



Revista de Calidad Asistencial

www.elsevier.es/calasis



ORIGINAL

Identificación de eventos adversos en pacientes hospitalizados por gripe



J.M. Aranaz-Andrés^{a,b,c,*}, M.T. Gea-Velázquez de Castro^{b,c,d}, F. Jiménez-Pericás^d,
A.I. Balbuena-Segura^d, M.C. Meyer-García^d, N. López-Fresneña^a,
J.J. Miralles-Bueno^c, B. Obón-Azuara^{e,f}, J. Moliner-Lahoz^{e,f} y C. Aibar-Remón^{e,f}

^a Servicio de Medicina Preventiva y Salud Pública, Hospital Universitario Ramón y Cajal, Madrid, España

^b Centro de Investigación Biomédica en Red en Epidemiología y Salud Pública (CIBERESP), España

^c Departamento de Salud Pública, Historia de la Ciencia y Ginecología, Universidad Miguel Hernández, Sant Joan d'Alacant, Alicante, España

^d Servicio de Medicina Preventiva y Calidad Asistencial, Hospital Universitari Sant Joan d'Alacant, Sant Joan d'Alacant, Alicante, España

^e Servicio de Medicina Preventiva y Salud Pública, Hospital Clínico Universitario Lozano Blesa, Zaragoza, España

^f Instituto de Investigación Sanitaria de Aragón (IIS), Zaragoza, España

Recibido el 29 de julio de 2014; aceptado el 7 de enero de 2015

Disponible en Internet el 7 de febrero de 2015

PALABRAS CLAVE

Eventos adversos;
Seguridad del
paciente;
Fiabilidad

Resumen

Objetivos: Testar la fiabilidad interobservador de la guía de cribado y del formulario modular de revisión (MRF2) para identificar eventos adversos (EA). Comparar la incidencia de EA entre un grupo de pacientes sometidos a precauciones de aislamiento de contacto y gotas por gripe frente a un grupo de pacientes no aislados.

Método: Estudio de casos y controles anidados en una cohorte. Los casos son 50 pacientes sometidos a precauciones de aislamiento por gripe, y los controles son 50 pacientes ingresados y no aislados.

Resultados: La incidencia de EA varía entre el 10 y el 26% por efecto del observador (26% [IC 95%: 17,4-34,60%]; 10% [IC 95%: 4,12-15,88%]; 23% [IC 95%: 14,75-31,25%]). Es siempre menor en la cohorte expuesta al aislamiento que en la no expuesta, y esa diferencia alcanza significación estadística cuando aplicamos el criterio de máxima exigencia en la definición de caso.

La concordancia respecto del cribado fue buena (porcentaje simple de concordancia superior al 76%; índice Kappa entre 0,29 y 0,81).

La concordancia respecto de la identificación precisa de EA relacionado con la asistencia fue menor (porcentaje de concordancia entre el 50 y el 93,3%; índice Kappa de 0,20 a 0,70).

Conclusiones: Antes de realizar un estudio epidemiológico sobre EA hay que analizar la fiabilidad interobservador para mejorar la precisión de los resultados y, por ende, la validez del estudio.

Los revisores presentan diferentes grados de concordancia. El índice Kappa muestra niveles elevados para la guía de cribado, no así para la identificación de EA.

* Autor para correspondencia.

Correo electrónico: jesusmaria.aranaz@salud.madrid.org (J.M. Aranaz-Andrés).

KEYWORDS

Adverse events;
Patient safety;
Reliability

Sin una buena metodología no se pueden garantizar los resultados alcanzados y, consecuentemente, las decisiones tomadas a partir de ellos. Por tanto, los investigadores han de ser garantes del método utilizado y este debe acercarse lo más posible al óptimo alcanzable.

© 2014 SECA. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

Identification of adverse events in hospitalised influenza patients

Abstract

Objective: To test the inter-observer agreement in identifying adverse events (AE) in patients hospitalized by flu and undergoing precautionary isolation measures.

Methods: Historical cohort study, 50 patients undergoing isolation measures due to flu, and 50 patients without any isolation measures.

Results: The AE incidence ranges from 10 to 26% depending on the observer (26% [95% CI: 17.4%-34.60%], 10% [95% CI: 4.12%-15.88%], and 23% [95% CI: 14.75%-31.25%]). It was always lower in the cohort undergoing the isolation measures. This difference is statistically significant when the accurate definition of a case is applied.

The agreement as regards the screening was good (higher than 76%; Kappa index between 0.29 and 0.81).

The agreement as regards the accurate identification of AE related to care was lower (from 50 to 93.3%, Kappa index from 0.20 to 0.70).

Conclusions: Before performing an epidemiological study on AE, interobserver concordance must be analyzed to improve the accuracy of the results and the validity of the study.

Studies have different levels of reliability. Kappa index shows high levels for the screening guide, but not for the identification of AE.

Without a good methodology the results achieved, and thus the decisions made from them, cannot be guaranteed. Researchers have to be sure of the method used, which should be as close as possible to the optimal achievable.

© 2014 SECA. Published by Elsevier España, S.L.U. All rights reserved.

Introducción

Los estudios epidemiológicos sobre eventos adversos relacionados con la asistencia sanitaria (EARA) son el punto de partida para conocer su frecuencia y poder implementar prácticas seguras que minimicen su impacto en los pacientes.

En España se han desarrollado varias iniciativas y estudios para determinar la frecuencia de EARA. El primero de ellos, el estudio ENEAS, puso de manifiesto que la incidencia acumulada de efectos adversos en nuestro país en 2005 fue del 9,3%. De todos ellos, el 37,4% de los EARA están relacionados con la medicación, el 25,3% con la infección relacionada con la asistencia sanitaria (IRAS) y el 25% con problemas técnicos durante un procedimiento.

La definición de EARA utilizada en Estados Unidos y Australia en estudios de cohortes retrospectivas con revisión de historias clínicas fue la de «daño no intencionado provocado por un acto médico más que por el proceso nosológico en sí»¹.

Sin embargo, el estudio ENEAS utilizó una definición operativa más precisa y acorde con la práctica clínica, considerando como EARA al conjunto de efectos e incidentes adversos. El efecto adverso supondrá una lesión o complicación que prolongue la estancia hospitalaria, que precise procedimientos diagnósticos o terapéuticos adicionales, o

que esté relacionado con el fallecimiento o la incapacidad al alta hospitalaria mientras que el incidente no tiene repercusión en el paciente^{2,3}.

El segundo EARA más frecuente es la IRAS, según el estudio ENEAS⁴. El estudio de Prevalencia de Infección Nosocomial en España (EPINE)⁵, que se realiza desde hace más de 2 décadas, y del *Point Prevalence Study*, efectuado en diversos países de Europa durante el año 2010⁶, estableció que alrededor del 7% de los pacientes hospitalizados presentan una IRAS⁷. Estas comportan una elevada morbilidad, un mayor coste económico y, según el estudio ENEAS, son evitables en un 56,6%⁸. Los principales tipos de IRAS son la infección respiratoria, la quirúrgica, la urinaria y la bacteriemia de catéter vascular. Esta distribución es similar a la que ya se había puesto de manifiesto en otros estudios semejantes⁵.

Para reducir las IRAS existen recomendaciones internacionales de prevención de las mismas. Desde 1996, los Centros para el Control y Prevención de Enfermedades (CDC) recomiendan añadir a las precauciones estándar las precauciones específicas basadas en la transmisión (contacto, aéreas o gotas) en pacientes en los que se sabe o se sospecha que están infectados o colonizados con patógenos importantes desde el punto de vista epidemiológico. Las precauciones basadas en la transmisión o «aislamiento» deben aplicarse ante la sospecha clínica de infección, y no solo

tras la confirmación microbiológica, ya que de esta forma se minimiza el riesgo de transmisión a otros pacientes. En el caso de sospecha de gripe se requiere adoptar precauciones de transmisión por contacto y gotas, que incluyen: ubicación en habitación individual, colocación de bata, guantes y mascarilla quirúrgica al entrar a la habitación, además de las precauciones estándar. Sin embargo, la adopción de estas precauciones puede suponer un incremento de eventos adversos para el paciente; en concreto, se han descrito mayor puntuación en las escalas de ansiedad y depresión en pacientes aislados, menor frecuencia de atención por parte de los profesionales y menor satisfacción con la atención recibida⁹.

En las últimas décadas, los estudios para detectar EARA han utilizado diferentes enfoques, los métodos basados en la revisión retrospectiva de historias clínicas han sido los más comunes, siendo la referencia (*gold standard*) el *Harvard Medical Practice Study*, que fue un estudio pionero en el que los registros médicos eran sometidos a cribado para detectar indicadores de EARA, seguido esto de una revisión confirmatoria¹⁰. De acuerdo con esta metodología, los criterios de detección son «marcadores de EARA» que son detectados en una fase de cribado, usualmente realizada por enfermeras, que serán confirmados o no en una segunda fase de revisión, realizada por médicos. Dos estudios en el ámbito iberoamericano realizaron una adaptación transcultural de la metodología¹¹⁻¹³.

Para poder identificar los EARA en la historia clínica de los pacientes se hace necesario disponer de una guía de cribado validada que pueda ser utilizada por personal entrenado en esta, así como un formulario de revisión que nos permita después realizar una valoración exhaustiva del EARA y valorar su grado de relación con la asistencia sanitaria, su gravedad y su evitabilidad¹⁴⁻¹⁶. Además, la clave para mejorar la validez y la precisión de los estudios epidemiológicos que utilizan como fuente de datos la historia clínica está en la capacitación de los revisores; es necesario que estén entrenados y tengan criterios comunes, dado que es imprescindible reducir la variabilidad interobservador.

El objetivo principal de este estudio es testar la fiabilidad de la Guía de Cribado de Eventos Adversos y el Formulario Modular de Revisión (MRF2), del Proyecto IDEA, para la identificación de EARA en pacientes con gripe sometidos a precauciones basadas en la transmisión por gotas y contacto.

Como objetivo secundario se propone evaluar la diferencia de EARA en el grupo de pacientes aislados por gripe frente a un grupo control de pacientes no aislados por gripe.

Método

Diseño del estudio: estudio de casos y controles anidados en una cohorte. La cohorte la formaron los pacientes ingresados en el Hospital San Juan de Alicante durante el año 2.013. Los casos fueron 50 pacientes sometidos a precauciones de aislamiento por gotas y contacto por gripe, y los controles, otros 50 pacientes ingresados y no aislados.

Se seleccionó un control (paciente no aislado) por cada caso (paciente aislado por gripe) (1:1). Cada control se emparejó con el caso por las siguientes variables: fecha de ingreso, servicio de ingreso, planta de hospitalización, estancia hospitalaria y edad (± 5 años). Se excluyeron los

pacientes menores de 18 años y los ingresos de menos de 24h.

Para testar la fiabilidad interobservador en la detección de EARA, 3 médicos revisaron las historias clínicas de forma retrospectiva. Uno de ellos se consideró investigador de referencia (*gold standard*) y era un médico residente de medicina preventiva y salud pública de cuarto año (FJP), asesorado por un especialista con amplia experiencia en seguridad del paciente. Sus resultados se compararon con los de los otros 2 investigadores MIR (AIBS, MCMG).

Cada uno de los 3 MIR revisó un total de 100 historias clínicas seleccionadas de pacientes ingresados: 50 historias de pacientes hospitalizados por gripe y sometidos a precauciones de aislamiento y 50 historias de pacientes hospitalizados por otras causas y sin precauciones de aislamiento.

Las historias clínicas se seleccionaron a través del sistema ORACLE/ALCOR del Hospital y fueron solicitadas al Archivo a través del Servicio de Admisión y Documentación Clínica, para la revisión manual por parte de los 3 MIR.

Para la identificación de los eventos adversos se utilizó la Guía de Cribado validada del Proyecto IDEA (Identificación de Efectos Adversos), un cuestionario elaborado a partir de una investigación previa, de una lista de condiciones similar a la de los estudios de Nueva York¹⁰, Utah y Colorado¹⁷, bajo técnicas de consenso¹⁸. Las historias clínicas que cumplían al menos uno de los 19 criterios de la Guía de Cribado fueron examinadas con detalle con posterioridad, mediante el Formulario Modular de Revisión (MRF2), para la caracterización precisa del evento adverso.

Los resultados se introdujeron en una base de datos *ad hoc* para posteriormente realizar un análisis de concordancia que nos permitió evaluar la fiabilidad y la precisión en la detección de EARA de 2 de los 3 MIR tomando como referencia al tercero (*gold standard*).

Se calculó la incidencia de EARA detectados de forma global, y por separado en el grupo de pacientes aislados frente al de no aislados. Se comparó la diferencia de proporciones de EARA con el test estadístico de ji cuadrado, en el grupo de aislamiento por gripe frente al de no aislamiento. Esta diferencia se analizó para cada uno de los 3 revisores.

Se establecieron 3 escenarios para definir evento adverso, en función del grado de acuerdo entre los 3 evaluadores para identificar este:

- *Escenario de acuerdo máximo:* se consideró EARA si los 3 evaluadores coincidían en su detección.
- *Escenario de referencia:* EARA identificado solo por el revisor de referencia o *gold standard*.
- *Escenario de acuerdo mínimo:* se consideró EARA si cualquiera de los revisores lo identificó como tal.

El grado de concordancia entre los 3 médicos residentes fue medido usando el coeficiente Kappa¹⁹ (que considera acuerdo pobre por debajo de 0,2; débil entre 0,21 y 0,4; moderado entre 0,41 y 0,60; bueno entre 0,61 y 0,80, y muy bueno de 0,81 a 1) y también por el porcentaje simple de concordancia.

Se determinó el porcentaje simple de concordancia y el índice Kappa en cada uno de los siguientes aspectos: el acuerdo en el cribado, el grado de relación del evento adverso con la asistencia sanitaria y su evitabilidad.

Tabla 1 Incidencia de eventos adversos (EA), general y según exposición de la cohorte

Escenarios	Incidencia general	EA en aislados por gripe	EA en no aislados	p-valor
Opción de acuerdo mínimo ^a	26% [IC 95: 17,4-34,60%]	22%	30%	p = 0,362
Opción de incidencia estándar ^b	23% [IC 95: 14,75-31,25%]	20%	26%	p = 0,476
Opción de acuerdo máximo ^c	10% [IC 95: 4,12-15,88%]	2%	18%	p = 0,008

^a EA identificado por cualquiera de los revisores.

^b EA identificado por el revisor *gold standard*.

^c EA identificado por los 3 revisores.

Este estudio se realizó siguiendo las recomendaciones de la Organización Mundial de la Salud (OMS) y de la Ley de Cohesión del Sistema Nacional de Salud (SNS)²⁰. El director del estudio estableció las condiciones necesarias para garantizar el cumplimiento de la Ley Orgánica 15/1999, de Protección de Datos de Carácter Personal. La recogida inicial de datos fue nominal, pero se mantuvo la identificación individual exclusivamente hasta superar los controles de calidad de la base de datos.

El estudio fue sometido a la consideración del Comité de Ética e Investigación Clínica del hospital Universitario San Juan de Alicante, que dio su aprobación.

Resultados

El revisor de referencia identificó un 27% de alertas positivas en la guía de cribado, mientras que los otros 2 revisores identificaron respectivamente un 24 y un 28%.

Considerando los 3 escenarios para el cálculo de la incidencia, vemos que esta varió de un 10 a un 26% según lo estricto en el nivel de exigencia para considerar EARA (tabla 1). Además, fue siempre mayor en el grupo de casos expuestos al aislamiento por gripe que en el grupo control de pacientes no aislados, y esa diferencia alcanzó significación estadística cuando aplicamos el criterio de máxima exigencia en la definición de caso.

El porcentaje simple de concordancia de la guía de cribado entre el revisor estándar y el revisor 1 fue del 90%, coincidiendo en 8 cribados positivos y 37 cribados negativos en los aislados. Con respecto a los 50 no aislados que se cribaron, observamos un porcentaje simple de concordancia

entre ambos del 92%, coincidiendo en 13 cribados positivos y 33 negativos (tabla 2). El índice Kappa fue 0,70 para los aislados y 0,81 para los no aislados entre el revisor *gold standard* y el revisor 1.

El porcentaje simple de concordancia de la guía de cribado entre el revisor estándar y el revisor 2 fue del 78%, coincidiendo en 4 cribados positivos y 35 cribados negativos en los 50 aislados. Con respecto a los 50 no aislados que se cribaron, observamos un porcentaje simple de concordancia entre ambos del 76%, coincidiendo en 12 cribados positivos y 26 negativos (tabla 2). El índice Kappa para el revisor 2 fue 0,29 para los aislados y 0,48 para los no aislados.

Comparando los resultados obtenidos para la determinación de EARA del revisor *gold standard* y del revisor 1 mediante el uso del MRF2, observamos que, dentro de los 23 EARA determinados por el revisor de referencia o *gold standard*, el porcentaje simple de concordancia entre ambos fue del 87,5% para los aislados (7 coincidencias de 8). Con respecto a los no aislados, observamos un porcentaje simple de concordancia del 93,3% (14 coincidencias de 15) (tabla 3). El índice Kappa para el revisor 1 fue 0,60 para los aislados y no es calculable para los no aislados.

El porcentaje simple de concordancia entre el revisor estándar y el revisor 2 en la identificación de EARA usando el MRF2 fue del 50% para los aislados (2 coincidencias de 4). Con respecto a los no aislados, observamos un porcentaje simple de concordancia del 71,4% (10 coincidencias de 14) (tabla 3). El índice Kappa para el revisor 2 fue 0,20 para los aislados y no es calculable para los no aislados.

Comparando los resultados obtenidos para la determinación del grado de evitabilidad de los EARA del revisor estándar y del revisor 1 mediante el uso del MRF2,

Tabla 2 Resultados del cribado del revisor estándar con los revisores 1 y 2

	Revisor <i>gold standard</i>					
	Pacientes aislados			Pacientes no aislados		
	+	—	Total	+	—	Total
Revisor 1						
+	8	1	9	13	2	15
—	4	37	41	2	33	35
Total	12	38	50	15	35	50
Revisor 2						
+	4	3	7	12	9	21
—	8	35	43	3	26	29
Total	12	38	50	15	35	50

Tabla 3 Resultados del MRF2 para la identificación de evento adverso relacionado con la asistencia (EARA) del revisor estándar con los revisores 1 y 2

	Revisor <i>gold standard</i>					
	Pacientes aislados			Pacientes no aislados		
	Sí EARA	No EARA	Total	Sí EARA	No EARA	Total
Revisor 1						
Sí EARA	6	1	7	14	1	15
No EARA	0	1	1			
Total	6	2	8	14	1	15
Revisor 2						
Sí EARA	1	0	1	10	1	11
No EARA	2	1	3	3	0	3
Total	3	1	4	13	1	14

observamos que, dentro de los 23 EARA determinados por el revisor estándar, el porcentaje simple de concordancia en la valoración del grado de evitabilidad entre ambos revisores fue del 75% para los aislados (6 coincidencias de 8). Con respecto a los no aislados, observamos un porcentaje simple de concordancia del 86,6% (13 coincidencias de 15). El índice Kappa para el revisor 1 fue 0,38 para los aislados y 0,70 para los no aislados.

El porcentaje simple de concordancia entre el revisor estándar y el revisor 2 en la determinación de la evitabilidad del EARA usando el MRF2 fue del 66,6% para los aislados (2 coincidencias de 3). Con respecto a los no aislados, observamos un porcentaje simple de concordancia del 71,4% (10 coincidencias de 14) (tabla 4). El índice Kappa para el revisor 2 fue 0,46 para los no aislados y no se pudo calcular para los aislados.

Discusión

La definición de caso, junto con la fiabilidad de las mediciones, son elementos esenciales de los estudios epidemiológicos, de ahí sea preciso ser especialmente cuidadosos con el diseño metodológico para garantizar la validez de

los resultados y la eficiencia de las medidas consecuentes a ellos.

La precisión en la definición de caso y en el rigor en la revisión de la historia clínica acaba siendo definitiva en la estimación de la incidencia de eventos adversos¹, que en nuestro estudio hemos visto que va del 10 al 26%. Es decir, el efecto observador puede multiplicar por 2,6 el resultado, de ahí la importancia de realizar una buena formación de los revisores que hagan estudios de esta naturaleza, a la vez que la de publicar los resultados de la concordancia para que el lector pueda enmarcar convenientemente la precisión del estudio y su validez. Curiosamente, la incidencia es siempre menor en la cohorte expuesta al aislamiento que en la no expuesta, y esa diferencia alcanza significación estadística cuando aplicamos el criterio de máxima exigencia en la definición de caso. Resultado este contrario a lo señalado por un trabajo previo que evaluaba el impacto del aislamiento sobre la seguridad del paciente, y que encontró que los pacientes aislados tenían el doble de probabilidad de sufrir un evento adverso por cada 1.000 días de estancia que los pacientes no aislados, así como una probabilidad 8 veces mayor de experimentar fallos en el cuidado, como caídas, úlceras por presión o alteraciones hidroelectrolíticas^{9,21}. Es posible que nuestro resultado esté condicionado por diferentes circunstancias, como el

Tabla 4 Evitabilidad determinada por el revisor estándar y los revisores 1 y 2 con el formulario modular de revisión (MRF2)

	Revisor <i>gold standard</i>					
	Pacientes aislados			Pacientes no aislados		
	Evitable	Inevitable	Total	Evitable	Inevitable	Total
Revisor 1						
Evitable	5	0	5	9	0	9
Inevitable	2	1	3	2	4	6
Total	7	1	8	11	4	15
Revisor 2						
Evitable	2	0	2	6	1	7
Inevitable	1	0	1	4	4	8
Total	3	0	3	10	5	15

hecho de tratarse de pacientes con gripe grave, sometidos a especial vigilancia clínica, epidemiológica y social; no en vano había un interés desmedido de los medios de comunicación, durante el otoño-invierno pasados, por conocer los pormenores de la gripe, sus posibilidades de transmisión y su impacto en términos de morbilidad, de tal modo que el efecto de la presión mediática pudiera haber condicionado un cierto «efecto Hawthorne». Otra causa podría ser la presencia de posibles factores de confusión tales como una mayor comorbilidad en el grupo control de pacientes no aislados que en el grupo de pacientes aislados por gripe.

Por otra parte, este estudio concluye que los 2 revisores que son comparados con el estándar presentan diferentes grados de concordancia. Cuando medimos el porcentaje simple de concordancia, vemos que ambos revisores presentan en líneas generales elevados porcentajes de acuerdo con el estándar, sobre todo el revisor 1, con más del 90% de coincidencia en el cribado del evento adverso, siendo este porcentaje de coincidencia algo menor con el revisor 2. Sin embargo, cuando medimos solo el porcentaje simple, no estamos eliminando el efecto del azar. Para ello disponemos de otra herramienta que es el índice Kappa. Cuando medimos el índice Kappa vemos que los niveles de concordancia siguen manteniéndose bastante elevados entre el revisor estándar y el revisor 1 (grados de acuerdo de «moderado» a «muy bueno») tanto en los aislados como en los no aislados, siendo este índice menor en el caso de la concordancia con el revisor 2. La diferencia de concordancia entre los revisores 1 y 2 apoya la necesidad de formación y entrenamiento de los revisores para minimizar la fiabilidad interobservador. La fiabilidad entre 2 equipos de observadores al utilizar herramientas como los «*trigger tool*», en un trabajo reciente, mostró que solo el 31% de los eventos adversos fueron identificados por los 2 equipos. El acuerdo entre observadores se consideró moderado ($\kappa = 0,45$, produciendo diferentes conclusiones. El análisis encontró un elevado error aleatorio y un pequeño error sistemático²². La fiabilidad del *global trigger tool* para la presencia, número y gravedad de los eventos adversos osciló entre $\kappa = 0,40$ y $\kappa = 0,60$. Cuando se comparó con un grupo de expertos revisores, la sensibilidad fue moderada (49%) y la especificidad alta (94%). La fiabilidad intra e interobservador se consideró apropiada para llevar a cabo estudios de comparación de tasas entre centros²³.

Observamos que el grado de concordancia es mayor al inicio del estudio, en la guía de cribado, y va disminuyendo conforme avanza el análisis del EARA con el cuestionario modular MRF2. Esto tiene lógica, teniendo en cuenta que conforme avanza el estudio del EARA se va concretando cada vez más en las variables del análisis, a la par que la muestra va siendo cada vez más reducida. Es decir, al principio del estudio, en el cribado inicial de detección de EARA partimos de una muestra de 100 historias clínicas (50 aislados y 50 no aislados), y es aquí donde observamos mayores niveles de concordancia; sin embargo, cuando medimos la concordancia para el EARA ya partimos de una muestra de 23 historias consideradas como «EARA» por el estándar, y finalmente para valorar la «evitabilidad» del EARA únicamente nos quedamos con los casos del estándar que coinciden en su valoración de EARA evitable («sí» o «no») con los otros revisores, por lo que la muestra es ya muy pequeña. Es por ello que hay que ser especialmente cuidadosos en la capacitación de los revisores para disminuir los acuerdos debidos al

azar, pero sobre todo para mitigar el efecto de infraestimación de los eventos adversos que limitaría la capacidad de elaborar estrategias para mejorar la seguridad de los pacientes. A pesar de esto, la concordancia para la determinación de EARA se mantiene moderadamente alta (sobre todo con el revisor 1).

El descenso progresivo de la muestra puede justificar en parte el declive observado en el grado de acuerdo de los revisores conforme avanza el estudio. Es por esto que para valorar la «evitabilidad» del evento adverso puede ser más adecuado realizar una puesta en común de los casos y consenso de los revisores, en vez de realizar la valoración individual más subjetiva, ya que la muestra a este nivel es muy reducida y hace imposible el cálculo del índice Kappa. Además, conscientes de esta realidad, para minimizar este efecto usaremos en el formulario MRF2 una escala tipo Likert, de 6 tramos.

Por otro lado, para aumentar el grado de acuerdo con el revisor 2 será necesario realizar una puesta en común entre los revisores para unificar criterios en el uso del MRF2 y mejorar así el nivel de concordancia para un futuro análisis.

En estudios realizados en diferentes países, las diferencias metodológicas parecen explicar, en parte, la variación en los resultados^{17,24}. La investigación sobre la seguridad del paciente a través de la revisión de historias debe ser fiable y precisa, sin perder eficiencia. Los pocos estudios publicados sobre la fiabilidad de la detección de eventos adversos indican un nivel de concordancia entre médicos y enfermeros moderado²⁵⁻²⁷. Del mismo modo, los estudios de evaluación de la exactitud de la detección de eventos adversos son escasos y por lo general muestran mayor sensibilidad que especificidad^{28,29}, excepto en un reciente estudio, en el que la especificidad fue mayor³⁰.

Para concluir, podemos afirmar que sin una buena metodología, pero sobre todo sin una buena regla de medir, no se pueden garantizar los resultados alcanzados y, consecuentemente, las decisiones tomadas a partir de ellos, pudiendo ser desde acertadas a muy disparatadas. Es por ello que los investigadores han de ser garantes del método utilizado, y este ha de acercarse lo más posible al óptimo alcanzable.

Financiación

Este estudio forma parte de un proyecto multicéntrico, sobre la seguridad del paciente hospitalizado con precauciones de aislamiento, financiado por el Fondo de Investigación Sanitaria, referencias PI13/01921 y PI013/02161.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Bibliografía

1. Aranaz JM, Aibar C, Gea MT, León MT. Los efectos adversos en la asistencia hospitalaria. Una revisión crítica. *Med Clin (Barc)*. 2004;123:21-5.
2. Aranaz-Andrés JM, Aibar-Remón C, Vítaller-Burillo J, Ruiz-López P, Limón-Ramírez R, Terol-García E, the ENEAS work group. Incidence of adverse events (AEs) related to health care in Spain.

- Results of the Spanish National Study of Adverse Events (ENEAS). *J Epidemiol Community Health*. 2008;62:1022–9.
3. Aranaz JM. Sucesos adversos relacionados con el uso del medicamento: ¿qué podemos aprender? *Med Clin (Barc)*. 2006;126:97–8.
 4. Aranaz JM, Aibar C, Vitaller J, Ruiz P. Estudio Nacional sobre los Efectos Adversos ligados a la Hospitalización. ENEAS 2005. Madrid: Ministerio de Sanidad y Consumo; 2006.
 5. Vaqué J, Rosselló J, Arribas JL. Prevalence of nosocomial infections in Spain: EPINE study 1990–1997. EPINE Working Group. *J Hosp Infect*. 1999;43 Suppl:S105–11.
 6. Zarb P, Coignard B, Griskeviciene J, Muller A, Vankerckhoven V, Weist K, et al. The European Centre for Disease Prevention and Control (ECDC) pilot point prevalence survey of healthcare-associated infections and antimicrobial use. *Euro Surveill*. 2012;17:20316.
 7. Pujol M, Limón E. Epidemiología general de las infecciones nosocomiales. Sistemas y programas de vigilancia. *Enferm Infecc Microbiol Clin*. 2013;31:108–13.
 8. Aranaz-Andrés JM, Aibar-Remón C, Vitaller-Burillo J, Gea-Velázquez de Castro MT, Requena-Puche J, Terol-García E, y grupo de trabajo ENEAS. Impact and preventability of adverse events in Spanish public hospitals. Results of the Spanish National Study of adverse events (ENEAS). *Int J Qual Health Care*. 2009;21:408–14.
 9. Abad C, Fearday A, Safdar N. Adverse effects of isolation in hospitalised patients: A systematic review. *J Hosp Infect*. 2010;76:97–102.
 10. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med*. 1991;324:370–6.
 11. Aranaz-Andrés JM, Limón R, Mira JJ, Aiba C, Gea MT, Agra Y, and ENEAS Wordking Group. What makes hospitalized patients more vulnerable and increases their risk of experiencing an adverse event? *Int J Qual Health Care*. 2011;23:705–12.
 12. Aranaz-Andrés JM, Aibar-Remón C, Limón-Ramírez R, Amarilla A, Restrepo FR, Urroz O, et al. Diseño del estudio IBEAS: prevalencia de efectos adversos en hospitales de Latinoamérica. *Rev Calid Asist*. 2011;26:194–200.
 13. Aranaz-Andrés JM, Aibar-Remón C, Limón-Ramírez R, Amarilla AC, Restrepo-Parra FR, Urroz-Torres O, et al. Prevalence of adverse events in the hospitals of five Latin American countries: Results of the Iberoamerican study of adverse events (IBEAS). *BMJ Qual Saf*. 2011;20:1043–51.
 14. Gea-Velázquez de Castro MT, Aranaz-Andrés JM. Utilidad de la guía de cribado del estudio ENEAS en un hospital de atención a enfermos pluripatológicos. *Med Clin (Barc)*. 2010;135:24–30.
 15. Requena J, Aranaz JM, Mira JJ, Gea MT, Miralles JJ, Limón R, et al. ¿Cómo identificar los sucesos adversos en urgencias? Una guía consensuada para el cribado. *Rev Calid Asist*. 2009;24:272–9.
 16. Aibar L, Rabanaque MJ, Mozas J, Puertas A, Aranaz JM, Aibar C. Improving patient safety by detecting obstetric care-related adverse events. Application of a new screening guide. *Arch Gynecol Obstet*. 2014;289:945–52.
 17. Thomas EJ, Studdert DM, Burstin HR, Orav EJ, Zeena T, Williams EJ, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care*. 2000;38:261–71.
 18. Aranaz JM, Limón R, Requena J, Gea MT, Núñez V, Bermúdez MI, et al., Grupo de trabajo del proyecto IDEA. Incidencia e impacto de los efectos adversos en dos hospitales. *Rev Calid Asist*. 2005;20:53–60.
 19. Guggenmoos-Holzmahn I. The meaning of Kappa: Probabilistic concepts of reliability and validity revisited. *J Clin Epidemiol*. 1996;49:775–82.
 20. Ley 16/2003, de 28 de mayo, de cohesión y calidad del Sistema Nacional de Salud. BOE núm. 128 de 29 mayo de 2003. [consultado 23 Jul 2014]. Disponible en: <http://www.boe.es/boe/dias/2003/05/29/pdfs/A20567-20588.pdf>
 21. Morgan DJ, Diekema DJ, Sepkowitz K, Perencevich EN. Adverse outcomes associated with contact precautions: A review of the literature. *Am J Infect Control*. 2009;37:85–93.
 22. Mattsson TO, Knudsen JL, Lauritsen J, Brixen K, Herrstedt J. Assessment of the global trigger tool to measure, monitor and evaluate patient safety in cancer patients: Reliability concerns are raised. *BMJ Qual Saf*. 2013;22:571–9.
 23. Sharek PJ, Parry G, Goldmann D, Bones K, Hackbarth A, Resar R, et al. Performance characteristics of a methodology to quantify adverse events over time in hospitalized patients. *Health Serv Res*. 2011;46:654–78.
 24. Michel P, Quenon JL, de Sarasqueta AM, Scemama O. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. *BMJ*. 2004;328:199–202.
 25. Silver MP, Hougland P, Elder S, Haug J, Pritchett T, Donnelly S, et al. Statewide identification of adverse events using retrospective nurse review: Methods and outcomes. *J Nurs Meas*. 2007;15:220–32.
 26. Weingart SN, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Discrepancies between explicit and implicit review: Physician and nurse assessments of complications and quality. *Health Serv Res*. 2002;37:483–98.
 27. Naessens JM, O'Byrne TJ, Johnson MG, Vansuch MB, McGlone CM, Huddleston JM. Measuring hospital adverse events: Assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Health Care*. 2010;22:266–74.
 28. Brennan TA, Localio RJ, Laird NL. Reliability and validity of judgements concerning adverse events suffered by hospitalized patients. *Med Care*. 1989;27:1148–58.
 29. Camacho LAB, Rubin HR. Assessment of the validity and reliability of three systems of medical record screening for quality of care assessment. *Med Care*. 1998;36:748–51.
 30. Braz Pavao AL, Camacho LAB, Martins M, Mendes W, Travassos C. Reliability and accuracy of the screening for adverse events in Brazilian hospitals. *Inter J Qual Health Care*. 2012;1–6.