

# Evaluación de las solicitudes de mantenimiento correctivo usando técnicas de agrupamiento y reglas de asociación

Antonio M. Cruz<sup>1,ψ</sup>, Cameron Barr, Norberto Castilla Casado<sup>3</sup>

<sup>1</sup>*Departamento de Ciencias Básicas, Escuela de Medicina, Universidad del Rosario, Colombia*

<sup>3</sup>*Centro de Bioingeniería (CEBIO), Instituto Superior Politécnico “José Antonio Echeverría” (ISPJAE), Ciudad Habana, Cuba*

Recibido 3 de enero de 2008. Aceptado 7 de abril de 2008.

---

**Resumen**— En este trabajo se exponen los resultados obtenidos de la aplicación de técnicas de descubrimiento de asociaciones y de agrupamiento para resolver el problema de la baja eficiencia presentado en un servicio de esterilización de un hospital en estudio. El objetivo fue detectar y discriminar las causas fundamentales que contribuyeron al surgimiento del problema presentado para luego solucionarlo. Para realizar esta investigación se recabó la información contenida en las solicitudes de servicio de mantenimiento correctivo y las órdenes de trabajos durante el período 2002-2004. Primeramente se segmentó la información contenida en el indicador en estudio: razón de las solicitudes de servicio de mantenimiento correctivo vs. cantidad de equipos por tipos de equipos, por servicios, por fabricante (OEM, del inglés Original Equipment Manufacturer) y por modelos. Luego con las técnicas de descubrimiento de asociaciones aplicadas se encontraron las causas fundamentales por las cuales se solicitaban los reportes de servicios. Éstas fueron: falta de entrenamiento en usuarios, fallos intrínsecos en los dispositivos médicos y malas políticas en el establecimiento de la frecuencia del mantenimiento programado. Las técnicas de agrupamientos pudieron discriminar las causas fundamentales por las cuales los dispositivos médicos del servicio de esterilización fallaban. Éstas fueron debido a fallos en el sistema de suministro de vapor y agua que alimenta las unidades de esterilización (en un 75% de los casos). Se tomaron medidas correctoras durante el período 2005-2006, que contribuyeron a que el indicador bajo estudio disminuyera de 6,4 a 0,4 unidades.

**Palabras clave**— Descubrimiento de asociaciones, Gestión tecnológica, Ingeniería clínica, Minería de datos, Técnicas de agrupamientos.

**Abstract**— In this research association discovering and clustering techniques for the resolution of the low efficiency problem in the sterilization service in a hospital under study were used. The aim was to find and to discriminate the main causes of the problem under study and then to apply corrective solutions. To conduct this research the information contained in corrective maintenance work orders and service requests in the period under study (2002-2004) was collected. First a segmentation of the information was carried out using the indicator: corrective service request versus number of medical devices. The levels of the information segmentation were: equipment types, services or cost centre, original equipment manufacturer and models. Then the association discovery technique was used. It revealed that the main causes of low efficiency in sterilization service were: users' training (errors in operation procedures), intrinsic failures in medical devices, and bad scheduled maintenance policies. Clustering technique uncovered the main causes of failures: malfunctioning of the power supply system (steam and water, in 75% of all cases). With the evidence obtained corrective actions were taken. The service requests dropped dramatically from 6.4 to 0.4 during the period 2005-2006.

**Keyword**— Association discovery, Clustering techniques, Clinical engineering, Data mining, Technology management.

---

## I. INTRODUCCIÓN

Los departamentos de Ingeniería Clínica (DIC) analizan y evalúan la información relativa a los datos históricos de mantenimiento, tanto preventivo como correctivo, los protocolos de seguridad, los eventos adversos, los fallos de equipos, entre otros aspectos [1]. Esto permite encontrar patrones de comportamiento, obteniendo así un “nuevo conocimiento” que puede contribuir a un mejor proceso de toma de decisiones. Actualmente las técnicas de agrupamientos (*clustering*) y descubrimiento de asociaciones (*association discovery*) se emplean frecuentemente para cumplir estos objetivos en muchos campos del saber. Estos métodos se encuentran dentro de una rama de la inteligencia artificial llamada minería de datos [2]. Los algoritmos computacionales que se encargan de hacer los descubrimientos de asociaciones lo que persiguen es encontrar reglas de asociación (por ejemplo en la forma Si A y B implica C entonces D,) entre los campos de uno o varios registros de una base de datos [2-3]. Esta técnica ha sido usada en la detección de fraudes bancarios y de telecomunicaciones [2]. En el campo del mantenimiento industrial se muestra como Boeing, empresa dedicada a la fabricación de aviones para vuelos comerciales que ha ahorrado más de US \$300.000 por avión/año gracias al uso de técnicas de descubrimiento de asociaciones en fallos [4]. Adicionalmente, la empresa consultora Exclusive Ore Inc. reportó que en un conjunto de 114 locomotoras ahorró más de US \$53.580 por locomotora/año al aplicar reglas de descubrimiento de asociaciones en reportes de servicios y fallos en dicho conjunto [5].

Ahora bien, las técnicas de agrupamiento de datos lo que buscan es “agrupar” o “conglomerar” grupos de datos con características similares entre ellos [2-3].

Realizando una exhaustiva revisión del estado del arte en lo que se refiere a la aplicación de estas técnicas en el campo de la Ingeniería Clínica, se puede decir que no se observan trabajos publicados que apliquen las técnicas de agrupamiento y descubrimiento de asociaciones [1, 6]. Por ende, el comportamiento de la solicitud de los reportes de servicios por métodos de minería de datos tampoco ha sido reportado.

Estudiar el comportamiento del patrón de los reportes de servicios de mantenimiento correctivo tiene particular importancia, por cuanto se ha demostrado por algunos estudios que alrededor del 68% de las solicitudes de los reportes de servicios de mantenimiento correctivo son falsas [7]. La solicitud de reportes de servicios que no son fallos reales hace consumir mucho tiempo al personal del DIC que se dedica a las labores de mantenimiento. Además, denota que los usuarios no están lo suficientemente entrenados para el uso adecuado de la tecnología.

Las técnicas de descubrimiento de asociaciones que se aplican en este trabajo se realizaron empleando el algoritmo llamado *Apriori* [2, 3, 8]. Este es uno de los algoritmos más populares usados por su simplicidad y robustez [2]. De manera somera se puede decir que los algoritmos de descubrimiento de asociaciones encuentran sus reglas ordenando los datos y contando las apariciones de las variables que se están analizando, maximizando y calculando los parámetros confianza, soporte y *lift* (del inglés *confidence* y *support*, ver más adelante la definición de cada uno de estos términos).

Por otra parte, el algoritmo utilizado para realizar las técnicas de agrupamiento fue el llamado EM, por su nombre en inglés, *Expectation-Maximization Algorithm*, que asigna instancias a clases de datos de forma probabilística [3]. La razón fundamental por la cual fue seleccionado este algoritmo es que permite formar agrupamientos con datos numéricos y no numéricos [3].

Estos dos algoritmos ya se encuentran implementados en productos informáticos (por ejemplo, el programa WEKA) [9]. Por esta razón los autores de este artículo desean aclarar que no es objetivo de este trabajo que los lectores conozcan a fondo el modo en que fueron desarrollados desde el punto de vista de programación, sino que conozcan su uso o aplicación. Así las cosas, esta investigación propone tres objetivos fundamentales:

1. Ofrecer un método a la comunidad de ingenieros clínicos que permita discriminar cuáles equipos médicos en una institución de salud aportan la mayor contribución al número total de reportes de servicios de mantenimiento correctivo.
2. Demostrar que el uso de métodos “inteligentes”, como es el de descubrimiento de asociaciones, es efectivo para obtener patrones que permiten dilucidar las principales causas que provocan la solicitud de los reportes de servicios de mantenimiento correctivo.
3. Usar técnicas de agrupamiento para comprobar hipótesis o preguntas generadas en las reglas que resulten de la aplicación de las técnicas de descubrimiento de asociaciones, con lo cual se pueden tomar acciones correctivas basadas en la evidencia y solucionar los problemas presentados.

## II. MATERIALES Y MÉTODOS

El método empleado para desarrollar esta investigación fue el siguiente:

1. Primero se realizó una caracterización de la muestra de datos bajo en estudio.

2. Luego se efectuó una segmentación [2] de la información contenida en la muestra de datos con el objetivo de identificar los agrupamientos de equipos que poseían las mismas características. La segmentación se realizó por tipos de equipos (Tabla 1). Se comparó la cantidad de equipos médicos por tipo y la cantidad de reportes de servicios de mantenimiento correctivo por cada tipo en cuestión.
3. Seguidamente se identificaron los patrones en el indicador: cantidad de equipos médicos por tipo contra la cantidad de reportes de servicios de mantenimiento correctivo por cada tipo.
4. Después se aplicaron las técnicas de descubrimiento de asociaciones con el objetivo de encontrar las reglas que permitieron establecer las principales causas por la cuales los reportes de servicios de mantenimiento correctivo son llevados a cabo o solicitados.
5. Finalmente, se aplicaron técnicas de agrupamiento para identificar los grupos que más contribuyen a los patrones de reportes de servicios de mantenimiento correctivo, para poder tomar acciones correctivas sobre estos.

#### 1. Caracterización de la muestra de datos en estudio

La muestra de datos utilizada para realizar este estudio proviene de un hospital de tercer nivel con un total de 100 camas y 749 equipos médicos en el

inventario localizados en 25 centros de costos. Estos equipos médicos pertenecen a 400 modelos de 180 fabricantes (OEM) y proveedores diferentes. La Tabla 1 muestra un resumen de los indicadores que caracterizan la muestra de datos; nótese que:

- El inventario puede ser catalogado de “joven”, pues como promedio no ha alcanzado su vida útil. La columna 8 de la Tabla 1 muestra cómo el promedio del indicador razón tiempo de explotación contra vida útil es menor que la unidad ( $AVU=ET/UL$ ) [12], excepto para los equipos de electromecánica que tienen un valor de 1,20 unidades
- Durante el período que se está estudiando (2002-2004) hubo un total de 719 reportes de servicio de mantenimiento correctivo en un total de 416 dispositivos pertenecientes a todas las categorías.

#### 2. Segmentación de la base de datos

De las Fig. 1, 2, 3 y 4 se puede observar que:

- La categoría “B” de Tipos de Equipos es la que tiene el mayor valor del indicador razón solicitud de reportes de servicios de mantenimiento frente a cantidad de equipos por tipo, con un valor de 6,4 unidades. Nótese que este grupo posee un total de 128 solicitudes de mantenimiento correctivo y solamente 20 equipos en dicho grupo o categoría (Fig. 1 y 2).

**Tabla 1.** Principales indicadores de la muestra de datos de equipos del hospital.

Tipo de Equipos	Código	No Equipos	CSS	PC en %	Promedio CS/CA en %	TAT promedio (horas) [11]	Promedio AVU	NP [13]
Imágenes	F	23	54	37,20	3,20	1,8	0,2	70
Electrónica médica	C	73	117	15,20	4,30	2,39	0,4	56
Electromecánica	A	122	179	10,10	2,30	1,42	1,2	43
Soprote de Vida	D	23	70	8,60	5,20	2,1	0,3	108
Esterilización	B	20	128	6,10	4,60	6,4	0,18	98
Laboratorio	G	17	33	5,60	9,90	0,79	0,35	67
Fibra Óptica	K	17	9	4,30	7,20	5,1	0,12	32
Odontología	E	41	58	4,30	3,70	4,1	0,7	45
Óptica	I	28	27	3,70	6,30	1,07	0,11	24
Electro óptica	L	5	2	2,70	5,90	0,88	0,54	21
Salón y Vacío	J	9	12	1,10	7,10	0,53	0,8	85
Equipos de medición	H	38	30	1,10	2,10	2,17	0,57	32
Promedio		34,67	59,92	8,33	5,15	2,40	0,46	56,75
Máximo		122,00	179,00	37,20	9,90	6,40	1,20	108,00
Mínimo		5,00	2,00	1,10	2,10	0,53	0,11	21,00

Donde:

CSS es la cantidad de solicitudes de servicios

PC es la penetración del costo de adquisición [10]

AVU es el arribo de la vida útil expresado como la razón entre ET/UL, donde ET es tiempo de explotación y UL es vida útil

NP es el nivel de prioridad, definido en [13]

- La categoría “B” de Tipos de Equipos tiene el mayor valor del indicador TAT, con un valor de 6,4 horas y el segundo mayor valor de prioridad, siendo este de 98 unidades como promedio en toda la categoría (Fig. 3 y 4).
- La categoría “D” de Tipos de Equipos es la que tiene el segundo mayor valor del indicador razón solicitud de reportes de servicios de mantenimiento frente a cantidad de equipos por tipo, con un valor de 3,04 unidades.
- Aunque la categoría “A” de Tipos de Equipos es la que tiene la mayor cantidad de reportes de servicios

de mantenimiento, con 179 casos, el indicador razón solicitud de reportes de servicios de mantenimiento contra cantidad de equipos por tipo es sólo de 1,47 unidades, siendo uno de los menores valores.

- La categoría “C” de Tipos de Equipos exhibe las mismas características que el grupo “A”, con una cantidad de reportes de servicios igual a 117 unidades y un indicador razón de solicitud de reportes de servicios de mantenimiento correctivo contra cantidad de equipos por tipo igual a 1,6 unidades (Fig. 1 y 2).

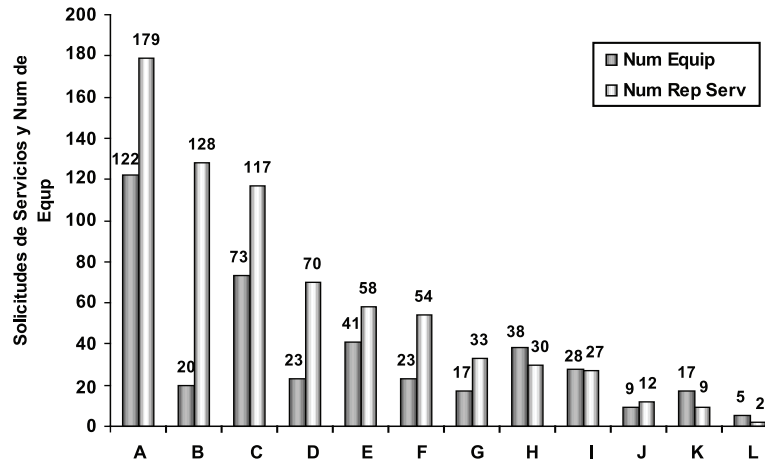


Fig. 1. Distribución por tipos de equipos o categoría del número total de reportes de servicio en comparación con el número total de equipos presentes.

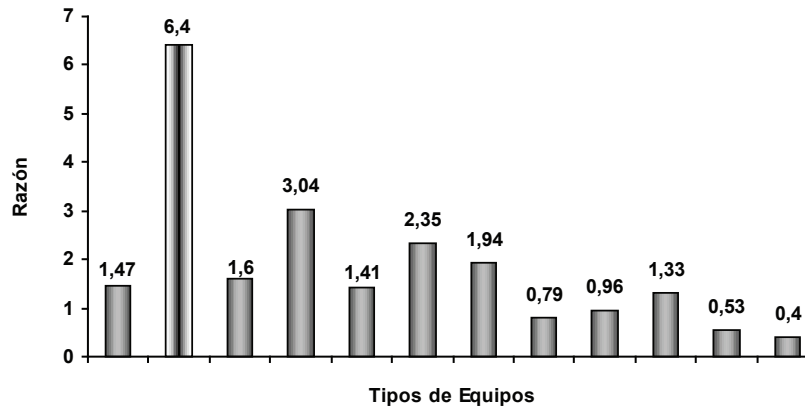


Fig. 2. Razón solicitud reportes de servicios contra cantidad de equipos por tipo de equipos.

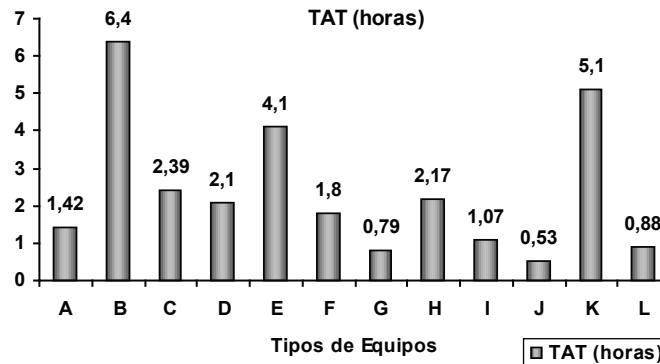


Fig. 3. Valores promedio del indicador TAT (en horas) por tipo de equipos.

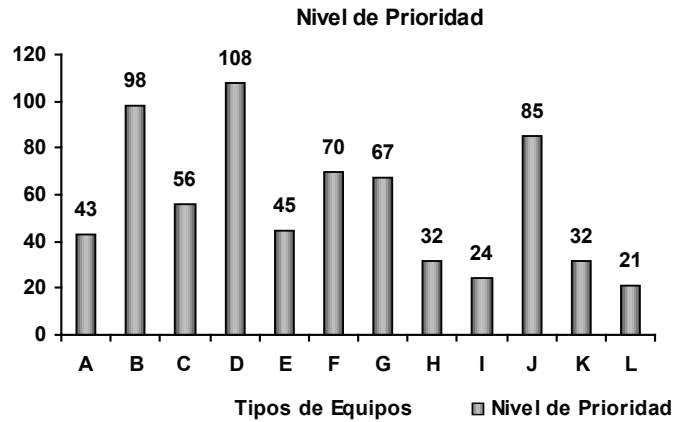


Fig. 4 Valores promedios de la prioridad (en unidades) por tipo de equipos.

Seguidamente se analizó la contribución de los reportes de servicios de mantenimiento correctivo por fabricante. El objetivo de este paso es el de encontrar la existencia o no de un patrón de contribución de los reportes de servicios de mantenimiento de un fabricante específico. Es cierto que analizando la muestra completa de equipos se podría encontrar mucha más información relativa a los fallos y a los reportes de servicios de mantenimiento. Sin embargo, es una práctica común reducir la cantidad total de datos bajo análisis en una muestra representativa llamada tamaño de la muestra. Existen dos métodos conocidos para seleccionar el tamaño de la muestra: métodos probabilísticos y no probabilísticos [15].

Los autores decidieron determinar el tamaño de la muestra usando el método no probabilístico. En este tipo de método, la selección del tamaño de la muestra no depende de la probabilidad; depende de causas relativas muy particulares a la investigación que se está llevando a cabo, o sea del objetivo que se esté persiguiendo [15]. Entonces el tamaño de muestra fue reducido quedándose con los equipos pertenecientes a la categoría “B”. La decisión fue basada en los siguientes criterios:

- La categoría “B” de Tipo de Equipos muestra el mayor valor del indicador razón solicitud de reportes de servicios de mantenimiento correctivo contra cantidad de equipos por tipo, con un valor de 6,4 unidades y el mayor valor en el indicador TAT con 6,4 horas (coincidentalmente el mismo valor).
- La categoría “B” de Tipo de Equipos muestra el segundo mayor valor en la prioridad numérica.
- En esta categoría se encuentran los equipos de esterilización. Un fallo inesperado o parada debidos a un reporte de servicio por mala operación de un usuario hace que se comprometan el servicio de esterilización y el funcionamiento del hospital en su totalidad.

La Fig. 5 muestra la distribución de reportes de servicios de mantenimiento correctivo por fabricante pertenecientes a la categoría de equipos de tipo “B”. Observe cómo los equipos del fabricante “AA” y “BB” tienen la mayor cantidad de reportes de servicios con un 50,78% y 20% del total respectivamente. La Fig. 6 muestra la distribución del número total de reportes de servicios de mantenimientos por modelos pertenecientes al fabricante “AA”.

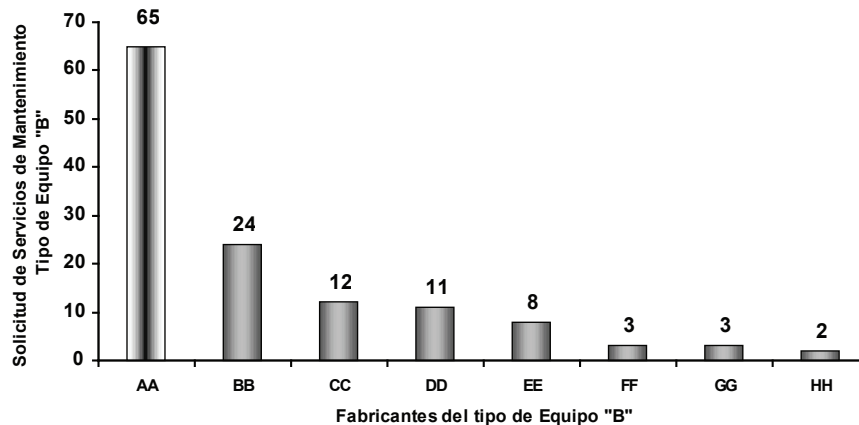


Fig. 5. Distribución del número total reportes de servicios de mantenimiento por fabricante. Nota aclaratoria: se codificó el nombre de los fabricantes, los autores se reservan el derecho de no revelarlos por razones obvias de privacidad.

Se consideraron para continuar con el análisis de los reportes de servicios de mantenimiento de los modelos asociados al fabricante "AA". Los modelos pertenecientes al fabricante "BB" fueron descartados en el análisis posterior, pues los 24 fallos están distribuidos en 17 equipos médicos en todo el período de análisis, y por lo tanto, esto representa, a criterio de estos autores, un patrón normal de comportamiento.

La Fig. 7 muestra el patrón de solicitud de reportes de servicios de mantenimiento en el período del 2002-2006 (siendo las barras pertenecientes al período comprendido entre el 2002-2004). Nótese cómo se observa un patrón muy particular, un bajo número de solicitud de reportes de servicios hacia el año 2002, luego ese número se "dispara", alcanzando un máximo hacia el año 2003 con un total de 36 solicitudes de servicios de mantenimiento (18 por cada modelo) y de nuevo el número total de reportes de servicios vuelve a

caer para el período comprendido entre los años 2005 y 2006.

### 3. Análisis de datos y generación de preguntas

Ahora que se ha cumplido el primer paso se pueden identificar los siguientes patrones en los datos:

- Las categorías de Tipo de Equipos "B", "D", "F", "G" son las que tienen las mayores razones del indicador reportes de servicios de mantenimientos correctivo contra cantidad de equipos en la categoría y sus valores van desde 6,4 a 1,94 unidades (Fig. 2).
- Ya se conocía que la categoría de equipo tipo "B" es la que tiene el mayor valor de indicador de reportes de servicios de mantenimiento contra cantidad de equipos en la categoría, en porcentaje; esto quiere decir que solamente el 4,80% de los equipos del inventario en estudio provoca el 17,8% de los reportes de servicios en el período que se está analizando.

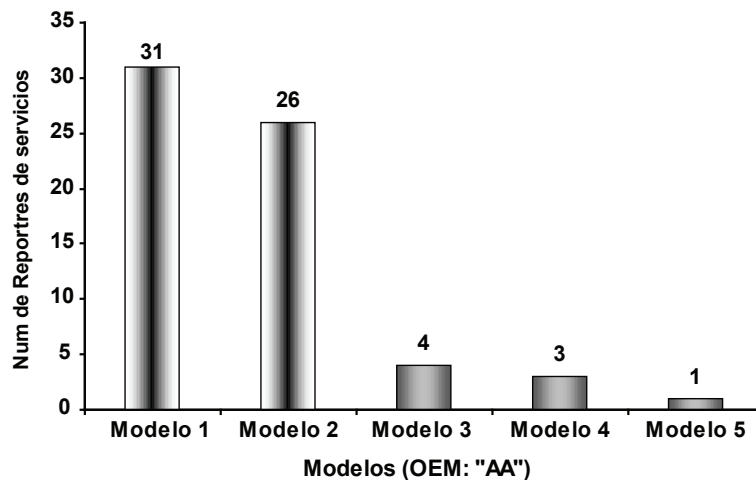


Fig. 6. Distribución del total de solicitud de servicios de mantenimiento por modelos pertenecientes al fabricante "AA".

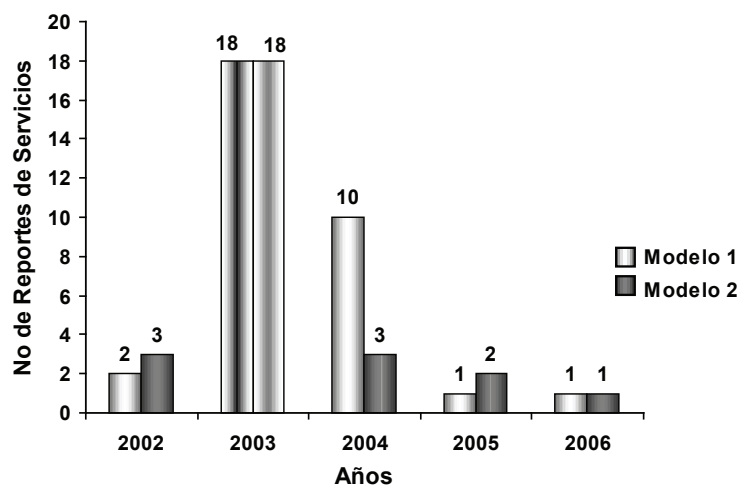


Fig. 7. Patrón de las solicitudes de los reportes de servicios por años pertenecientes a los modelos 1 y 2. Nota aclaratoria: Los 59 reportes de servicios de mantenimiento de los modelos 1 y 2 de esta figura corresponden con las 59 solicitudes de servicios de mantenimiento del fabricante "AA" de la Fig. 6.

- Los modelos 1 y 2 del fabricante “AA” contribuyen con el 44,53% del total de los reportes de servicios de la categoría “B”, los cuales, como ya se conocía, son los equipos de esterilización. Como información adicional se encontró que los modelos 1 y 2 están ubicados en el mismo centro de costo (Estación Central de Esterilización).

Después de obtener estas “pistas” o patrones sobre los datos, se generaron las siguientes preguntas de investigación:

1. ¿Las solicitudes de reportes de servicios de mantenimiento son debidas al mal funcionamiento de los equipos en la mayoría de los casos, o debidas a una falta de entrenamiento de los usuarios? ¿o es una combinación de ambas posibilidades? Si es así, ¿con qué frecuencia ocurren ambos sucesos?
2. ¿Son las solicitudes de mantenimiento por fallos reales del equipamiento producto de frecuencias de mantenimiento programadas mal establecidas? En caso de ser así, ¿las frecuencias de mantenimiento programadas son demasiado altas o bajas?
3. ¿Ocurren las solicitudes de servicios de mantenimiento debido a las causas establecidas por una combinación de las preguntas 1 y 2 formuladas anteriormente?

Para responder estas preguntas de investigación se definió un grupo de variables. Las variables definidas y su propósito dentro del análisis se exponen a continuación (Tabla 2).

- Variable: *Error de operador o usuario*. El propósito de esta variable es encontrar cuáles reportes de servicios de mantenimiento fueron fallos reales y cuáles no lo fueron. Es obvio que este dictamen puede ser dado por el ingeniero clínico después que realiza la orden de trabajo.
- Variable: *Fallo real*. El propósito de esta variable es encontrar cuáles reportes de servicios de mantenimiento fueron fallos reales, sin importar la causa que lo produjo.
- Variable: *Mantenimiento programado* en la semana que se solicitó el reporte de servicio o fallo. El objetivo de esta variable es de determinar la necesidad de un ajuste de la frecuencia de mantenimiento en caso de existir muchas reglas que evidencian fallos en semanas en que no existan mantenimientos programados.
- Variable: *Mantenimiento programado comienza en tiempo o no*. Esta variable tiene un propósito de control y medida de la eficiencia del mantenimiento programado. Por ejemplo, si un reporte de servicio de mantenimiento correctivo es solicitado en un período particular (digamos el viernes) y además es un fallo real no causado por el operador o usuario; y existía un mantenimiento programado en esa semana y no se comenzó o no se realizó en tiempo, pero además este patrón es repetitivo, esto implica dos cosas: primero, la frecuencia de mantenimiento es adecuada; segundo, se debe ejercer más control sobre el cumplimiento de las actividades programadas para que se hagan en el tiempo estipulado.

**Tabla 2.** Variables seleccionadas.

Código de la Variable	Descripción de la variable y tipo	Posibles valores
ErrOp	Error de usuario u operador (ordinal)	TRUE (Error del operador o usuario) (T) FALSE (No error) (F)
FailSSys	Fallos del sistema de suministro (agua o vapor), (ordinal)	TRUE (Fallo del sistema de suministro) (T) FALSE (Otro fallos que no es el sistema de suministro) (F)
Real	Fallo real (ordinal)	TRUE (cuando realmente un fallo ha ocurrido) (T) FALSE (no ha ocurrido un fallo) (F)
Turn	Equipo o turno de trabajo (ordinal)	M (turno de la mañana) AF (turno de la tarde )
Sched	Mantenimiento programado en la semana del fallo (ordinal)	YES (si existía mantenimiento programado en la semana que ocurrió el fallo) (Y) NO (no existía mantenimiento programada en la semana que ocurrió el fallo) (N)
StartSched	Mantenimiento programado comienza en tiempo (ordinal)	YES (si comenzó el mantenimiento programado en tiempo en la semana que ocurrió el fallo) (Y) NO (no comenzó el mantenimiento programado en tiempo en la semana que ocurrió el fallo) (N) DNM (no existía mantenimiento programado)

- Variable: *Equipo de trabajo o turno* en que se realiza el reporte de servicio de mantenimiento. El propósito de esta variable es determinar o discriminar el período exacto en el cual se realiza la mayor cantidad de reportes de servicios de mantenimiento. Esto ayudaría a los gestores a determinar si existe algún patrón de sabotaje en los usuarios o si necesitan entrenamiento.
- Variable: *Fallo del sistema de suministro*. El propósito de esta variable es determinar cuáles son las causas del fallo en cuestión. Es bien conocido que en algunos casos el equipo médico se encuentra en su vida útil, fué adquirido con buenas condiciones contractuales de un buen fabricante y, sin embargo, el equipo falla. Usualmente se olvidan las condiciones ambientales de trabajo en las cuales el equipo está funcionando, tales como la temperatura, alimentación eléctrica, calidad del agua y el vapor [16]. Un patrón de fallos repetitivos del equipo debido al sistema de suministro ayudaría a descartar si el fallo viene “de dentro” o “de fuera” del equipo.

### III. RESULTADOS Y DISCUSIÓN

La Tabla 3 muestra las mejores 12 reglas encontradas. Para una mayor comprensión teórica sobre cómo los valores Soporte, Lift y Confidencia de la Tabla 3 fueron calculados se recomienda ver el Anexo 1.

De las reglas encontradas mostradas en la Tabla 3 se puede responder a las preguntas formuladas.

1. Siempre que un reporte de servicio de mantenimiento correctivo es solicitado en el cual un fallo real ocurre

(Real=TRUE) y no existe mantenimiento programado en esa semana (Sched=NO), que no haya sido debido a un error de operación (ErrOp=FALSE) es 2,33 veces más probable (*lift*) que las demás combinaciones. Esta combinación tiene una frecuencia de ocurrencia del 35,00% (soporte) (Regla 2 en la Tabla 3). En otras palabras, una de las causas más probables por las cuales ocurre un reporte de servicio donde está presente un fallo es por bajas frecuencias de mantenimiento planificado y no por errores de operación.

2. Siempre que un reporte de servicio de mantenimiento correctivo es solicitado en el cual un fallo real ocurre (Real=TRUE), la probabilidad de que no haya sido debido a un error de operación (ErrOp=FALSE) es del 87,00% (confianza). Esta combinación ocurre en el 51,00% (soporte) de todos los casos (57) (Regla 6 en Tabla 3). En otras palabras, cuando un fallo real ocurre es debido usualmente a mal funcionamiento de los equipos y no a mala operación de los usuarios.
3. Siempre que un reporte de servicio de mantenimiento correctivo es solicitado, en el cual un fallo no ocurre (Real=FALSE) se debe a un error de operación (ErrOp=TRUE). Esta combinación ocurre en un 44,00% (soporte) de todos los 59 casos (Regla 4 en la Tabla 3). En otras palabras, cuando un fallo real no está presente como dictamen en una orden de trabajo que surge de un reporte de servicios es debido a errores de operación.

**Tabla 3.** Mejores reglas encontradas.

Regla	Parte izquierda Antecedente	➔	Parte derecha Consecuente	Confidencia	Lift Ratio	Soporte
1	ErrOp=False	➔	Real=True Sched=NO	1,00	2,33	0,35
2	Real=True Sched=NO	➔	ErrOp=False	0,83	2,33	0,35
3	ErrOp=True	➔	Real=False	0,86	2,00	0,44
4	Real=False	➔	ErrOp=True	1,00	2,00	0,4
5	ErrOp=False	➔	Real=True	1,00	1,87	0,51
6	Real=True	➔	ErrOp=False	0,87	1,87	0,51
7	ErrOp=False Sched=NO StartSched=DNM	➔	Real=True	1,00	1,87	0,35
8	Real=True	➔	ErrOp=False Sched=NO StartSched=DNM	0,67	1,87	0,35
9	ErrOp=False StartSched=DNM	➔	Real=True	1,00	1,87	0,35
10	Real=True	➔	ErrOp=False StartSched=DNM	0,67	1,87	0,35
11	ErrOp=False Sched=NO	➔	Real=True	1,00	1,87	0,37
12	Real=True	➔	ErrOp=False Sched=NO	0,67	1,87	0,37



**Tabla 4.** Agrupamientos formados.

Nombre del Conglom	Características del Conglomerado	Variables														
		V1 Modelos		V2 FailSSys		V3 Real		V4 ErrOp		V5 Work Team or Turn		V6 Sched		V7 StartSched		
		M1	M2	T	F	F	T	T	F	M	AF	Y	N	Y	DNM	N
C-0	Número total de elementos: 6 % del total: 11% Probabilidad de pertenencia: 0,11	1	6	5	1	5	1	5	1	5	1	1	5	0	5	1
C-1	Número total de elementos: 20 % del Total: 36% Probabilidad de pertenencia: 0,34	7	13	8	12	19	1	19	2	17	3	1	19	0	19	1
C-2	Número total de elementos: 23 % del Total: 41% Probabilidad de pertenencia: 0,40	23	0	22	1	1	22	1	22	22	2	2	22	1	22	1
C-3	Número total de elementos: 7 % del Total: 13% Probabilidad de pertenencia: 0,12	2	5	7	1	2	6	1	6	5	2	1	6	1	6	0

Donde:

V1 son los Modelos Modelo 1 = M 1 = GE2609 AR-2 y Modelo 2 = M 2 = GA2609 EM-2

V2 es la variable FailSSys (Tabla 2)

V3 es la variable Real (Tabla 2)

V4 es la variable ErrOp (Tabla 2)

V5 es la variable Work Team or Turn (Tabla 2)

V6 es la variable Sched (Tabla 2)

V7 es la variable StartSched (Tabla 2)

La Tabla 4 muestra el número total de agrupamientos obtenidos con las variables definidas anteriormente.

Con la evidencia encontrada se pudieron tomar las decisiones pertinentes, esto a su vez permitió llevar a cabo las siguientes acciones correctivas en el año 2004:

La mayoría de los casos de reportes de servicio de mantenimiento que generaron fallos reales estuvieron relacionados con el sistema de suministro de agua y vapor de los equipos de esterilización (75,00%) y además no existía mantenimiento programado en el momento que ocurrieron esos reportes de servicio (Regla de asociación 1, 7 y 9 en la Tabla 3 y los totales de la variable V2 en la Tabla 4). Se incrementó la frecuencia de mantenimiento de una vez por año (cada 12 meses) a 3 veces por año (cada 4 meses). Las reglas 1, 7 y 9 en la Tabla 3 y la variable V3 en la Tabla 4 ayudan a discriminar el número total de fallos reales y a soportar la evidencia del incremento de la frecuencia de mantenimiento a 3 veces por año.

Esta decisión está sustentada en un análisis estadístico realizado con los 30 reportes de servicios de mantenimiento que fueron dictaminados como “fallos reales” (V3 en Tabla 4 donde se muestran un total de 30 fallos reales en contraposición a 27 que no lo fueron). Donde:

- a) El tiempo medio entre fallos (MTBF) de los fallos reales siguen una distribución normal con parámetros  $N(\mu, \sigma) = N(5,35; 2,32)$  (en meses); con intervalos de confianza para estos parámetros de  $(4,48 \leq \mu \leq 6,22)$  y  $(1,85 \leq \sigma \leq 3,12)$  respectivamente. La Fig 8(a) muestra el histograma de frecuencias y la Fig. 8(b) muestra la función de densidad de probabilidad.
- b) Como 4,48 meses es el límite inferior en el intervalo de confianza del parámetro  $\mu$ , de manera conservadora, se tomó como 4 meses la frecuencia de mantenimiento programado para el sistema de suministro de los equipos de esterilización.

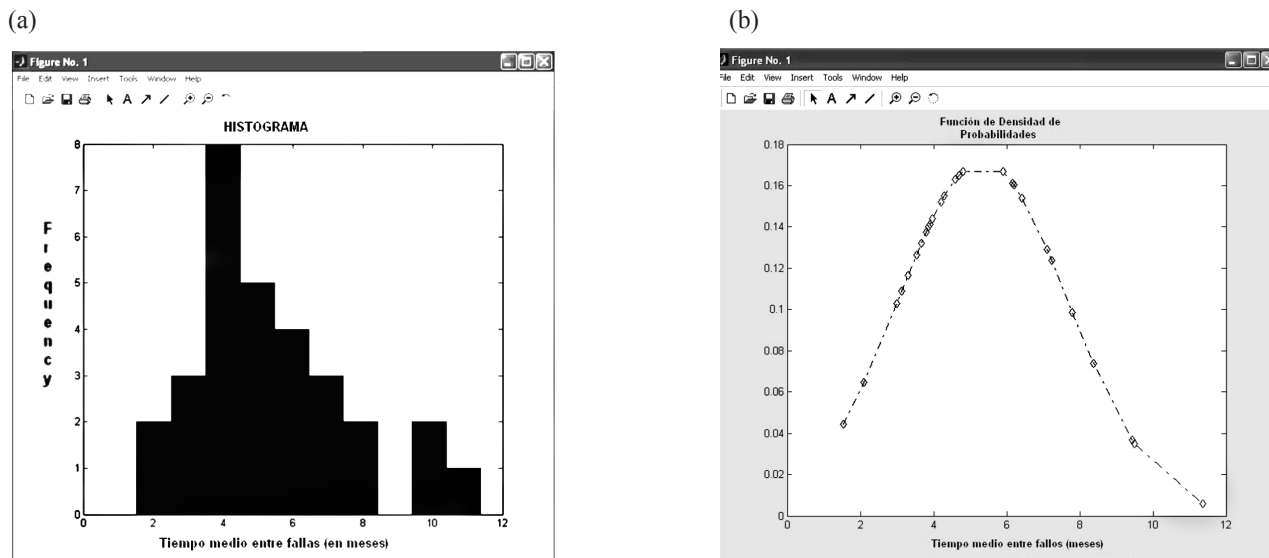


Fig. 8. Distribución estadística del tiempo medio entre fallos. (a) Histograma. (b) Función de densidad de probabilidades.

Como en la mayoría de los casos los reportes de servicios de mantenimiento correctivos que se solicitaron y que no resultaron ser fallos reales ocurrieron en el turno de la mañana (87,50%) (Regla 4 en la Tabla 3 y la variable V5 en la Tabla 4) se llevó a cabo un programa integral de entrenamiento a este turno.

Después de tomar estas acciones correctivas las solicitudes de servicio de mantenimiento correctivos de los modelos 1 y 2 del fabricante “AA” cayeron notablemente del 2004 al período 2005-2006.

Finalmente, las herramientas informáticas empleadas para la realización de esta investigación fueron:

La información fue recopilada, actualizada y mantenida gracias al sistema SMACOR™. Este es un Sistema de Mantenimiento Asistido por Computador (SMAC) disponible en [17]<sup>1</sup>.

El procesamiento computacional para el análisis de agrupamiento y descubrimiento de las asociaciones se llevó a cabo usando el programa informático WEKA® versión 3.4.7 disponible en [9].

El procesamiento estadístico se llevó a cabo usando el Toolbox Estadístico de MathWorks, Inc [18].

La computadora personal donde se llevó a cabo el procesamiento computacional fue una Pentium IV PC

a 2,3 GHz con 512 MB RAM. El tiempo medio que demoró la corrida para la construcción de los modelos computacionales fue de 30-50 segundos.

#### IV. CONCLUSIÓN

1. Después de realizar la segmentación de la muestra de datos, las reglas de asociación demostraron ser un método efectivo para encontrar las causas más generales por las cuales se solicitaban los reportes de servicios de mantenimiento. Éstas fueron: inadecuadas políticas de frecuencias de mantenimiento programado, falta de entrenamiento de los usuarios y fallos intrínsecos en los equipos de esterilización.
2. La aplicación de la técnica de agrupamientos demostró ser un método efectivo para mejorar la eficiencia del servicio de esterilización del hospital en estudio. Por una parte, permitió discriminar la causa de los fallos reales intrínsecos de los equipos de esterilización, originados en los fallos en el sistema de suministro de agua y vapor, y, por otra parte contribuyó a que se tomaran medidas correctivas para solucionar el problema. El indicador en estudio disminuyó de un valor de 6,4 a 0,4 unidades. Esto representó un mejoramiento de 16 veces.

<sup>1</sup> Estos sistemas son comúnmente conocidos por sus siglas en inglés como CMMS (Computerized Maintenance Management System). En español se conocen como GMAO o GMAC (sistemas de Mantenimiento Asistido por Ordenador o Computadora). Actualmente existe una versión de este producto llamada SYSMANCOR registrada en la oficina de derecho de autor de la Republica de Colombia, para obtener versión de este producto contactar al autor principal de este trabajo.

## V. REFERENCIAS

- [1] Grimes L.S. The Future of Clinical Engineering: The Challenge of Change. *IEEE Engineering in Medicine and Biology Magazine*, 22, 91, March-April, 2003.
- [2] Two Crows Corporation. Introduction to Data Mining and Knowledge Discovery, 3<sup>rd</sup> Ed., Maryland, 1999.
- [3] Witten, L.H., Franck, E. **Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations**. *Morgan Kaufmann, 2000. Cap. 1, 80*.
- [4] Casebank Technologies Inc. Available at: <http://www.casebank.com/about/aerospace.asp>
- [5] Exclusive Ore, Data Mining in Equipment Maintenance and Repair: Augmenting PM with AM. 2007. Available at: <http://www.exclusiveore.com>
- [6] O’Dea Schlabig W., Williams J.S., Larissa H. Clinical Engineering Management: An Annotated Bibliography 1989-1993. *Biomedical Instrumentation and Technology*, 10, 101-111, 1994.
- [7] ECRI. Health Technology. *Special report on managing service contract*, 3, 4, Winter, 89-95, 1989.
- [8] Agrawal R. and Srikant R. Fast Algorithms for Mining Association Rules. In Proc. 1994 Int. Conference Very Large Data Bases (VLDB’94), 487-499, Santiago, Chile, septiembre 1994.
- [9] Weka, The University of Waikato. Available at: <http://www.cs.waikato.ac.nz/ml/weka>
- [10] Cruz M.A., Barr C., Pozo E.P. **Improving Corrective Maintenance Efficiency in Clinical Engineering Departments**, *IEEE Engineering in Medicina and Biology*, 26, 3, May/June, 60-65, 2007.
- [11] Cohen T. Benchmark indicators for medical equipment repair and maintenance. *Biomedical Instrumentation and Technology*, 29, 308-320, 1995.
- [12] Cruz M.A., Rodriguez D.E., Sanchez V.M.C, Manuel L. An event-tree-based mathematical formula for the removal of biomedical equipment from a hospital inventory. *Journal of Clinical Engineering*. Winter, 27, 1, 63-71, 2005.
- [13] Capuano M., Koritko S. Risk Oriented Maintenance System. *Biomedical Instrumentation and Technology*, 30, 1, 25-35, 1996.
- [14] Hernandez S. R, Fernandez C. C, Baptista L. P, Research Methodology 4th edition, McGraw-Hill, 2006. Chapter 4, 235-253.
- [16] Mosquera C.G. and et al. Availability and reliability of industrial systems. Venezuela. 1995, 52-60.
- [17] [online] Available at <http://www.gestecno.sld.cu>
- [18] User’s Guide Version 2. Fuzzy Logic Toolbox for use with MATLAB 5.3.

## ANEXO 1

Las técnicas de descubrimiento de asociaciones forman reglas con los registros de datos en los cuales están involucradas las variables bajo estudio. Estas reglas siguen la forma  $A \rightarrow B$ , donde A es el antecedente de la regla o la parte izquierda (LHS), y la parte B es el consecuente de la regla o parte de derecha (RHS). Por ejemplo, en la siguiente regla de asociación “siempre que un usuario reporta un servicio de mantenimiento entonces el equipo ha fallado”, el antecedente es “solicita un servicio” y el consecuente es “el equipo ha fallado”. Luego, en un grupo de transacciones de una base de datos donde están involucradas las variables en estudio es fácil determinar la proporción de transacciones que poseen un valor específico de dicha variable (por ejemplo  $ErrOP = TRUE$ ): sencillamente contando. Para cada regla encontrada se calcula, existe un soporte (*support*), una confianza (*confidence*), y una relación entre el antecedente y consecuente (*lift*) entre todas las reglas.

El soporte se refiere a la presencia o frecuencia de una regla en particular. Por ejemplo, si se tienen 15 transacciones de 1000 reportes de servicios en las cuales las variables “Fallo Real es TRUE y Error del Operador o Usuario es FALSE”, entonces se dice que el soporte de la asociación es de 1,5% ( $100 \cdot (15/1000)$ ). Un valor muy bajo del parámetro soporte (digamos uno en un millón de transacciones) pudiera indicar que esa regla en particular no es muy importante.

Ahora bien, para poder descubrir reglas significativas, se debe buscar la frecuencia relativa de una variable en la transacción en combinación con otras. Esto es, dada la ocurrencia del elemento o variable “A” en la transacción, ¿cuán frecuentemente ocurre la variable “B” también? En otras palabras, ¿cuál es la probabilidad predictiva condicional de “B” dado “A”? Siguiendo con el ejemplo anterior sería, “cuando un usuario u operador reporta un servicio (con fallo real), ¿cuán frecuentemente comete un error que causa dicho fallo (variable  $ErrOP = TRUE$ )”.

Para esto se define la confianza. Esta se ofrece en forma de razón como frecuencia de la variable “A” y “B”/ frecuencia “A”. Ejemplifiquemos los conceptos anteriores con una base de datos hipotética con un tamaño de muestra de 1000 transacciones de reportes de servicios con las siguientes condiciones:

- 60 transacciones correspondientes a reportes de servicios en los cuales el error de operación está presente, esto es la variable  $ErrOP = TRUE$ .
- 70 transacciones correspondientes a reportes de servicios en los cuales el error de operación no está presente, esto es la variable  $ErrOP = FALSE$ .
- 200 transacciones correspondientes a reportes de servicios en los cuales un fallo real del equipo médico está presente, esto es la variable  $Real = TRUE$ .

- 20 transacciones correspondientes a reportes de servicios en los cuales un fallo real del equipo medico no está presente, esto es la variable Real= FALSE.
- 15 transacciones correspondientes a reportes de servicios en los cuales el error de operación de un usuario está presente y un fallo real del equipo medico está presente. Esto es, las variables ErrOP= TRUE y Real= TRUE.
- 15 transacciones correspondientes a reportes de servicios en los cuales el error de operación de un usuario está presente y un fallo real del equipo medico está presente. Esto es, las variables ErrOP= FALSE y Real= TRUE.

Ahora bien el cálculo del soporte y la confianza es como sigue:

$$\text{Soporte para (ErrOP= TRUE)} = 60/1000=6\%$$

$$\text{Soporte para (ErrOP= FALSE)} = 70/1000 = 7\%$$

$$\text{Soporte para Real= TRUE}=200/1000=20\%$$

$$\text{Soporte para Real= FALSE} = 20/1000 = 2\%$$

$$\text{Confianza para (ErrOP = TRUE} \rightarrow \text{Real = TRUE)} = 15/60=25\%$$

$$\text{Confianza para (ErrOP = FALSE} \rightarrow \text{Real= TRUE)} = 15/70=21,4\%$$

Entonces de los ejemplos anteriores uno puede decir que: cuando un usuario u operador reporta un servicio de mantenimiento, la probabilidad de que un error de operación cause un fallo real en el equipamiento es del (25%) y es mayor que la probabilidad de que un error no esté presente y se provoque un fallo real (21,4%).

El *Lift* es otra medida de la potencia o solidez de una asociación. Cuanto mayor sea este parámetro, mayor será la influencia que tiene la ocurrencia de "A" sobre la probabilidad de ocurrencia de B. El Lift es calculado en forma de razón según la forma (confianza de A→B)/(Soporte de B), así continuando con el ejemplo anterior se tiene que:

$$\text{Lift de (ErrOP = TRUE} \rightarrow \text{Real= TRUE)} = (25\%)/(6\%)= 4,17$$

$$\text{Lift de (ErrOP = FALSE} \rightarrow \text{Real= TRUE)} = (21,4\%)/(6\%)= 3,57$$