



Original

Evaluación de la variabilidad interobservador en la clasificación sistemática de la morbilidad operatoria en resección pulmonar

Gonzalo Varela* y Nuria M. Novoa

Servicio de Cirugía Torácica, Hospital Universitario de Salamanca, Salamanca, España

INFORMACIÓN DEL ARTÍCULO

Historia del artículo:

Recibido el 3 de abril de 2011

Aceptado el 14 de mayo de 2011

On-line el 12 de julio de 2011

Palabras clave:

Resección pulmonar

Morbilidad operatoria

Bases de datos

Concordancia interobservador

R E S U M E N

Objetivo: Cuantificar la variabilidad interobservador al establecer una clasificación sistemática de la morbilidad operatoria de la resección pulmonar.

Método: Se ha clasificado la morbilidad de forma retrospectiva en una serie de 499 registros prospectivos de pacientes sometidos a resección pulmonar mayor (458 lobectomías y 51 neumonectomías). Se utilizó la clasificación sistemática propuesta por Seely et al en 2010. Cada uno de los autores clasificó de forma independiente las complicaciones y se calculó el estadístico kappa ponderado.

Resultados y comentarios: El índice kappa fue de 0,79. Aunque el valor es alto, introduce un sesgo sistemático en la clasificación de la morbilidad de los pacientes que indica la necesidad de valorar muy cuidadosamente los datos introducidos en los registros multiinstitucionales para poder obtener conclusiones válidas.

© 2011 SEPAR. Publicado por Elsevier España, S.L. Todos los derechos reservados.

Evaluation of the Interobserver Variability in the Systematic Classification of Operative Morbidity in Lung Resection

A B S T R A C T

Objective: To quantify the interobserver variability in establishing a systematic classification for the operative morbidity of lung resection.

Methods: Morbidity was classified retrospectively in a series of 499 prospective registries of patients who underwent major lung resection (458 lobectomies and 51 pneumonectomies). The systematic classification proposed by Seely et al. in 2010 was used. Each one of the authors independently classified the complications and the weighted kappa statistic was calculated.

Results and comments: The kappa index was 0.79. Although the value is high, it introduces a systematic bias in the classification of patient morbidity that indicates the need to very carefully evaluate the data entered into the multi-institutional registers in order to be able to obtain valid conclusions.

© 2011 SEPAR. Published by Elsevier España, S.L. All rights reserved.

Keywords:

Lung resection

Operative morbidity

Data bases

Interobserver agreement

Introducción

Para la evaluación de la calidad de la cirugía es imprescindible conocer los efectos adversos de la misma, fundamentalmente la mortalidad y la morbilidad operatorias. Un estudio previo realizado en España sobre pacientes sometidos a resección pulmonar y en el que se utilizaron bases de datos administrativas¹ demostró que existen diferencias notables entre centros al comparar la mortalidad hospitalaria bruta y ajustada por riesgo. La mortalidad es una variable muy robusta en las bases de datos administrativas, ya que

el fallecimiento de los pacientes siempre consta si se produce. Sin embargo, cuando se analizó la morbilidad posquirúrgica los resultados fueron de algún modo desconcertantes, ya que se encontró que en los centros con mayor mortalidad no se registraban las mayores cifras de morbilidad, como era de esperar. Este hallazgo, que podría ser debido a variaciones en las prácticas clínicas locales y en la calidad de los registros clínicos, hace pensar que el análisis de la morbilidad de grandes series de pacientes incorporados en bases de datos multicéntricas podría estar sometido a un sesgo sistemático que invalide las conclusiones.

Recientemente, Seely et al² han propuesto la aplicación de una clasificación sistemática de la morbilidad operatoria en cinco grupos estandarizados (tabla 1) y concluyen que el uso de este tipo de clasificación facilita la comparación objetiva entre procedimientos

* Autor para correspondencia.

Correo electrónico: gvs@usal.es (G. Varela).

Tabla 1
Clasificación sistemática de las complicaciones en intervenciones torácicas

Grado	Definición
<i>Complicación menor</i>	
I	No es necesario tratamiento farmacológico ni de otro tipo
II	Se requiere tratamiento farmacológico o una intervención menor
<i>Complicación mayor</i>	
III	Se requiere una intervención quirúrgica, radiológica o endoscópica o múltiples tratamientos
IIIa	La intervención no requiere anestesia general
IIIb	La intervención requiere anestesia general
IV	Se necesita tratamiento en la unidad de cuidados intensivos y soporte vital
IVa	Disfunción de un solo órgano
IVb	Disfunción multiorgánica
V	La complicación lleva al paciente a la muerte

Modificado de Seely et al².

quirúrgicos y series de pacientes, y entre cirujanos y equipos quirúrgicos diferentes. Sin embargo, aunque los autores hacen referencia al tema en su publicación, se desconoce la variabilidad interobservador a la hora de clasificar las complicaciones posquirúrgicas, lo cual sería fundamental antes de adoptar este tipo de clasificaciones para el análisis de grandes bases de datos clínicas.

En este estudio pretendemos cuantificar el acuerdo entre dos observadores independientes al clasificar las complicaciones postoperatorias en resección pulmonar.

Método

Análisis retrospectivo de los registros de todos los paciente a los que se ha realizado una resección pulmonar anatómica (lobectomía o neumonectomía) en nuestro centro entre enero de 2007 y diciembre de 2010. Los registros se obtuvieron de una base de datos prospectiva que contiene los datos de todos los pacientes operados desde enero de 1994 hasta el momento actual. La calidad de los datos se asegura mediante dos controles de calidad realizados por un *data manager*, el primero en el momento de emitir el informe de alta y el segundo antes de remitir la documentación clínica al archivo general del hospital, una vez incorporados los estudios histológicos u otros que estuvieran pendientes. Las complicaciones postoperatorias se encuentran codificadas, y existe un acuerdo escrito en la unidad sobre la definición de cada tipo de complicación posquirúrgica.

Para este trabajo se ha realizado una explotación de la información contenida en la base de datos que incluye las siguientes variables: número de historia clínica, procedimiento quirúrgico principal, mortalidad a 30 días del procedimiento, complicación postoperatoria (sí/no) y tipo de complicación (hasta 4 códigos de complicaciones por paciente), y además, la necesidad de reoperación y el fallecimiento del enfermo en los 30 días siguientes a la operación. El primer firmante realizó una revisión manual de las incongruencias o datos ausentes apreciados en el listado inicial y estableció el listado definitivo, que no contenía datos ausentes ni incongruencias.

Una vez realizada la explotación de los datos, se preparó un listado de casos con sus complicaciones y los dos autores, de forma independiente, clasificaron las complicaciones en 5 niveles, de acuerdo con las definiciones presentadas en el artículo de Seely et al². Finalmente, las dos gradaciones se unieron en un único archivo y se calculó el estadístico kappa ponderado de Cohen. Para la cuantificación de la importancia del desacuerdo se aceptaron los valores por defecto que proporciona el programa estadístico utilizado (Stata 10.1).

Tabla 2
Complicaciones registradas en la serie de casos

Tipo de complicación	Complicación registrada (n.º de casos)					
	A	B	C	D	E	F
Accidente vascular cerebral	1					
Arritmia cardíaca	18	4	1			
Atelectasia	5	1	3			
Edema pulmonar	1					
Embolia pulmonar	1					
Empiema pleural	4	3	1			
Fístula bronquial	2	1	2			
Fuga aérea >5 días	45	3	1			
Hematoma de herida	1					
Hemorragia	2	1				
Hemotórax	8	1				
Hernia diafragmática		1				
Íleo paralítico	2					
Infección de herida	6	4		2		
Infección urinaria	2					
Insuficiencia cardíaca	7	1	2			
Insuficiencia renal	2		1	1		
Lesión del nervio periférico	3	1				
Necesidad de oxigenoterapia al alta	5	4	1			
Neumonía nosocomial	9	5	1			
Neumotórax	7	2				
Paro cardíaco	1					
Pericarditis	1					
Reacción medicamentosa	1					
Retención urinaria	1					
Sepsis	1					
Ventilación mecánica postoperatoria	1					
Reintervención mayor					18	
Muerte						5

Tabla 3
Clasificaciones asignadas por los dos observadores

Observador 2	Observador 1							Total
	I	II	IIIa	IIIb	IVa	IVb	V	
I	40	5	0	0	0	0	0	45
II	1	43	4	0	0	0	0	48
IIIa	0	8	10	0	0	1	0	19
IIIb	0	1	1	12	0	0	0	14
IVa	0	1	1	1	2	0	0	5
IVb	0	2	0	2	0	0	0	4
V	0	0	0	0	0	0	5	5
Total	41	60	16	15	2	1	5	140

Kappa 0,79 (error estándar, 0,05; p < 0,0001).

Resultados

Se ha obtenido una serie de 499 pacientes (41 neumonectomías y 458 lobectomías). La mortalidad a 30 días es del 1% (5 pacientes, 2 post lobectomía y 3 post neumonectomía). Se han registrado una o más complicaciones en 140 casos (28%). Las complicaciones incluidas en el listado que se utilizó por los evaluadores figuran en la tabla 2.

En la tabla 3 se puede observar la clasificación por niveles de complicación realizada por los dos observadores. El valor del estadístico kappa ponderado es de 0,79 (error estándar, 0,05; p < 0,0001).

Discusión

El análisis de bases de grandes datos clínicas es en la actualidad un medio necesario para la mejora de la calidad asistencial, ya que permite obtener en poco tiempo datos robustos basados en grandes poblaciones de pacientes que sirven de estándar para comparar los datos de la propia institución e introducir las medidas correctoras necesarias en caso de observar desviaciones importantes^{3,4}. Si

bien es cierto que la calidad del trabajo quirúrgico no debe basarse solamente en los resultados (mortalidad y morbilidad), hasta ahora existen muy pocas iniciativas publicadas sobre el uso de otras variables, tales como la adecuación del diseño de los procesos clínicos a las mejores evidencias científicas disponibles. En cirugía torácica se ha propuesto la construcción de un índice que reúne los resultados y el diseño de procesos para la evaluación de la calidad⁵, pero su aplicación práctica está todavía en desarrollo⁶.

La clasificación de la severidad de las complicaciones propuesta por Seely et al² parece a priori un método adecuado para estandarizar la evaluación de los resultados de morbilidad operatoria, de tal forma que la comparación entre centros tenga mayor validez. Además, introduce un concepto innovador, como es la consideración de los recursos utilizados para tratar las complicaciones.

El cálculo de índice kappa de acuerdo entre observadores es un método muy utilizado en epidemiología clínica cuando se desea cuantificar la variabilidad entre varios observadores de sucesos clínicos y es un parámetro mucho más representativo que la simple tasa de acuerdo, ya que ésta está muy influenciada por el azar⁷. Cuando el orden de los valores que se obtienen en la observación tiene importancia, como es el caso de esta investigación, es recomendable el uso del índice kappa ponderado, ya que no es igualmente valorable que exista discrepancia clasificando la complicación como grado 1 o grado 2, o que la discrepancia sea entre los grados 1 y 5. Aunque el uso del estadístico kappa se ha criticado por su dependencia de la prevalencia del suceso estudiado y por la subjetividad de la ponderación de las discrepancias⁸, se considera que un índice kappa ponderado por encima de 0,8 indica un excelente grado de concordancia entre dos observadores⁷. En nuestro caso, el índice kappa obtenido de 0,79 merece algunos comentarios.

Aunque sin duda este valor demuestra que existe un elevado acuerdo entre dos observadores al clasificar la severidad de las complicaciones registradas, hay que tener en cuenta que los dos observadores participamos diariamente en la discusión de los mismos casos, y que en el centro existe un acuerdo previo acerca de la definición de las complicaciones. En este contexto, el acuerdo esperado entre observadores probablemente debería ser del 100%. Es fácil suponer que, cuando se trabaje en una base de datos multiinstitucional, el grado de acuerdo a la hora de definir y clasificar

las complicaciones debe ser mucho menor o, lo que es lo mismo, el esfuerzo requerido para la validez de los datos al analizar la morbilidad es mucho mayor.

El estudio que hemos presentado demuestra que, incluso entre cirujanos del mismo centro, existe un cierto grado de discrepancia en la clasificación de las complicaciones posquirúrgicas. Por lo tanto, antes de adoptar un sistema de estandarización de la clasificación de la morbilidad a nivel multiinstitucional es preciso estudiar la variabilidad e introducir las medidas correctivas necesarias para la construcción de índices de riesgo en intervenciones torácicas. Estas medidas incluirían la definición precisa e inequívoca de cada tipo de complicación y una clasificación genérica de los tratamientos requeridos para resolverlas. Solo garantizando la calidad de los datos introducidos en las bases de datos clínicas se pueden obtener conclusiones válidas⁹.

Bibliografía

1. Varela G, Molins L, Astudillo J, Borro JM, Canalís E, Freixinet J, et al. Experiencia piloto de benchmarking en Cirugía Torácica: comparación de la casuística e indicadores de calidad en resección pulmonar. Arch Bronconeumol. 2006;42:267-72.
2. Seely AJE, Ivanovich J, Threader J, Al-Hussaini A, Al-Shehab D, Ramsay T, et al. Systematic classification of morbidity and mortality alter thoracic surgery. Ann Thorac Surg. 2010;90:936-42.
3. Brunelli A, Varela G, Berrisford R, Rocco G. Audit, quality control and performance in Thoracic Surgery: an European perspective. Thorac Surg Clin. 2007;17:387-93.
4. Brunelli A, Varela G, Van Schil P, Salati M, Novoa N, Hendriks JM, et al. on behalf of the ESTS Audit and Clinical Excellence Committee. Multicentric analysis of performance after major lung resections by using the European Society Objective Score (ESOS). Eur J Cardiothorac Surg. 2008;33:282-8.
5. Brunelli A, Berrisford RG, Rocco G, Varela G, on behalf of the European Society of Thoracic Surgeons Database Committee. The European Thoracic Database project: composite performance score to measure quality of care after major lung resection. Eur J Cardiothorac Surg. 2009;35:769-74.
6. Brunelli A, Rocco G, Van Raemdonck D, Varela G, Dahan M. Lessons learned from the European thoracic surgery database: The composite performance score. Eur J Surg Oncol. 2010;36 Suppl 1:S93-99.
7. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation and sample size requirements. Phys Ther. 2005;85:267-8.
8. Thompson WD, Walters SD. A reappraisal of the Kappa coefficient. J Clin Epidemiol. 1998;41:949-58.
9. Pass HI, Medical registries. Continued attempts for robust quality of data. J Thorac Oncol. 2010;5:S198-9.