínica ambulatoria (evoluciones médicas, laboratorios, informes de diagnóstico por imágenes y de anatomía patológica) en búsqueda de antecedentes y comorbilidades. Cada episodio de

internación fue discutido al menos en una ocasión con dos médicos con experiencia en el registro y codificación de pacientes internados. Posteriormente uno de esos médicos, encargado de la formación de codificadores y revisores de las epicrisis realizó las codificaciones. Todo el grupo médico permaneció ciego a los datos relevados en el registro electrónico. Los médicos tratantes desconocían la realización del estudio, así como también los codificadores.

Análisis comparativo

Se evaluó la correctibilidad mediante el % de correctos en cada uno de los campos con sus IC95%. Para los campos comorbilidades, complicaciones y antecedentes se muestra mediana, media y percentilo 5 y 95 de las diferencias de datos registrados en la base GS y la base; y la diferencia entre los datos registrados base GS y datos correctos en la base. Se comparó la proporción de diagnóstico principal, comorbilidades correctas en los grupos de GRD correctamente e incorrectamente asignados mediante test de chi cuadrado. Se creó una clasificación que compara ambos registros en cada uno de los campos (Tabla 2).

Campo	Clasificación
<i>GRD</i>	Correcto GRD diferente, dentro de categoría diagnóstica mayor (CDM). GRD diferente, fuera de una CDM
<i>Diagnóstico principal</i>	Correcto Incorrecto por falta de granularidad Incorrecto dentro de una CDM Incorrecto dentro de una CDM
<i>Comorbilidades (co)</i> <i>Antecedentes (atc)</i> <i>Complicaciones (comp)</i>	Número de co, atc y comp correctas Número de incorrectas por mala asignación de campo Número de co, atc, comp, incorrectas por otra causa Número de co, atc, comp, por diagnóstico inexistente Diferencias entre co, atc, comp de base GS y base. Diferencias entre co, atc, comp correctas de base GS y base
<i>Procedimiento principal</i>	Correcto Incorrecto por falta de granularidad Incorrecto dentro de una CDM Incorrecto fuera de una CDM Ausente
<i>Procedimientos secundarios</i>	Correcto Incorrecto por falta de granularidad Incorrecto dentro de una CDM Incorrecto fuera de una CDM Ausente Diferencia entre proa. Secundarios entre base GS y base

Tabla 2: Clasificación de posibles errores

Resultados

No se hallaron diagnósticos incompletos en los campos diagnósticos principal y procedimiento principal, así como en los códigos de DRG.

El número de GRD asignados correctamente fue de 81/107 (75,70%. IC95% 0,66-0,83).

De los 26/107 errores, 18 (16,8%) son diferencias dentro de una Categoría Diagnóstica Mayor (CDM) y 8 (7,5%) fuera de la misma.

La asignación del diagnóstico principal correcto fue hecha en 83/107 (77,6%. IC95% 0,68-0,85). De los

diagnósticos erróneos 8 (7,5%) lo fueron por falta de granularidad, 12 (11,2%) correspondían a errores dentro de la misma CDM y 4 (3,7%) por fuera de la misma. El 61,4% (16 / 26 GRD incorrectos) de los GRD mal asignados tienen errores en el diagnóstico principal. La proporción de diagnósticos principales correctos en los GRD s correctamente asignados es 0,90 versus 0,38 en los GRD incorrectos ($\chi^2 p=0,0000$).

Comorbilidades: El 28/107 (26.2%) de los registros tienen algún error. La mediana de la diferencia de comorbilidades entre la base GS y la base fue de 0 (media de 0,55) el percentilo 5 fue de -1 equivalente una comorbilidad inexistente y el 95 de 3.

La mediana de la diferencia entre las comorbilidades registradas en el GS y las comorbilidades correctas fue de 1 con p5 y 95 0 á 4, es decir que el 95% de los registros tienen menos de 5 comorbilidades presentes en el gold estándar que no figuran en la base.

Con la intención de medir el impacto de tener comorbilidades mal registradas en la asignación del GRD se dicotomizó la variable comorbilidad según esté correctamente asignada o tenga alguno de los errores de la clasificación. La proporción de comorbilidades con errores en la base es 0,22 en caso de tener el GRD correctamente asignado y 0,38. Esta diferencia no es estadísticamente significativa ($\chi^2 p=0,10$).

Complicaciones: Menos del 1% con algún tipo de error. La mediana de la diferencia de complicaciones entre el gold estándar y las complicaciones registradas es 0 (media 0,16) Percentilo 5 -95 de 0 y 1, respectivamente. La mediana de la diferencia de complicaciones entre GS y complicaciones registradas correctas los resultados son los mismos.

Es decir que el 95% de los registros tienen menos de 2 complicaciones no registradas en un mismo episodio de internación.

Antecedentes: la mediana de la diferencia entre el gold estándar y los antecedentes registrados es 0 (media 0,33). Percentilo 5-95 de -2 y 2. La mediana de las diferencias entre los antecedentes de gold estándar y antecedentes correctamente registrados es 1(media 0,75 p5-95 0-2).

Procedimiento principal: 89/107 (83.2%) de los episodios tienen procedimiento principal. El 82% de ellos correcto (IC95% 72-89). De los 89 episodios con procedimientos, 18 (20.2%) tienen errores en la asignación del DRG. La proporción de errores en el diagnóstico principal en los GRD quirúrgicos incorrectos es 38.9%.

La proporción de procedimientos correctos es significativamente mayor en el grupo de episodios con GRD correctos (0.87 vs. 0.61, respectivamente. $\chi^2 p = 0.0097$). Procedimientos secundarios: la mediana de las diferencias entre los registros de ambas bases fue de 0 (p5-95 0 y 0). Hay sólo 1 subregistro en 3 episodios y un registro inexistente en la base GS.

La mediana de las diferencias entre los registros de la base GS y los registros correctos es 0 (p5-95 0 y 0). Hay dos episodios con un subregistro en un procedimiento y otros dos con subregistro de dos procedimientos secundarios.

Discusión

En esta etapa se intentó cuantificar tres aspectos que creemos pertinentes para el diagnóstico inicial de la calidad de la base: la completitud de los campos, la correctibilidad, y la frecuencia de los distintos errores en los GRD incorrectos. La proporción de GRD correctos está cercana a reportes previos [9, 14]. La mayor proporción de GRD incorrectos es atribuible a errores dentro de una CDM, lo que implica un menor impacto en reflejar la distribución de las patologías de pacientes. Aunque produzca inevitables variaciones en el case mix. En el caso del diagnóstico principal el 85% son diagnósticos correctos o errores por baja granularidad

.Los errores en el diagnóstico principal son la causa más frecuentemente observada en las fallas en la asignación del GRD. En cuanto a las comorbilidades, el subregistro (baja sensibilidad) es mayor que el registro de diagnósticos inexistentes. Hay un 16% de diferencias en las comorbilidades mal asignadas entre el grupo de registros con GRD correcto e incorrecto. La ausencia de diferencias estadísticamente significativas podría deberse al escaso tamaño muestral o a la diferente ponderación de las comorbilidades por parte del grouper, para modificar DRG. El campo complicaciones tiene una adecuación casi perfecta, Hobo sólo dos complicaciones faltantes en un episodio y una proporción total de correctas de 99%, por encima de lo reportado [15]. El interrogante es si este porcentaje refleja una adecuada recolección de los datos del campo o de un sesgo de registro de la base GS. Esto último puede explicarse por la metodología de captura del dato en la base GS: consulta los antecedentes por múltiples vías mientras que los nuevos diagnósticos (como las complicaciones) así como los procedimientos los extrae de la epicrisis y a través del seguimiento directo del paciente. La proporción de procedimientos principales incorrectos es algo superior a la de otros reportes [15]. Una vez más la mayor proporción de errores se debe a falta de granularidad. Así como ocurre con el diagnóstico principal el error éste campo tiene un impacto mayor en la frecuencia de DRG correctos. El número de procedimientos subregistrados es escaso. Creemos válido traspolar conceptos del campo complicaciones.

En conclusión, en esta primera medición observamos que la base de datos tiene distinto grado de correctibilidad y completitud en los diferentes campos. Éstos no tienen el mismo impacto en la capacidad de modificar el DRG, reflejada por la frecuencia de errores en los GRD incorrectos. Una limitación del trabajo en este punto es que no se utilizó la tasa de modificación de GRD luego de la corrección de los datos [7]. Una debilidad planteada en el trabajo es la posible fuente de sesgo de recolección de datos de campos cuya única fuente de información es la historia clínica y el seguimiento del paciente. Aunque el método de comparación mediante base GS ha sido recomendado por otros en comparación a otras metodologías [6] A pesar de ello el impacto en la asignación de GRD parece ser menor. Por otra parte la experiencia del hospital italiano es particular en cuanto a la codificación por DRG ya que al no utilizar el mismo como medio de financiación no existen sesgos de sobreutilización de códigos más onerosos que pueden alterar el case mix hospitalario. El hallazgo de campos con distinta capacidad de modificación de GRD debería orientarnos a focalizar las estrategias correctivas en futuros ciclos de mejora de la calidad de la base de datos.

Datos de contacto

Dr. Javier Pollán: javier.pollan@hospitalitaliano.org.ar – Gerencia Médica. Plan de Salud. Hospital Italiano de Buenos Aires. Gascón 450. Ciudad Autónoma de Buenos Aires. Argentina. (C1181ACH)

Referencias

1. Whiting-O'Keefe, Q.E., D.W. Simborg, W.V. Epstein and A. Warger, *A computerized summary medical record system can provide more information than the standard medical record*. *Jama*, 1985. **254**(9): p. 1185-92.
2. U.S. Dept. of Health and Human Services - Centers for Disease Control and Prevention - Health Care Financing Administration, *ICD-9-CM international classification of diseases, ninth revision, clinical modification, sixth edition*. Official ed. 2000, [Washington, D.C.].
3. Simborg, D.W., *DRG creep: a new hospital-acquired disease*. *N Engl J Med*, 1981. **304**(26): p. 1602-4.
4. Luna, D., F.G. Bernaldo de Quiros, L. Garfi, E. Soriano and M. O'Flaherty, *Reliability of secondary central coding of medical problems in primary care by non medical coders, using the International Classification of Primary Care (ICPC)*. *Medinfo*, 2001. **10**(Pt 2): p. 300.
5. Luna, D., P. Otero, A. Gomez, M. Martinez, S. García Martí, M. Schpilberg, A. Lopez Osornio, and F.G. Bernaldo de Quiros. *Implementación de una Historia Clínica Electrónica Ambulatoria: "Proyecto ITALICA"*. in *Argentine Conference on Computer Science and Operational Research - 6to Simposio de Informática en Salud - 32 JAIHO*. 2003. Buenos Aires,

Argentina: SADIO.

6. Gonzalez Bernaldo de Quiros, F., E. Soriano, D. Luna, A. Gomez, M. Martinez, M. Schpilberg, and A. Lopez Osornio. *Desarrollo e implementación de una Historia Clínica Electrónica de Internación en un Hospital de alta complejidad*. in *Argentine Conference on Computer Science and Operational Research - 6to Simposio de Informática en Salud - 32 JAIIO*. 2003. Buenos Aires, Argentina: SADIO.
7. Lorenzoni, L., R. Da Cas and U.L. Aparo, *The quality of abstracting medical information from the medical record: the impact of training programmes*. *Int J Qual Health Care*, 1999. **11**(3): p. 209-13.
8. Hogan, W.R. and M.M. Wagner, *Accuracy of data in computer-based patient records*. *J Am Med Inform Assoc*, 1997. **4**(5): p. 342-55.
9. Hsia, D.C., *Diagnosis related group coding accuracy of the peer review organizations*. *J Ahima*, 1992. **63**(9): p. 56-64.
10. Hsia, D.C., W.M. Krushat, A.B. Fagan, J.A. Tebbutt and R.P. Kusserow, *Accuracy of diagnostic coding for Medicare patients under the prospective-payment system*. *N Engl J Med*, 1988. **318**(6): p. 352-5.
11. Lorence, D.P. and I.A. Ibrahim, *Benchmarking variation in coding accuracy across the United States*. *J Health Care Finance*, 2003. **29**(4): p. 29-42.
12. Yao, P., B.R. Wiggs, C. Gregor, R. Sigurnjak and P. Dodek, *Discordance between physicians and coders in assignment of diagnoses*. *Int J Qual Health Care*, 1999. **11**(2): p. 147-53.
13. Parkin, D.M. and C.S. Muir, *Cancer Incidence in Five Continents. Comparability and quality of data*. *IARC Sci Publ*, 1992(120): p. 45-173.
14. Lorenzoni, L., R. Da Cas and U.L. Aparo, *Continuous training as a key to increase the accuracy of administrative data*. *J Eval Clin Pract*, 2000. **6**(4): p. 371-7.
15. Barrie, J.L. and D.R. Marsh, *Quality of data in the Manchester orthopaedic database*. *Bmj*, 1992. **304**(6820): p. 159-62.