

Captura de datos clínicos con verificación aleatoria: una nueva técnica para controlar y mejorar la calidad del registro

Rosario Granero y José M. Doménech
Universitat Autònoma de Barcelona

Este trabajo plantea las principales dificultades de la grabación de datos clínicos. Se ha diseñado una experiencia que consiste en capturar en un ordenador, en ausencia de protecciones automáticas, 335 entrevistas estructuradas codificadas previamente a mano por entrevistadores entrenados. El objetivo general del estudio consiste en valorar la eficacia de un nuevo sistema para controlar y mejorar la grabación: la verificación aleatoria (VA) de un porcentaje de cuestionarios, basada en el principio de «feedback». Para ello se han medido y tipificado los errores que cometen operadores profesionales durante la captura. Los objetivos específicos son: 1) valorar la eficacia de la VA respecto a la doble entrada (DE) –esta última constituye la estrategia habitual utilizada para mejorar la calidad del registro–; 2) comprobar qué modalidades de error contienen los datos grabados; y 3) valorar la eficacia de los operadores para resolver las discrepancias aparecidas durante la verificación (midiendo qué porcentaje de errores persiste en los campos con discordancias durante la segunda captura de la VA y la DE). Los resultados principales indican que la VA ofrece un rendimiento similar a la DE con distintos operadores, mientras que la entrada simple y la DE con un mismo operador no constituyen sistemas eficaces para garantizar la calidad del registro. En la discusión se destacan qué aspectos de la medición y de la captura de datos se mejoran usando controles informáticos, qué requisitos mínimos de calidad deben prevalecer durante la grabación y qué elementos requieren los sistemas informáticos para operar de forma eficiente.

Clinical data capture with random verification: A new procedure for controlling and improving the data capture quality. This work contemplates the main difficulties related to the capture of clinical data into the computer. We have designed an experience that consists of capturing into the computer, in absence of automatic protections, 335 structured interviews coded on paper by trained interviewers. The general objective of this study consists of rating the efficiency of a new computerized system (based on the feedback principle) for controlling and improving the capture: the random verification (RV) of a percentage of questionnaires. To accomplish this objective, we have measured and classified the mistakes that professional operators commit during the data capture. The specific objectives are: 1) to rate the efficiency of the RV compared to double entry (DE) –this last strategy is the most commonly used for improving the quality of the capture–; 2) to check the types of mistakes contained in the recorded data; and 3) to rate the operators' efficiency for resolving the discrepancies that appeared during verification (through the measurement of the mistakes that persist in the fields with disagreements during the second capture of the RV and the DE). The principle results obtained indicate that the RV and the DE with different operators offer similar performance, while the simple entry and the DE with only one operator do not constitute effective systems for guaranteeing the quality of the capture. In the discussion we emphasize what aspects of the measurement and the data capture are improved using computerized controls, what minimum quality requirements should be present during the recording and what elements the computerized systems require for operating efficiently.

Durante los últimos años han aparecido diversos estudios que abordan aspectos relacionados con la calidad del proceso de gestión de datos (Rondel, Varley y Webb, 1999) y con las consecuen-

cias que comporta la falta de controles rutinarios de calidad (Groves, 1989; Rowley, 1989; Magnusson y Bergman, 1990; Redman, 1992; Granero, 1999). En estos trabajos es habitual contemplar la gestión de datos como una tarea compleja que conlleva diferentes fases: 1) el diseño de la base de datos (definición de la información a registrar, creación de la estructura de las tablas y especificación de las protecciones de registro); 2) la captura de la información; 3) la preparación de la matriz de datos para el análisis estadístico; y 4) el análisis de datos. Se ha argumentado que la falta de planificación o la realización incorrecta de cualquiera de estas etapas

Correspondencia: Rosario Granero Pérez
Facultad de psicología
Universitat Autònoma de Barcelona
08193 Barcelona (Spain)
E-mail: Roser.Granero@uab.es

contribuye a la introducción de sesgos y errores que afectan a la validez interna y externa de un estudio, y que pueden comprometer seriamente las conclusiones que se formulan (Sarís, 1991; Freedland y Camey, 1992; Moritz et al., 1995). Grisso et al. (1991) presentan diferentes propuestas para mejorar la calidad y evitar estas fuentes de error.

El interés que suscita el estudio de la calidad de datos tiende a afianzarse con la introducción de las nuevas prestaciones que hoy ofrece la informática. Aunque todavía persiste la creencia de que el mero uso de ordenador incrementa la calidad de los estudios (sobre todo porque permite adoptar diseños y análisis estadísticos más complejos), cada vez más investigadores advierten que no siempre se consigue mejorar la calidad de los resultados finales publicados (Butcher, 1994; Cobos, 1995).

Hasta el presente, sin embargo, sólo en algunos ámbitos de la investigación en Ciencias de la Salud el proceso de gestión de datos está sometido a una auditoría, por ejemplo en los ensayos clínicos que disponen de los suficientes recursos y de una estricta legislación tanto de la F.D.A. como de las agencias de Europa y Japón (Lewis, Jones y Röhmel, 1995; I.C.H., 2000). Pero en muchos sectores aún no existen incentivos suficientes para que los investigadores utilicen estrategias para controlar y mejorar la calidad de los datos, y tampoco se han promovido planes de formación específicos para orientar a los profesionales en estos aspectos. Esto implica que a los datos que se están publicando en nuestra ciencia (incluso en revistas prestigiosas) se les presupone la calidad a pesar de que muy pocos investigadores han utilizado las estrategias y filtros de control requeridos.

Afortunadamente, este panorama general tiende a cambiar porque muchas disciplinas comienzan a asumir que la fase específica de captura de los datos constituye una etapa tan crítica y crucial como la definición de la matriz de datos o la realización de los análisis estadísticos. Por ello la investigación coetánea atribuye una gran relevancia al estudio de las fuentes de error que afectan a la calidad del registro/grabación, y también a las estrategias de control que se pueden adoptar para que los datos no contengan errores antes de ser analizados.

Al abordar este problema en Ciencias de la Salud se comprueba que es habitual recoger la información de los sujetos a través de cuestionarios administrados en lápiz y papel, y posteriormente se introduce en el ordenador. Esta forma de trabajo no permite disponer de un indicador de la calidad de la información cuando se realiza una «entrada simple» (ES) de los datos registrados en los formularios. Una de las alternativas que tradicionalmente se ha considerado más eficaz para garantizar que la captura es correcta y para disponer de indicadores de calidad ha consistido en efectuar una «doble entrada» (DE). Más recientemente, y como consecuencia del elevado coste de la DE, se han propuesto otras alternativas como la «verificación aleatoria» (VA) de un porcentaje de los casos elegidos al azar por el programa de entrada de datos y presentados al mismo operador con el suficiente retardo para evitar los efectos de recuerdo. Estudios previos realizados con datos ficticios constatan que, en ausencia de controles automáticos -esto es, sin que se hayan definido validaciones por rango, condiciones lógicas o saltos automáticos-, la VA aumenta la eficacia de la captura de los datos respecto a la ES, sin afectar a la eficiencia, debido al «feedback» que ejerce sobre el operador (Doménech, Losilla y Portell, 1998).

El propósito general de este trabajo es el estudio de los errores cometidos por operadores profesionales al capturar datos diagnós-

ticos reales registrados en formularios codificados a mano por clínicos expertos. Para ello hemos diseñado una experiencia que consiste en grabar en un ordenador, en ausencia de controles automáticos, entrevistas estructuradas codificadas a mano por entrevistadores entrenados en el uso de dichos protocolos. Los objetivos específicos del trabajo consisten en: 1) demostrar la eficacia de la VA frente a la doble DE (realizada por uno o dos operadores); 2) definir y cuantificar qué tipos de error contienen los datos tras ser grabados (omisión de valores, cambios de código o completar campos que deberían quedar vacíos); y 3) valorar si las discrepancias que aparecen durante la verificación (esto es, durante la segunda captura que se realiza en los cuestionarios sometidos a VA y DE) se resuelven eficazmente. La captura de los datos se ha realizado con el Sistema DAT, que permite implementar las técnicas de VA y DE (Doménech et al., 1998).

Esta experiencia reproduce en gran medida la estrategia procedimental que subyace en un gran número de trabajos que se llevan a cabo en nuestra disciplina. Por ello, las medidas obtenidas con relación a los errores de grabación que comete un operador en el registro de datos, unido al estudio de su distribución y tipología, permiten obtener indicadores de la calidad de los datos aportados en dichos estudios. La experiencia también ayuda a determinar qué variables pueden estar relacionadas con los incrementos y/o decrementos de la calidad de la captura, y qué características debería contemplar el software para funcionar de forma adecuada durante la grabación de datos.

Método

Sujetos

En el estudio ha intervenido un operador principal, encargado de realizar la grabación completa de los registros (una parte con DE y otra con VA) y un segundo operador encargado de realizar otra captura independiente de los registros sometidos a DE. Esta segunda figura es necesaria porque diferentes trabajos indican que la eficacia de la DE depende de que la segunda captura sea realizada por un operador diferente al que llevó a cabo la primera (Blumenstein, 1993; Gassman, Owen, Kuntz, Martin y Amoroso, 1995). Por otro lado, disponer de dos operadores permite comparar la DE con una interpretación y con dos interpretaciones independientes de los formularios en papel.

Las condiciones de inclusión de los operadores consistían en: un mínimo de 1 año de experiencia en la introducción de datos y demostrar eficiencia en la captura (se exigió un ritmo de grabación superior a 30 palabras por minuto sin error).

Material

Se ha dispuesto de 335 entrevistas estructuradas DICA-R («Diagnostic Interview for Children and Adolescents-Revised»; Reich, Shayka y Taibleson, 1991), codificadas en papel por clínicos entrenados. La información diagnóstica se obtuvo de familias que acudieron a consulta de distintos Servicios de Psicología y Psiquiatría Infantil de hospitales públicos de Barcelona, entre octubre de 1993 y abril de 1996, y de sujetos escolarizados de la población general.

Se ha creado un formulario informatizado con el Sistema DAT que consta de 983 campos y carece de protecciones automáticas para la captura de datos (no se han definido validaciones ni saltos),

de modo que los operadores pueden grabar cualquier valor en cualquier campo siempre y cuando éste sea consistente con el formato definido. Asimismo, y para facilitar el registro de la información, el diseño de las pantallas se ha realizado para que la disposición de los campos reproduzca lo más fielmente posible el aspecto del protocolo en papel con los datos originales.

Diseño

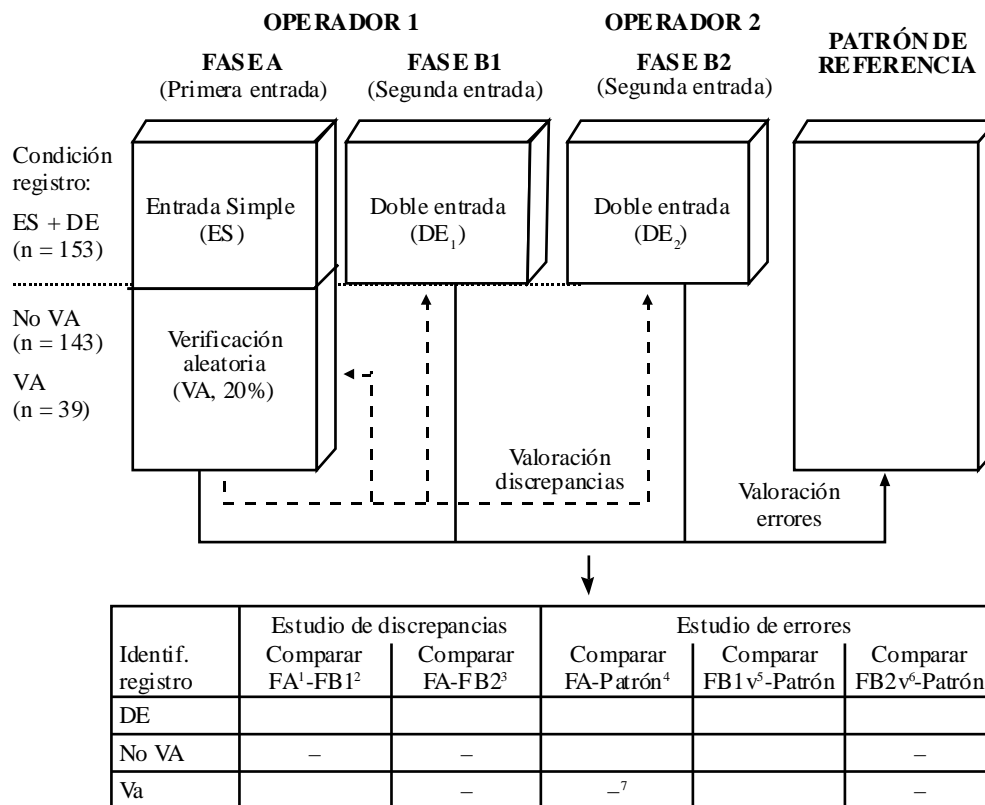
El diseño realizado viene representado en la parte superior de la Figura 1. La variable independiente del estudio es el tipo de registro. La condición ES se ha caracterizado por una consigna verbal informando al operador de la importancia de llevar a cabo una captura correcta de los datos que estaban codificados manualmente en las hojas de papel. La condición VA, definida con un 20% de formularios a verificar y con un retardo en la doble entrada igual a 2 formularios, se ha caracterizado por una consigna verbal informando que a partir de ese momento el programa controlaría los errores cometidos durante la entrada de datos, obligando a repetir la introducción de algunos registros; el operador también fue informado de que si se cometían discrepancias respecto a la primera entrada, el programa presentaría un mensaje de aviso y le solicitaría que verificase los campos con discordancia. La condición DE se ha caracterizado por una consigna verbal informando de que el programa volvería a solicitar la in-

troducción de todos los registros grabados al inicio, de forma similar a la VA.

Se han definido dos variables dependientes, las discrepancias y los errores:

1. Las discrepancias que se producen entre las Fases A y B se han valorado como la falta de coincidencia al grabar una misma serie de valores en dos ocasiones diferentes (Figura 1). Este indicador es una medida indirecta de la frecuencia de errores que cometen los operadores durante la captura de los datos. Cuando no se producen discrepancias entre dos entradas se supone que la información ha sido capturada prácticamente sin errores; por otro lado, cuando se producen discrepancias, se asume que al menos una de las dos capturas es incorrecta. Asimismo, y puesto que la frecuencia relativa de discrepancias es un indicador de la consistencia (o aptitud) de cada operador durante la grabación, este índice se considera una medida adecuada para comparar diferentes estrategias de registro de datos.

2. Los errores se han valorado como la falta de coincidencia entre los datos grabados por los operadores y la información registrada en las fuentes documentales originales (en nuestro estudio, dichas fuentes están representadas por el patrón o entrada de referencia). La frecuencia relativa de errores constituye un indicador especialmente adecuado para valorar la eficacia de los procedimientos que se emplean para el registro de datos.



¹Fase A; ²Fase B1; ³Fase B2; ⁴Entrada de referencia; ⁵Datos verificados por el mismo operador; ⁶Datos verificados por un operador diferente; ⁷Esta comparación no se ha podido realizar debido a que la versión utilizada del Sistema DAT, tras efectuar la VA, únicamente mantiene almacenado el valor final que el operador ha considerado adecuado a cada campo.

Figura 1. Esquema del diseño de la experiencia

Procedimiento

La tarea de grabación fue presentada a los operadores como un trabajo rutinario de entrada de datos clínicos, pero al finalizar la experiencia fueron informados sobre los verdaderos propósitos del estudio. La captura de datos se ha estructurado en dos etapas:

1. Entrenamiento: se presenta el programa y se ayuda a practicar con él la entrada de 10 cuestionarios de prueba, y a realizar copias de seguridad. De este modo, los operadores tomaban contacto con el «software» de entrada de datos y se familiarizaban con los protocolos. También se informó que podían realizar los descansos que necesitaran durante cada sesión, y se incentivó la rapidez de la entrada de datos.

2. Fase experimental: finalizada la fase de entrenamiento, se iniciaron las sesiones de captura de datos. En esta fase, las entrevistas se presentaron a los operadores en orden aleatorio. A medida que avanzaba el registro se daban las consignas verbales ES, VA y DE. Cada sesión duraba unas 4 horas y los operadores podían realizar descansos regulares según su propia voluntad. Asimismo, el operador que llevó a cabo la entrada de las Fases A y B1 (Figura 1) no fue informado de que los datos volverían a ser entrados por otro operador (Fase B2).

Tras finalizar la fase experimental se elaboró el patrón o entrada de referencia generado con un sistema de filtros sucesivos. Para ello se pidió al segundo operador que volviera a grabar la totalidad de los cuestionarios (n=335), resolviendo las discrepancias que se producían respecto a la entrada verificada. A continuación, tres evaluadores de nuestro grupo de investigación, que utilizan la entrevista DICA-R, realizaron tres nuevas comprobaciones independientes de todos los registros.

Los datos disponibles permiten realizar las comparaciones que aparecen destacadas en la tabla de la parte inferior de la Figura 1. El análisis estadístico se ha efectuado con el Sistema SPSS 9.0 para Windows (SPSS, 1999). La estrategia de análisis ha consistido en obtener las proporciones de discrepancias y errores (consideradas en este diseño como incidencias acumuladas), y las correspondientes diferencias (absolutas) entre las distintas condiciones de registro. Los intervalos de confianza se han obtenido a partir de macros.

Finalmente, debido a que la entrevista DICA-R posee una estructura ramificada (en forma de árbol lógico) que hace que sólo se complete una parte de la información diferente en cada entrevista, es conveniente definir variables que miden el total de discrepancias/errores en función de cómo aparecen los campos en el patrón de referencia (completos o vacíos).

Resultados

En relación al estudio de las discrepancias entre las Fases A y B se comprueba que, independientemente del contenido del patrón, la DE efectuada por un operador distinto constituye la condición de registro que obtiene una mayor proporción de discrepancias. En el caso de que los datos sean grabados por el mismo operador en ambas ocasiones, la DE obtiene más discrepancias que la VA. En concreto, las comparaciones realizadas sobre el total de campos verificados indican que la DE realizada con un único operador incrementa significativamente la proporción (absoluta) de discrepancias (Dif.=6.2%, IC 95%: 5 a 7.5) respecto a la VA. Por su parte, la DE realizada por un operador diferente incrementa la proporción de discrepancias un 14.89% (IC 95%: 13.5 a 16.2) res-

pecto a la VA, y un 8.6% (IC 95%: 7.6 a 9.7) respecto a la DE con una sola interpretación.

Los resultados relativos a los errores existentes cuando se comparan los datos capturados y el patrón quedan recogidos en la parte superior de la Tabla 1. En primer lugar aparecen las proporciones totales de errores para cada condición de registro al comparar la Fase A con la entrada de referencia, y los datos que permiten valorar el efecto del «feedback» de la VA respecto a la simple entrada. Estos resultados sugieren que la consigna que se facilita al activar la VA unida al control que ejerce el programa sobre el operador actúan decrementando la proporción de errores. Por ejemplo, para el total de registros la diferencia es del 4.8% (IC 95%: 4 a 5.6). La parte superior de la Tabla 1 también recoge la proporción total de errores para cada condición de registro al comparar los datos verificados (esto es, al finalizar la Fase B) con el patrón, los resultados de las comparaciones que permiten valorar el efecto del «feedback» de la VA respecto a la DE, y el efecto de la doble interpretación de los datos. Estos resultados sugieren que la mayor proporción de errores corresponde a la condición de DE realizada por el mismo operador. Tanto la VA como la DE realizada por un operador diferente obtienen proporciones de error notablemente inferiores. Finalmente, los resultados de la parte superior de la Tabla 1 también muestran que cuando se compara la Fase B verificada con el patrón las proporciones de error son notablemente menores que cuando se compara la Fase A con el patrón, independientemente del tipo de registro.

La parte central de la Tabla 1 permite estudiar el tipo de error cometido durante la captura. Esta parte de la tabla recoge la proporción de errores en función de su tipología y de la condición de registro. La estimación se ha efectuado respecto al total de campos grabados (cuestionarios completos) y respecto al total de campos en los que se detectaron errores. El tipo de error se ha valorado comparando los datos de la Fase A y los datos verificados (al finalizar la Fase B) con el patrón, y se ha clasificado en tres categorías: error por cambio de código, error por omitir valores que figuran en el patrón y error por llenar campos que aparecen vacíos en el patrón. Los resultados de comparar la Fase A y el patrón sugieren que tanto para la consigna de la ES como para la consigna de la VA, el tipo de error menos frecuente se produce como consecuencia de llenar campos que deberían estar vacíos. De otro lado, en el caso de la ES la proporción de errores por omisión de valores es más frecuente que la proporción de errores por cambio de código. Esta relación se invierte cuando se activa la VA. Al comparar la Fase B verificada y el patrón, los resultados indican que tanto para la DE con el mismo operador como para la VA, el tipo de error menos frecuente se produce como consecuencia de llenar campos que deberían estar vacíos, y el más frecuente como consecuencia del cambio de códigos. De otro lado, cuando la DE se efectúa por un operador diferente, el tipo de error menos frecuente corresponde a la omisión de valores, y el más frecuente al cambio de códigos; en esta última condición de registro, sin embargo, las diferencias entre los tres tipos de error son menos evidentes.

La parte inferior de la Tabla 1 permite estudiar la eficacia en la resolución de las discrepancias. Esta parte de la tabla recoge la proporción de errores que persisten al comparar los datos verificados (al finalizar la Fase B) y el patrón en los campos en los que se habían detectado discrepancias durante el proceso de verificación. En este caso la estimación también se ha efectuado respecto al total de campos grabados y respecto al total de campos con discrepancias durante la segunda captura. Los resultados aportados al

comparar la VA frente a la DE con un mismo operador, y ésta con la DE con diferente operador, indican que en todas las condiciones de registro persisten errores en los campos en los que se habían producido discrepancias durante el proceso de verificación y los operadores se habían visto obligados a resolverlas. La situación más desfavorable para el registro corresponde a la DE con una sola interpretación de la información, y la más favorable a la DE con dos interpretaciones.

Conclusión

Los resultados de este estudio constatan que la ES y la DE con un único operador no constituyen sistemas eficaces para garantizar la calidad del registro. Por el contrario, la VA contribuye a incrementar la eficacia global del proceso de captura de datos, resulta más eficaz que la DE con un único operador, y ofrece un rendimiento similar a la DE con dos operadores distintos. Esto se evi-

<i>Tabla 1</i>				
Proporciones de errores en tantos por mil (entre paréntesis figuran los intervalos de confianza del 95%)				
ERRORES CONTENIDOS EN LOS DATOS TRAS LA CAPTURA (%)				
Comparación de la Fase A con el patrón				
Condición de entrada	Patrón vacío	Patrón con valor	TOTAL	
Entrada simple	4.11 (3.73 a 4.49)	38.4 (36.6 a 40.3)	13.7 (13.1 a 14.3)	
No VA	2.34 (2.05 a 2.63)	29.0 (27.3 a 30.8)	8.89 (8.40 a 9.38)	
Efecto VA: ES-NoVA	1.77 (1.29 a 2.25)	9.38 (6.83 a 11.9)	4.78 (4.02 a 5.55)	
Comparación de la Fase B con el patrón				
Condición de entrada	Patrón vacío	Patrón con valor	TOTAL	
DE ₁ (con operador)	2.17 (1.89 a 2.44)	24.8 (23.3 a 26.3)	8.47 (8.01 a 8.93)	
VAT ₂ (NO VA + VA)	0.89 (0.73 a 1.04)	11.3 (10.3 a 12.3)	3.41 (3.14 a 3.68)	
DE ₂ (con 2 operadores)	1.16 (0.96 a 1.36)	6.44 (5.67 a 7.21)	2.63 (2.37 a 2.89)	
Efecto VA: DE ₁ - VA _T	1.28 (0.96 a 1.60)	13.5 (11.7 a 15.3)	5.06 (4.52 a 5.60)	
VA _T - DE ₂	-0.27 (-0.53 a -0.02)	4.83 (3.58 a 6.08)	0.78 (0.40 a 1.15)	
Efecto DE ₂ : DE ₁ - DE ₂	1.00 (0.66 a 1.35)	18.3 (16.7 a 20.0)	5.84 (5.31 a 6.37)	
TIPOS DE ERROR (%)				
Comparación de la Fase A y el patrón				
Condición	Cálculo prop.	Cambiar códigos	Omitir valores	Llenar campos
Entrada simple	campos grabados	4.53 (4.20 a 4.87)	6.18 (5.78 a 6.57)	2.97 (2.69 a 3.24)
	campos con error	332 (311 a 352)	452 (430 a 473)	217 (199 a 235)
No VA	campos grabados	4.10 (3.77 a 4.44)	3.02 (2.74 a 3.31)	1.76 (1.54 a 1.98)
	campos con error	462 (434 a 489)	340 (314 a 366)	198 (176 a 221)
Comparación de la Fase B y el patrón				
Condición	Cálculo prop.	Cambiar códigos	Omitir valores	Llenar campos
DE ₁ (1 oper.)	campos grabados	3.66 (3.36 a 3.97)	3.24 (2.96 a 3.53)	1.56 (1.36 a 1.76)
	campos con error	433 (405 a 460)	383 (356 a 410)	184 (163 a 206)
VA _T NoVa+Va	campos grabados	1.90 (1.70 a 2.10)	0.84 (0.70 a 0.97)	0.67 (0.55 a 0.79)
	campos con error	577 (518 a 597)	246 (212 a 280)	197 (165 a 228)
DE ₂ (2 oper.)	campos grabados	0.98 (0.83 a 1.14)	0.81 (0.67 a 0.95)	0.84 (0.69 a 0.98)
	campos con error	374 (326 a 421)	308 (263 a 354)	318 (272 a 364)
ERRORES EN LOS CAMPOS CON DISCORDANCIA DURANTE LA VERIFICACION (%)				
Condición de entrada	% sobre el total de campos verificados		% sobre el total de campos con discrepancias	
DE ₁ (1 operador)	2.69 (2.42 a 2.96)		157 (143 a 172)	
Verificación aleatoria	0.56 (0.31 a 0.81)		51.7 (31.8 a 78.7)	
DE ₂ (2 operadores)	0.92 (0.76 a 1.07)		35.6 (29.7 a 42.2)	
Efecto VA: DE ₁ - VA	2.13 (1.76 a 2.49)		106 (79.1 a 132)	
Efecto VA: VA - DE ₂	-0.35 (-0.65 a -0.06)		16.1 (0.675 a 39.0)	
Efecto DE ₂ : DE ₁ - DE ₂	1.77 (1.46 a 2.09)		122 (106 a 137)	

dencia cuando: 1) se valoran las discrepancias que se cometen al entrar dos veces la misma serie de datos, ya que en este caso se ha comprobado que la condición más favorable para el registro sin discordancias corresponde al momento en que se activa la VA; 2) se valoran los errores que persisten cuando se compara la Fase A con el patrón (en este caso también se observa que la menor proporción de errores se obtiene al activar la VA); y 3) se comparan los datos verificados al finalizar la Fase B con el patrón (en este caso los resultados obtenidos indican que las proporciones de errores son similares en las condiciones de registro correspondientes a la DE con diferentes operadores y a la VA, y que estos valores son notablemente inferiores a la proporción de errores cuando se efectúa la DE de los registros con un único operador).

Por otro lado, los resultados también destacan que ninguna de las estrategias de captura de datos garantiza totalmente la grabación correcta, ni siquiera la DE como sugiere la literatura (Blumenstein, 1993; Wolf, 1993; Gibson, Harvey, Everett y Parmar, 1994; Gassman et al., 1995). Esto se evidencia cuando se comparan los datos verificados (al final de la Fase B) con el patrón: en todas las condiciones de registro aparecen diferentes tipos de error (por cambio de código, por omisión y por completar campos que deberían estar vacíos). Los errores persisten en un porcentaje no despreciable de campos con discrepancias durante la segunda grabación.

Discusión

En líneas generales, los resultados obtenidos indican que las condiciones más óptimas para el registro de datos aparecen cuando se efectúa una DE con diferentes interpretaciones de la información, y también cuando se facilita a un único operador información («feedback») sobre el control que el programa computerizado ejerce sobre la grabación de datos (por ejemplo, cuando se le solicita que vuelva a grabar algunas entrevistas escogidas de forma aleatoria y se le advierte de las discrepancias que se van produciendo). En este sentido, la VA se presenta como una técnica adecuada para controlar y mejorar la calidad de la captura de datos clínicos. Es importante destacar que la principal ventaja de la VA frente a la DE consiste en que reduce considerablemente el coste total del registro de datos (únicamente representa un ligero incremento respecto a la ES), y esto es muy importante en los contextos científico (general) y clínico (particular) porque se suelen manejar grandes volúmenes de datos y se dispone de bajo presupuesto para la grabación. También es importante tener presente que ambas estrategias constituyen sistemas de control de la calidad de datos compatibles. Así, por ejemplo, en aquellos estudios en los cuales se obtuvieran proporciones de discrepancias elevadas para la VA se aconsejaría realizar una DE de los registros con otro operador con objeto de reducir la proporción final de errores. Y también sería adecuado utilizar la VA durante la primera captura de los datos y completar la grabación con la DE del resto de los registros (de esta forma, por ejemplo, la VA permitiría disponer de una primera entrada con menor proporción de errores y mejoraría el rendimiento de la DE). Finalmente, ambas técnicas también se podrían utilizar simultáneamente en estudios con operadores, registros o campos que requieran diferentes niveles de control.

Por otro lado, y al margen de la valoración que se puede realizar sobre las técnicas de captura de datos utilizadas en este estudio, es importante advertir que parte de las discrepancias y los errores producidos durante la grabación podrían ser consecuencia del tipo de registro original. En este caso, parece lógico suponer

que cuando los datos se codifican a mano la interpretación por parte del operador se dificulta. Esta situación implica necesariamente plantear algunas consideraciones sobre los requisitos que debería cumplir el software de captura de datos. A continuación especificamos algunas de las más importantes.

En primer lugar, consideramos que todo sistema gestor de datos debería diferenciar entre las posibles situaciones que se producen durante la grabación (captura directa o diferida), facilitando diferentes atributos de «valor válido» y «valor no válido».

Dentro de la categoría de los valores válidos se debería poder diferenciar entre valor definitivo y valor provisional o difícil de interpretar. El primero se asignaría en aquellas situaciones en las que no existen dudas respecto al valor que se graba en un campo. De otro lado, cuando los operadores considerasen necesario confirmar con posterioridad la información que graban (por ejemplo consultando fuentes adicionales o solicitando interpretaciones independientes), los sistemas gestores deberían permitir asignar un atributo especial a dichos campos. Estos atributos permitirían realizar interesantes estudios sobre la captura y las causas que la dificultan, ya que los datos quedarían marcados durante la grabación. Así, por ejemplo, sería posible valorar si un alto porcentaje de errores es debido a una difícil interpretación (lo cual sería a su vez indicativo de una codificación deficiente de los datos en el papel o registro original).

Dentro de la categoría de los valores no válidos se debería poder diferenciar entre el código de valor pendiente y el atributo de valor inconsistente. En el primer caso, los valores pendientes designarían que la información es desconocida en el momento de grabar el valor, y que por lo tanto el campo queda pendiente de ser grabado tras la consulta de otras fuentes adicionales o de la misma fuente en otro momento. En el segundo caso, el atributo de valor inconsistente se asignaría en aquellas situaciones en las que se activan protecciones automáticas y un valor no cumple los criterios de validación de un campo. Esta última asignación permitiría: (1) disponer de las inconsistencias que se producen entre los datos y las protecciones durante la captura; (2) guardar en un registro de incidencias la causa exacta de dichas inconsistencias; y (3) que el responsable de gestionar la base de datos valorase el origen de la inconsistencia (valor erróneo o protección mal definida).

Siguiendo con los aspectos de «software», también resulta imprescindible de cara a la calidad de la gestión poder registrar todas aquellas incidencias que se producen durante la entrada de un valor en un campo. En este sentido, el sistema, además de generar y ofrecer mensajes de aviso y/o error ante datos inconsistentes o no coincidentes entre dos entradas, debería disponer de un archivo histórico de incidencias en el que quedarán grabados los sucesivos cambios realizados en cada campo, así como también el momento temporal en que se producen, el responsable que los autoriza, el operador que los realiza y los motivos del cambio, tal como sugieren Gassman et al. (1995). Este sistema de grabación resulta indispensable si tenemos presentes los resultados de este estudio, que apoyan la hipótesis de que el encargado de resolver las discrepancias o incongruencias debería ser un sujeto responsable de otros aspectos de la gestión de datos, altamente implicado con la investigación, conocedor del tipo de datos que se manejan y concienciado de la importancia de un registro correcto.

Como punto final nos gustaría destacar que el área de investigación encargada del estudio de la calidad de la captura de datos contempla muchos aspectos todavía pendientes de valorar. En este sentido, es deseable emprender nuevas líneas de trabajo que

aporten nuevas evidencias sobre diversas cuestiones: qué sistema de mensajes resulta más efectivo durante el proceso de verificación, qué acciones se deben llevar a cabo para resolver las discrepancias, con qué retardo resulta más efectiva la verificación, qué porcentaje de registros es necesario volver a introducir para asegurar el correcto funcionamiento de la VA, en qué grado las características personales y psicológicas de los sujetos pueden interactuar con el «feedback» que se facilita durante la verificación, o en qué grado se mejora la calidad del registro de los datos cuando

además de las técnicas de control utilizadas en esta experiencia se activa todo el sistema de protecciones automáticas posibles (validaciones, saltos automáticos, etc.).

Agradecimientos

Esta investigación se ha llevado a cabo gracias a las ayudas DGICYT PM95-126 y PM98-173 del Ministerio de Educación y Cultura.

Referencias

- Blumenstein, B.A. (1993). Verifying keyed medical research data. *Statistics in Medicine*, 12, 1535-1542.
- Butcher, J.N. (1994). Psychological assessment by computer: potential gains and problems to be avoided. *Psychiatric Annals*, 24, 20-24.
- Cobos, A. (1995). El síndrome GIGO. *JANO*, 49, 481-482.
- Doménech, J.M., Losilla, J.M., y Portell, M. (1998). La verificació aleatòria: una estratègia per millorar i avaluar la qualitat de l'entrada de dades. *QÜESTIO. Quaderns d'Estadística i Investigació Operativa*, 22, 493-510.
- Freedland, K.E., y Carney, R.M. (1992). Data management and accountability in behavioral and biomedical research. *American Psychologist*, 47, 640-645.
- Gassman, J.J., Owen, W.W., Kuntz, T.E., Martin, J.P., y Amoroso, W.P. (1995). Data quality assurance, monitoring and reporting. *Controlled Clinical Trials*, 16 (2 Suppl.), 104S-136S.
- Gibson, D., Harvey, A.J., Everett, V., y Parmar, M.K. (1994). Is double data entry necessary? The CHART trials: Continuous, Huperfractionated, Accelerated Radiotherapy. *Controlled Clinical Trials*, 15, 482-488.
- Granero, R. (1999). *Mejora de la calidad de la gestión de datos clínicos a través de sistemas computerizados. Aplicación a la informatización de la entrevista diagnóstica estructurada DICA-IV*. Tesis Doctoral. Universitat Autònoma de Barcelona.
- Grisso, T., Baldwin, E., Blanck, P.D., Rotheram-Borus, M.J., Schooler, N.R., y Thompson, T. (1991). Standards in research: APA's mechanism for monitoring the challenges. *American Psychologist*, 46, 758-766.
- Groves, R.M. (1989). *Survey errors and survey costs*. New York: John Wiley & Sons.
- ICH. (2000). International Conference on Harmonization. Dirección Internet: <http://www.eudra.org/humandocs/humans/ICH.htm>. (CPMP/ICH/363/96).
- Lewis, J.A., Jones, D.R., y Röhmel, J. (1995). Biostatistical methodology in clinical trials – A European Guideline. *Statistics in Medicine*, 14, 1655-1682.
- Magnusson, D., y Bergman, L.R. (1990). *Data quality in longitudinal research*. Cambridge: Cambridge University Press.
- Moritz, T.E., Ellis, N.K., Villanueva, C.B., Steeger, J.E., Ludwig, S.T., y Deegan, N.I. (1995). Development of an interactive data base management system for capturing large volumes of data. *Medical Care*, 33 (10 Suppl.).
- Redman, T.C. (1992). *Data quality. Management and technology*. New York: Bantam Books.
- Reich, W., Shayka, J., y Taibleson, Ch. (1991). *DICA-R DSM-III-R version*. Unpublished manuscript. Washington University, Division of Child Psychiatry. St. Louis.
- Rondel, R.K., Varley, S.A., y Weeb, C.F. (1999). *Clinical data management*. Chichester: John Wiley & sons.
- Rowley, G. (1989). Reliability and other indicators of data quality in classroom observation research. *Journal of Classroom Interaction*, 24, 22-29.
- Saris, W.E. (1991). *Computer-assisted interviewing*. Newbury Park: SAGE Publications.
- SPSS, Inc. (1999). *SPSS Base 9.0 Syntax Reference Guide*. Chicago, Illinois: SPSS, Inc.
- Wolf, R.M. (1993). Data quality and norms in international studies. Special Issue: Mandatory testing: Issues in policy-driven assessment. *Measurement and Evaluation in Counseling and Development*, 26, 35-40.

Aceptado el 11 de julio de 2000