

ESTUDIO DE LA EFICACIA DE LA TÉCNICA DE VERIFICACIÓN ALEATORIA COMO ALTERNATIVA A LA DOBLE ENTRADA DE LOS DATOS: VENTAJAS Y LIMITACIONES¹

Rosario Granero
José M^a Doménech
Albert Bonillo

Universitat Autònoma de Barcelona

RESUMEN

En este trabajo se plantean cuestiones básicas relativas al proceso de captura de datos en el ordenador. El trabajo comienza revisando las distintas formas en que puede llevarse a cabo el registro. A continuación se exponen las dificultades que conlleva la grabación de información contenida en formularios de papel. En primer lugar, la “entrada simple” (ES) no dispone de controles ni indicadores de calidad del registro, y por lo tanto no garantiza que la grabación sea correcta. Por otro lado, la “doble entrada” (DE) de la totalidad de los datos tiene el inconveniente de resultar muy costosa. En este trabajo se presenta una alternativa a la DE denominada “verificación aleatoria” (VA), que consiste en verificar un porcentaje del total de los registros escogidos al azar por el ordenador y presentados al operador con el suficiente retardo para evitar efectos de recuerdo. El objetivo general del estudio es valorar y comparar el grado de consistencia de los operadores cuando utilizan DE *versus* VA durante la captura de datos reales que previamente habían sido codificados en formularios de papel. Los resultados obtenidos indican que el *feedback* que el sistema ofrece al operador durante el proceso de la VA incrementa el grado de consistencia de los operadores en comparación con la DE. También se constata que la DE con un solo operador no garantiza la calidad del registro. En la discusión se plantea en qué situaciones y contextos de investigación resulta adecuado utilizar la VA como técnica de control.

Palabras clave: gestión de datos; captura de datos; calidad de datos; doble entrada; verificación aleatoria.

¹ Esta investigación se ha llevado a cabo gracias a las ayudas DGICYT PM95-126 y PM98-173 del Ministerio de Educación y Cultura

Introducción

La gestión de datos (“data management”) constituye un conjunto de etapas secuenciales cuyo objetivo prioritario es el manejo de la información para obtener conocimiento sobre el modo en que los fenómenos acontecen en la realidad y las relaciones funcionales que se establecen entre ellos. Las fases principales que definen este proceso son (Rondel, Varley y Webb, 1999):

1. Diseño de la base de datos a partir de las variables a registrar.
2. Captura de la información en el ordenador.
3. Preparación de la matriz de datos para el análisis estadístico (definición y generación de variables).
4. Análisis estadístico.

Durante las últimas décadas, la mayoría de investigadores y revisores sólo se han preocupado de que los datos que se publican estén correctamente analizados. Actualmente, sin embargo, diversos estudios advierten de la complejidad que conlleva el adecuado desarrollo de cualquiera de las fases de la gestión de datos, y alertan sobre los numerosos errores que se pueden cometer cuando no se planifican y utilizan controles rutinarios de calidad en cualquiera de estas etapas (Saris, 1991; Redman, 1992). En estas investigaciones también se indica que la introducción y propagación de estos errores puede tener consecuencias devastadoras sobre la validez de los estudios, y que el mero uso de ordenadores no mejora la calidad de los resultados finales que se publican (Butcher, 1994). Sin embargo, a pesar de estas serias advertencias, en muy pocos ámbitos de la investigación aplicada se adoptan sistemas de control de calidad para garantizar la adecuada gestión de datos, probablemente por falta de recursos y ausencia de legislación. Sólo en el caso de los estudios que evalúan productos médicos existe una normativa muy estricta y los datos son sometidos a una auditoría (Smith y Heywood, 1999; ICH, 2000a, 2000b, 2000c).

Dentro del contexto general de la gestión de datos, la captura de la información en el ordenador constituye una etapa crítica que no siempre ha gozado de la atención que merece. Recientemente están apareciendo distintos trabajos donde se revisan las fuentes de error que pueden atentar contra la calidad de la grabación, y donde se exponen los procedimientos de control que se pueden utilizar para asegurar que los datos que posteriormente serán analizados estadísticamente están exentos de errores (Groves, 1989; Granero, 1999). La Figura 1 inserta la gestión de datos en el proceso general de investigación científica. En la parte sombreada de esta ilustración se destaca la importancia de utilizar protecciones exhaustivas y, si son incompletas, se debe activar un proceso de depuración para garantizar la calidad de la información contenida en la base de datos de un estudio. Es importante que esta base de datos haya sido construida para contener todas las variables definidas en la etapa del diseño de la investigación, así como también un “archivo histórico” donde se recojan

de manera que en cualquier momento sea posible disponer de los valores previos a cualquier modificación (“role-back”). Sólo así es posible evaluar y asegurar la calidad de los datos de la investigación.

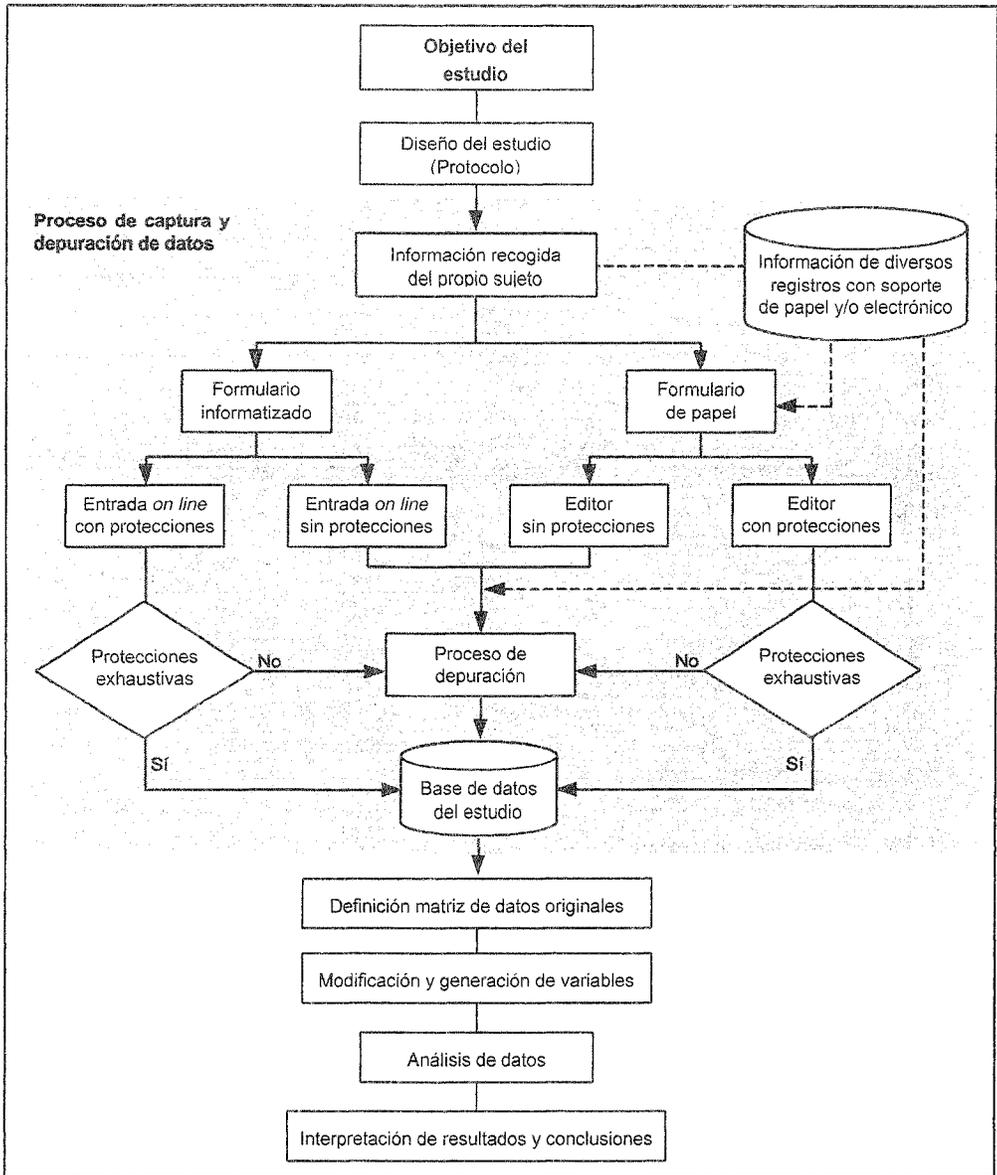


Figura 1. El proceso de entrada y depuración de datos dentro del contexto general de investigación científica.

En numerosos ámbitos de las Ciencias Sociales y de la Salud aún es frecuente recoger la información mediante formularios administrados en papel que posteriormente se graban en el ordenador. En demasiadas ocasiones este registro se realiza mediante “entrada simple” (ES), pero este sistema adolece de un problema fundamental: no dispone de controles ni indicadores de la calidad del registro y, por lo tanto, no puede garantizar que la grabación sea correcta. Una de las alternativas más comunes, y considerada más eficaz, para garantizar la captura ha sido la “doble entrada” (DE). Este procedimiento consiste en grabar dos veces la totalidad de los registros, ya sea mediante un único operador (encargado de grabar la misma serie de datos en dos ocasiones distintas) o con dos operadores (cada uno de los cuales graba en una ocasión la misma serie de datos). La DE con dos operadores es considerada la modalidad de elección, especialmente cuando la información que se registra está contenida en formularios de papel que pueden dificultar su correcta interpretación (Blumenstein, 1993; Gassman y otros, 1995). Tras el doble registro de los datos se procede a su comparación automática, y se corrigen las eventuales discrepancias aparecidas entre ambas capturas. En esta fase, se recomienda que sea un tercer sujeto, estrechamente vinculado con el estudio, el responsable último de resolver las discordancias (Stellman, 1989; Gibson, Harvey, Everett y Parmar, 1994). En cuanto a la eficacia de la DE, distintos estudios constatan que esta técnica puede reducir hasta en un 30% la proporción de inconsistencias respecto a la ES (Neaton, Duchene, Svendsen y Wentworth, 1990; Reynolds-Haertle y McBride, 1992; Wolf, 1993). En cuanto a sus dificultades, la DE constituye una técnica de control muy costosa, que sólo puede ser empleada cuando se dispone de suficientes recursos económicos.

Debido al coste de la DE, se han propuesto técnicas alternativas en las cuales unas variables son verificadas de forma más intensiva y exhaustiva que otras. Por ejemplo, Mullooly (1990) y Neaton y otros (1990) proponen el uso de sistemas de “verificación selectiva”, caracterizados por grabar con DE aquellos ítems considerados cruciales y con ES el resto de los datos. En este trabajo presentamos otra técnica alternativa denominada “verificación aleatoria” (VA), que consiste en verificar un porcentaje del total de los registros escogidos al azar por el ordenador y presentados al operador con el suficiente retardo para evitar efectos de recuerdo. La técnica de la VA requiere que sea necesariamente el mismo operador quien introduzca los datos, pero cuenta con una ventaja importante respecto a la DE: reduce considerablemente el coste de la grabación. Además, la VA también ofrece un índice de las discrepancias cometidas durante la captura que puede ser utilizado para controlar el grado de consistencia de los operadores durante la captura, y también para estimar el porcentaje de errores residuales contenidos en la base de datos del estudio. En una reciente investigación realizada con datos ficticios se ha demostrado que la VA, en comparación con la ES, aumenta la eficacia del registro sin afectar la eficiencia (Doménech, Losilla y Portell, 1998). La explicación más plausible de esta eficacia es el *feedback* que el sistema proporciona al operador durante el proceso de grabación.

El objetivo general de este trabajo es valorar la utilidad de la VA (en ausencia de otros controles automáticos) para la captura de datos obtenidos de entrevistas diagnósticas reales que previamente habían sido codificadas en formularios de papel por psicólogos clínicos. Los objetivos específicos son: 1) valorar el grado de consistencia de los operadores durante la captura de datos cuando se utiliza la VA frente a la DE; y 2) discutir en qué situaciones y/o contextos tiene interés el uso de la VA.

Método

Sujetos

Para realizar este estudio se dispuso de la colaboración de dos operadores profesionales en la grabación de datos. El primero se encargó de la captura de todos los formularios, una parte con DE y el resto con VA. El segundo operador realizó una nueva grabación independiente de los registros sometidos a DE. Esta segunda figura es necesaria porque permite comparar la DE con una interpretación y con dos interpretaciones independientes de los formularios en papel.

Las condiciones de inclusión de los operadores consistieron en un mínimo de un 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 disponía de 335 entrevistas diagnósticas estructuradas DICA-R (Reich, Shay-ka y Taibleson, 1991), instrumento de evaluación psicológica caracterizado por poseer muchas preguntas filtro (de salto), al igual que numerosas encuestas. Estos datos habían sido codificados previamente a mano en formularios de papel por entrevistadores entrenados en el uso de los protocolos. La información diagnóstica correspondía a entrevistas realizadas entre 1993 y 1996 a familias que acudían a servicios de Psicología y Psiquiatría infantil de hospitales públicos de Barcelona, y a familias de niños escolarizados de la población general.

Para llevar a cabo la captura se creó un formulario informatizado con el Sistema DAT, que permite implementar las técnicas de VA y DE (Doménech y Losilla, 1995). Este formulario constaba de 983 campos y carecía de protecciones automáticas (no se definieron validaciones ni saltos), de modo que permitía a los operadores grabar cualquier valor en cualquier campo siempre y cuando éste fuera consistente con el formato definido. Para facilitar el registro de la información, el diseño de las pantallas se realizó reproduciendo lo más fielmente posible el aspecto del formulario en papel.

Diseño

El diseño realizado se esquematiza en la Figura 2. La variable independiente es la condición de captura, y tiene tres niveles: ES, VA y DE. Mediante ES se grabaron

inicialmente 153 entrevistas, que después serían vueltas a grabar (DE). Las restantes 182 entrevistas fueron sometidas a VA, que se activó para una recaptura igual al 20% de los formularios con un retardo en la segunda entrada igual a 2 formularios. Durante la VA 39 entrevistas fueron capturadas en dos ocasiones y 143 en una única ocasión.

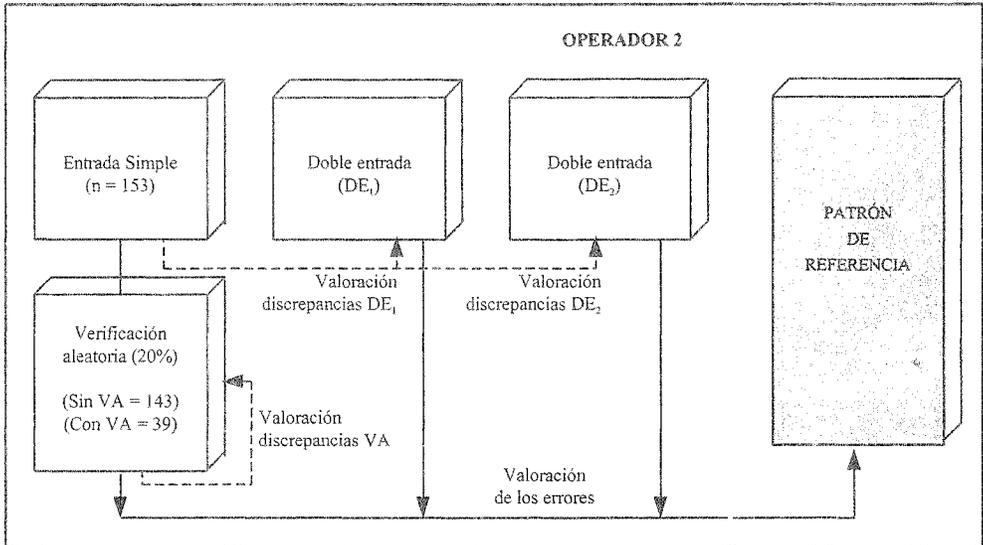


Figura 2: Esquema del procedimiento de captura de datos.

La ES se ha caracterizado por una consigna verbal informando al operador de la importancia de llevar a cabo una captura correcta de los datos. La condición de VA se ha caracterizado por una consigna verbal informando al operador que el programa controlaría las discrepancias que iba cometiendo durante la grabación obligándole a repetir la introducción de algunos registros; también se le informó que ante cada discrepancia el sistema presentaría un mensaje de aviso y le solicitaría la verificación del correspondiente campo. La condición DE se ha caracterizado por una consigna verbal informando que el programa volvería a solicitar la introducción de los registros de forma similar a la VA.

La variable dependiente principal es el total de discrepancias producidas durante la segunda captura de los registros, que se han definido como la falta de coincidencia al grabar una misma serie de valores en dos ocasiones diferentes. Se suele considerar que este indicador es una medida indirecta de la frecuencia de errores que cometen los operadores durante la captura de los datos. En este sentido, 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 fue 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.

Además de las discrepancias, en este estudio también se disponía de los errores reales, definidos como la falta de coincidencia entre los datos grabados por los operadores y la información registrada en los formularios en papel. Esta información estaba representada en el patrón o entrada de referencia (Figura 2). 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.

Finalmente, puesto que la DICA-R posee estructura ramificada en forma de árbol lógico que hace que sólo se complete una parte de la información diferente en cada entrevista, las variables que miden las discrepancias y los errores se han definido en función de cómo aparecían realmente los campos en el patrón de referencia (completos o vacíos).

Procedimiento

En un principio, 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 estructuró en dos etapas. La primera fase, de entrenamiento, tuvo como objetivos mostrar el *software* de entrada de datos a los operadores y familiarizarlos con los formularios en papel. Consistió en presentarles el programa DAT, ayudarles a practicar con él la captura de 10 cuestionarios de prueba y enseñarles a realizar las copias de seguridad. La segunda fase, experimental, consistió en presentar a los operadores las entrevistas que debían ser capturadas en orden aleatorio, incentivando la rapidez en la captura. Cada sesión duraba unas 4 horas, y los operadores podían realizar descansos regulares según su propia necesidad. El orden de captura de la fase experimental queda ilustrado en la Figura 2. Es importante destacar que el primer operador no fue informado al inicio de que se realizaría una DE de los formularios.

El patrón o entrada de referencia utilizado en este estudio había sido creado a través de un sistema de filtros sucesivos (Granero y Doménech, en prensa). En concreto, tras finalizar la captura el segundo operador procedió a grabar la totalidad de los formularios en papel ($n=335$), resolviendo las discrepancias que se iban produciendo respecto a los datos ya grabados. Posteriormente, tres nuevos evaluadores de nuestro grupo de investigación familiarizados con la entrevista DICA-R realizaron tres nuevas comprobaciones independientes de todos los registros.

El análisis estadístico se ha efectuado con el Sistema SPSS 10.0.6 para Windows (SPSS, 1999). El procedimiento de análisis ha consistido en la obtención de proporciones de discrepancias (que en este estudio representan “incidencias acumuladas”), y en el cálculo de la diferencia absoluta de proporciones y la razón entre las incidencias de discrepancias —también denominada “riesgo relativo” (Kleinbaum, Kupper y Morgenstern, 1982).

Resultados

Los resultados obtenidos al valorar las discrepancias que se producen aparecen en la Tabla 1. La parte superior de esta tabla sugiere que, independientemente del contenido del patrón (campo completo o vacío), 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 también más discrepancias que la VA.

Tabla 1: *Proporciones de discrepancias cometidas por los operadores en tantos por mil (entre paréntesis figuran los intervalos de confianza del 95%).*

Condición de entrada	Patrón vacío	Patrón con valor	TOTAL
Doble entrada (un operador)	4.1 (3.7 a 4.5)	48.7 (46.7 a 50.8)	17.1 (16.4 a 17.8)
Verificación Aleatoria	1.9 (1.4 a 2.4)	38.7 (34.6 a 42.7)	10.9 (9.8 a 11.9)
Doble entrada (dos operadores)	7.1 (6.6 a 7.6)	71.0 (68.5 a 73.5)	25.7 (24.9 a 26.6)

Feedback VA: DE ₁ ¹ versus VA	Patrón vacío	Patrón con valor	TOTAL
Diferencia (%)	2.2 (1.5 a 2.8)	10.1 (5.5 a 14.7)	6.2 (5.0 a 7.5)
Riesgo relativo	2.2 (1.6 a 2.9)	1.3 (1.1 a 1.4)	1.6 (1.4 a 1.7)

Feedback VA: DE ₂ ² versus VA	Patrón vacío	Patrón con valor	TOTAL
Diferencia (%)	5.2 (4.5 a 6.0)	32.3 (27.6 a 37.1)	14.9 (13.5 a 16.2)
Riesgo relativo	3.8 (2.8 a 5.0)	1.8 (1.6 a 2.1)	2.4 (2.2 a 2.6)

Doble interpret.: DE ₂ versus DE ₁	Patrón vacío	Patrón con valor	TOTAL
Diferencia (%)	3.1 (2.4 a 3.7)	22.3 (19.0 a 25.5)	8.6 (7.6 a 9.7)
Riesgo relativo	1.7 (1.5 a 2.0)	1.5 (1.4 a 1.5)	1.5 (1.4 a 1.6)

¹ DE realizada por el mismo operador ; ² DE realizada por un operador diferente

Las comparaciones realizadas sobre el total de campos verificados indican que la DE con un operador incrementa, respecto a la VA, la proporción (absoluta) de discrepancias ($d=6.2\%$, IC 95%: 5 a 7.5), y también el riesgo de cometerlas (RR=1.6, IC 95%: 1.4 a 1.7). La DE con dos operadores incrementa, respecto a la VA, la proporción de discrepancias un 14.9% (IC 95%: 13.5 a 16.2), y el riesgo de cometerlas se multiplica por 2.4 (IC 95%: 2.2 a 2.6). Finalmente, la DE con doble interpretación de la información aumenta, respecto a la DE con un único operador, la proporción de discrepancias en un 8.6% (IC 95%: 7.6 a 9.7), y el riesgo de cometerlas se multiplica por 1.5 (IC 95%: 1.4 a 1.6).

Finalmente, la comparación de los datos grabados por los operadores con el patrón de referencia indica que (Granero y Doménech, en prensa): la DE con un operador obtiene un 8.5% de errores finales (IC 95%: 8 a 8.9), la DE con dos operadores un 2.6% (IC 95%: 2.4 a 2.9) y la VA un 3.4% (IC 95%: 3.1 a 3.7).

Conclusión

Los resultados obtenidos en esta experiencia sugieren que la VA de un porcentaje de los registros contribuye a incrementar el grado de consistencia de los operadores durante la captura de datos en comparación con la DE. La condición de captura con menor grado de discrepancias se obtiene al activar la VA, y la menos consistente cuando se realiza una DE con dos interpretaciones.

Uno de los aspectos más llamativos de los resultados obtenidos en este estudio es la superioridad que aparenta la DE con una única interpretación frente a la DE con dos interpretaciones. En este sentido, es importante tener presente que la consistencia de la grabación se ha valorado en función del número de discrepancias, y no del total de errores que persisten tras la captura. Podría suceder que un valor alto de discordancias correspondiera a una proporción baja de errores finales. Para comprobar esta posible eventualidad hemos valorado las proporciones reales de errores contenidos en los datos al final de la grabación, y hemos constatado que la VA (con una proporción de errores del 3.4 %) ofrece un rendimiento similar a la DE con dos operadores distintos (cuya proporción de errores es del 2.6 %) y resulta mucho más eficaz que la DE con un único operador (con un 8.5 % de errores). Esta conclusión es consistente con los resultados obtenidos por Doménech y otros (1998) en un estudio experimental en el que se grabaron datos ficticios. Una de las explicaciones más plausibles de estos resultados es el hecho de que la VA ofrece un *feedback* permanente al operador, en tanto que la DE sólo lo ofrece al operador que realiza la segunda captura.

Discusión

La técnica de VA contemplada en abstracto podría verse como un procedimiento similar a la DE realizada por un mismo operador, con la diferencia de que sólo se recapturan un porcentaje del total de los registros. De acuerdo a este planteamiento, si la DE con una única interpretación se considera una técnica poco adecuada para garantizar la calidad del registro (Blumenstein, 1993; Gassman y otros, 1995), la VA todavía sería menos eficaz porque no se vuelven a grabar la totalidad de los datos. Los resultados obtenidos en este estudio, sin embargo, no confirman esta sospecha, sino que sugieren que la VA es un procedimiento útil para mejorar la calidad del registro. Estos resultados se explican si se tiene en cuenta que la VA es algo más que la mera reentrada de un conjunto de formularios por un mismo operador: el operador sabe que está siendo controlado y obtiene *feedback* inmediato y constante sobre la consistencia del registro de datos durante todo el proceso de gra-

bación. Es precisamente este *feedback* el que actúa incrementando la atención de los operadores.

Constatar que la VA resulta útil para mejorar la calidad del registro de datos implica valorar cuáles son los ámbitos de aplicación de esta técnica. En líneas generales, el uso de la VA se justifica cuando no se dispone de recursos suficientes para realizar una DE, por ejemplo cuando se dispone de grandes volúmenes de datos. También puede resultar interesante implementar la VA de forma compatible con la DE en tres situaciones distintas: a) en aquellos estudios que obtuvieran proporciones de discrepancias elevadas para la VA se aconsejaría realizar una DE de los registros con objeto de reducir la proporción de errores residuales; b) parece adecuado utilizar la VA durante la primera captura de los datos y completar la grabación con la DE del resto de los registros, ya que de esta forma la VA permitiría disponer de una primera entrada con menor proporción de errores que mejoraría el rendimiento final de la DE; y c) ambas técnicas se podrían utilizar simultáneamente en estudios con operadores, registros o campos que requirieran diferentes niveles de control. Esta última situación se concreta, por ejemplo, en la grabación de todos los partes de accidentes de una comunidad para efectuar la estadística anual de accidentabilidad laboral: los accidentes graves y mortales requieren DE con operadores distintos, pero el resto de partes podrían ser capturados mediante VA si no se dispone de suficientes recursos. Por otro lado, simultanear la DE con la VA cuando se dispone de operadores que requieren distinto nivel de control resulta relevante si se tienen en cuenta los resultados obtenidos en investigaciones que demuestran una gran variabilidad en la eficacia de los distintos operadores encargados de entrar los datos en un estudio (Reynolds-Haertle y McBride, 1992). Finalmente, también son sugerentes los resultados obtenidos por Gibson y otros (1994), que atribuyeron gran parte de la mejora en la grabación de los datos de su estudio (respecto a los porcentajes de incongruencias habitualmente referenciadas en la bibliografía) a la disponibilidad de operadores que habían sido entrenados en el uso de los formularios, que eran responsables de otros aspectos cruciales de la investigación, conocedores de las consecuencias de una captura deficiente y, por lo tanto, concienciados de la importancia de un registro correcto de los datos.

Por otro lado, establecer que la VA resulta eficaz para mejorar la calidad del registro conlleva un conjunto adicional de interrogantes sobre qué características psicológicas de los sujetos pueden interactuar con el *feedback* que facilita esta técnica, o cómo debe llevarse ésta a cabo desde el punto de vista procedimental para garantizar su funcionalidad. Con relación a este último punto, es imprescindible realizar nuevos estudios más ecológicos que el aquí presentado para determinar qué tipo de mensaje/s debe presentar el ordenador frente a una discrepancia, quién debe ser el encargado de resolverla y con qué secuencia de acciones, con qué retardo resulta más efectiva la verificación, y, por supuesto, qué porcentaje de registros es necesario volver a grabar para asegurar el correcto funcionamiento de esta estrategia.

Las consideraciones expuestas en el párrafo anterior alertan sobre la necesidad de contextualizar los resultados obtenidos en esta experiencia, que en ningún caso

pueden interpretarse como determinantes. Por ejemplo, en relación a las consignas facilitadas, el primer operador (Figura 2) no fue advertido al inicio de la captura de que los datos serían sometidos a DE. Si se le hubiera indicado que el sistema ejercería un control sobre las discrepancias cometidas durante ambas entradas y se hubiera implementado un incentivo de calidad de la grabación tal vez se hubiera conseguido una mayor reducción de los errores obtenidos para la DE con una y/o dos interpretaciones.

También quisiéramos destacar que, probablemente, parte de las discrepancias producidas durante la grabación de los datos de esta experiencia son consecuencia del tipo de registro original. Resulta evidente que cuando los datos se codifican a mano la interpretación por parte del operador se dificulta y se hace más libre que cuando se dispone de texto mecanografiado. Esto podría suponer una dificultad añadida para el uso de la VA, ya que cuando el registro resulta difícil de interpretar el operador tenderá a acentuar su atención durante la primera captura, lo cual podría contribuir a que durante la segunda entrada de la verificación recordase el valor que había grabado. Esta posible eventualidad, que reduciría de forma artificiosa el porcentaje de discordancias, sólo puede ser contrarrestada incrementando el retardo con el que se efectúa la VA.

En relación a lo anterior, resultaría muy práctico utilizar programas de captura que dispusieran de un atributo especial para identificar los campos que han sido difíciles de leer y/o interpretar por parte del operador encargado de la grabación, con objeto de contrastarlos posteriormente. Una discusión más detallada sobre las características que debe reunir el software para la captura de datos se puede encontrar en Granero y Doménech (2000).

Como reflexión de este trabajo también quisiéramos puntualizar que, a diferencia de una creencia muy popular (Blumenstein, 1993; Wolf, 1993; Gibson y otros, 1994; Gassman y otros, 1995), la DE no garantiza por sí misma la calidad del registro, y mucho menos que los datos estén libres de errores antes de proceder a su análisis estadístico (especialmente cuando se utilizan instrumentos de evaluación con estructura compleja, como muchas encuestas ramificadas). Por otro lado, aunque fuéramos capaces de garantizar (mediante VA o DE) que los datos capturados coinciden con el registro en papel no se eliminaría un porcentaje importante de errores que ya estaban contenidos en el registro original. En efecto, en ausencia de otros controles, no se puede descartar que además de los errores producidos como consecuencia de una captura deficiente, los datos posean otras inconsistencias debidas a que los entrevistadores originales hayan efectuado saltos incorrectos, hayan olvidado el registro de algunas variables o hayan codificado valores que no cumplen las condiciones de validación de los correspondientes campos. En este caso, se hace imprescindible activar todo un sistema de protecciones automáticas para detectar las inconsistencias de las fuentes documentales originales, impidiendo así los errores invisibles a la DE y la VA. El uso de estas protecciones automáticas es imprescindible cuando se manejan grandes volúmenes de datos, se trabaja con estructuras de captura de la información complejas o se utilizan instrumentos de evaluación que

requieren que los entrevistadores tomen decisiones sobre su manejo al aplicarlos a cada caso. Es frecuente que las empresas de grabación utilicen procedimientos de captura basados en la doble entrada simultaneada con el uso de protecciones automáticas. Es preciso, pues, realizar nuevos estudios para valorar cómo se mejora en estos casos la calidad final del registro.

En relación a lo dicho anteriormente, y como última reflexión, recordar que tanto el conjunto de protecciones automáticas como el proceso global de captura y depuración de datos (parte sombreada de la Figura 1) deben ser planificados durante el diseño del estudio.

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 avoid. *Psychiatric Annals*, 24, 20–24.
- Doménech, J.M., y Losilla, J.M. (1995). El Sistema DAT: gestor de datos científicos. Manual de referencia. Bellaterra: Laboratori d'Estadística Aplicada i de Modelització. Universitat Autònoma de Barcelona.
- 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.
- 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 doble 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.
- Granero, R., y Doménech, J.M. (2000). Problems related to the administration of structured interviews on paper forms: software requirements for assuring the quality of on line clinical data capture. Manuscrito en revisión editorial.
- Granero, R., y Doménech, J.M. (en prensa). Captura de datos clínicos con verificación aleatoria: una nueva técnica para controlar y mejorar la calidad del registro. *Psicothema*.
- Groves, R.M. (1989). *Survey errors and survey costs*. New York: John Wiley & Sons.

- ICH (International Conference on Harmonization) (2000a). Guidance on Statistical Principles for Clinical Trials. Acceso: <http://www.eudra.org/humandocs/humans/ICH.htm>. (CPMP/ICH/363/96).
- ICH (International Conference on Harmonization) (2000b). Note for Guidance on Clinical Safety Data Management: Periodic Reports for Marketed Drugs. Acceso Internet: <http://www.eudra.org/humandocs/humans/ICH.htm>. (CPMP/ICH/288/95).
- ICH (International Conference on Harmonization) (2000c). Note for Guidance on Good Clinical Safety Data Management: The Standards for Expedited Reporting. Acceso Internet: <http://www.eudra.org/humandocs/humans/ICH.htm>. (CPMP/ICH/377/95).
- Kleinbaum, D.G., Kupper, L.L., y Morgenstern, H. (1982). *Epidemiologic Research. Principles and Quantitative Methods*. Belmont: Van Nostrand Reinhold.
- Mullooly, J.P. (1990). The effects of data entry error: an analysis of partial verification. *Computers and Biomedical Research*, 23, 259–267.
- Neaton, J.D., Duchene, A.G., Svendsen, K.H., y Wentworth, D. (1990). An examination of the efficiency of some quality assurance methods commonly employed in clinical trials. *Statistics in Medicine*, 9, 115–124.
- 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.
- Reynolds-Haertle, R.A., y McBride, R. (1992). Single vs doble data entry in CAST. *Controlled Clinical Trials*, 13, 487–494.
- Rondel, R.K., Varley, S.A., y Weeb, C.F. (Eds.) (1999). *Clinical data management*. 2nd Edition. Chichester: John Wiley & Sons.
- Saris, W.E. (1991). *Computer-assisted interviewing*. Newbury Park: SAGE.
- Smith, B., y Heywood, L. (1999). The International Conference on Harmonisation and its Impact. En R.K. Rondel, S.A. Varley y C.F. Weeb (Eds.), *Clinical data management* (pp. 21-46). Chichester: John Wiley & Sons.
- Stellman, S.D. (1989). Brief reports. The case of the missing eights. An object lesson in data quality assurance. *American Journal of Epidemiology*, 129, 857–860.
- SPSS, Inc. (1999). *SPSS Base 10.0 Syntax Reference Guide*. Chicago, Illinois: Author.
- 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.

