

ANÁLISIS ESTADÍSTICO Y DETERMINACIÓN DE LOS FACTORES QUE INFLUYEN EN EL RETRASO DE LA TUBERCULOSIS PULMONAR. CASO: HOSPITAL MILITAR DE GUAYAQUIL

María Luisa Conforme Yagual¹ Soraya Solis²

¹Ingeniera en Estadística Informática; e-mail: mconform@espol.edu.ec

²Directora de Tesis, Master en Docencia y Educación, ESPOL, 2004; Ingeniera en Computación, ESPOL, 1998; e-mail: ssolis@espol.edu.ec

Resumen: El presente trabajo muestra un análisis estadístico que determina cuáles son los factores que influyen en el retraso del diagnóstico de la Tuberculosis Pulmonar. El estudio fue basado en la información registrada en el Hospital Militar de Guayaquil.

En la primera parte se realiza una descripción de la enfermedad. En el siguiente apartado se detallan cuáles son los factores relacionados con la enfermedad. En la tercera parte se presenta las herramientas estadísticas utilizadas en el estudio. Seguidamente, se muestra el análisis univariado de los datos. En la última parte se muestran las distribuciones de probabilidad conjunta de las variables, el análisis de contingencia que determina la independencia de las variables y el método de Regresión Logística Binomial que determina los factores influyentes en el retraso.

Summary: The present work shows a statistical analysis that determines which are the factors that influence in the delay of the diagnosis of the Pulmonary Tuberculosis. The study was based on the information registered in the Military Hospital of Guayaquil.

The first part presents a description of the illness. The following section details which are the factors related with the illness that were gathered of each patient's clinical history. In the third part, the statistical tools used in the study are exposed. Subsequently, the univariate analysis of the data is shown. The last section shows the distributions of combined probability of the variables, the contingency analysis that determines the independence of the variables and the method of Binomial Logistical Regression that it determines the influential factors in the delay.

INTRODUCCIÓN

La tuberculosis constituye la segunda causa de mortalidad en el mundo ⁽¹⁾. En 1993, la Organización de Salud Mundial (OMS) declaró un estado de emergencia global para la tuberculosis debido al sostenido crecimiento de la enfermedad en el mundo ⁽²⁾. Se estima que para finales del año 2005, 12 millones de casos de tuberculosis se identificarán en el mundo, es decir un aumento del 58%, de los 7.5 millones estimados para 1990 ⁽³⁾. En el Ecuador, para el año 2004, en la provincia del Guayas se reportaron 2688 casos ⁽⁴⁾. El diagnóstico temprano y el tratamiento correcto de los enfermos de Tuberculosis Pulmonar hasta su curación, son las finalidades esenciales de la lucha antituberculosa ⁽⁵⁾. Durante el tiempo que un enfermo de tuberculosis pulmonar no es diagnosticado ni tratado sus lesiones pueden progresar y verse abocado a un fatal desenlace ⁽⁶⁾.

OBJETIVO GENERAL DEL ESTUDIO

El objetivo de este estudio es conocer y evaluar los factores asociados al retraso diagnóstico total de la tuberculosis en pacientes sintomáticos no tratados anteriormente.

OBJETIVO ESPECÍFICO DEL ESTUDIO.

- ⊕ Conocer cuál es la oportunidad que tienen los pacientes de incurrir en un retraso en el diagnóstico cuando tienen uno de los factores influyentes.
- ⊕ Determinar qué cantidad de la población obtuvo excesos en cada uno de las etapas del diagnóstico total de la enfermedad.

OBJETO DE ESTUDIO.

Se definió como objeto de estudio a todo aquel paciente en quien se aisló el agente

causante de la TB pulmonar, Mycobacterium Tuberculosis, para lo cual era necesario que se cumpla cualquiera de los siguientes criterios:

- Pacientes en los que existían síntomas o signos clínicos compatibles con TB Pulmonar.
- Pacientes en los que existía evidencia radiológica de la enfermedad.
- Pacientes con baciloscopia de esputo o jugo gástrico cuyo resultado era positivo.

El estudio se llevó a cabo en el Hospital Militar de Guayaquil- Ecuador, los datos fueron recopilados en el Laboratorio Clínico y en el departamento de Estadística y Archivo. Se consideró como caso nuevo aquel que nunca había realizado tratamiento antituberculoso y que acudieron a consulta desde el 1 de enero del 2000 hasta el 30 de Junio del 2004.

1. BREVE DESCRIPCIÓN TÉCNICA DE LA TUBERCULOSIS PULMONAR

La tuberculosis pulmonar es una enfermedad crónica causada por Mycobacterium Tuberculosis. Las partes del pulmón más afectadas son los segmentos apicales y posteriores de los lóbulos superiores, y los segmentos superiores de los lóbulos inferiores. La infección pulmonar tiene generalmente un comienzo insidioso. Cuando el paciente presenta los primeros síntomas, la enfermedad puede estar ya muy avanzada desde el punto de vista radiológico.

El hombre es el principal reservorio de Mycobacterium tuberculosis. En teoría, el período de transmisibilidad dura todo el tiempo que se expulsan en el esputo bacilos tuberculosos viables. Algunos enfermos no tratados o tratados de manera inadecuada pueden mostrar intermitentemente bacilos en el esputo, durante años. El grado de transmisibilidad depende del número de bacilos expulsado y de su virulencia, la suficiencia de la ventilación, la exposición de los bacilos al sol o a la luz ultravioleta, y las oportunidades para dispersarse en aerosol por tos, estornudos, habla o canto.

Los enfermos con Tuberculosis pulmonar activa, con la espiración, al hablar, cantar, reír, estornudar y, sobre todo, al toser producen aerosoles contaminantes. Las gotas de secreción, en el exterior, pierden una parte de su contenido acuoso por evaporación y dejan el núcleo de las gotas con uno o pocos bacilos, que son los verdaderos vehículos de transmisión ya que quedan en suspensión en el aire y se dispersan sin dificultad. Los mecanismos de defensa del aparato respiratorio son incapaces de impedir que, cuando estos núcleos contaminantes son inhalados, lleguen

hasta los alveolos pulmonares, donde los bacilos encuentran las condiciones para multiplicarse.

El período de incubación puede durar entre 5 a 12 semanas desde que se produce la infección hasta que aparece la lesión primaria. El riesgo de presentar la enfermedad es máximo en los niños menores de 3 años, más bajo en etapas posteriores de la niñez y de nuevo alto en los adolescentes, los adultos jóvenes y los muy ancianos.

La Tuberculosis Pulmonar suele manifestarse habitualmente con tos productiva de larga evolución. El esputo suele ser escaso y no purulento. Conjuntamente, puede existir dolor torácico, y en ocasiones hemoptisis. Ésta última, es indicativa de enfermedad avanzada. Además se presenta otros síntomas como: Fiebre, astenia, pérdida de apetito, disminución de peso, sudoración nocturna excesiva.

2. RECOPIACIÓN Y DESCRIPCIÓN DE LOS FACTORES RELACIONADOS CON LA ENFERMEDAD.

2.1 LEVANTAMIENTO DE INFORMACIÓN.

Para el levantamiento de información se realizaron varias visitas al Laboratorio del hospital con la finalidad de registrar todas las baciloscopias de esputo que hayan resultado positivas para el Bacilo de Koch. Posteriormente se buscó en el departamento de Estadística y Archivo el número de la historia clínica de los pacientes con exámenes positivos.

Otra de las herramientas utilizadas para la captura de información fue la revisión de los partes médicos de los departamentos de Neumología y Medicina General, con el objetivo de encontrar pacientes cuya exploración física muestre síntomas compatibles con TB. Luego se buscó nuevamente en el departamento de Estadística y Archivo el número de la historia clínica de los pacientes.

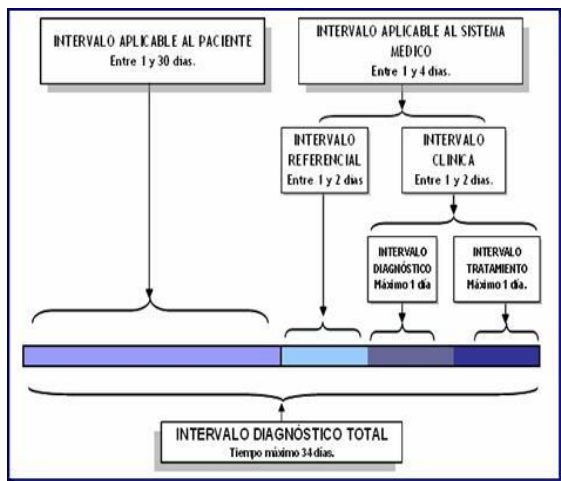
Posteriormente de los números de historias clínicas recopilados se seleccionó solo los casos nuevos. Específicamente la información que se recopiló fue:

- Datos Generales del paciente.
- Hábitos tóxicos.
- Sintomatología compatible con TB Pulmonar y resultados en la Auscultación Pulmonar (presencia o ausencia de sonidos anómalos).
- Diagnóstico Radiológico.
- Fecha de aparición de los primeros síntomas.
- Fecha de las visitas y exploraciones practicadas.
- Fecha de la primera sospecha de TB.
- Fecha de inicio del tratamiento específico.

2.2 DESCRIPCIÓN DE FACTORES: INTERVALOS DE TIEMPO DEL PROCESO DE DIAGNÓSTICO TOTAL.

A continuación se describen las variables utilizadas para medir el tiempo del proceso de diagnóstico total de la tuberculosis pulmonar y el valor límite de cada período.

GRÁFICO 1



"Esquema de los intervalos de tiempo del proceso de diagnóstico total de la Tuberculosis pulmonar"

➤ Intervalo Aplicable al paciente (IAP):

Días acontecidos entre el inicio de los síntomas compatibles con tuberculosis pulmonar hasta que el enfermo solicita la debida atención. Demora por parte del paciente fue considerada en intervalos mayores a 30 días.

➤ Intervalo Referencial (IR):

Tiempo transcurrido desde la primera consulta hasta su solicitud de admisión o el establecimiento de sospecha de tuberculosis pulmonar. Más de dos días fueron considerados como demora.

➤ Intervalo Diagnóstico (ID):

Tiempo transcurrido desde que se sospecha la existencia de tuberculosis pulmonar o desde que el paciente es hospitalizado hasta que se toma la primera muestra para TB. Tiempo superior a 1 día fue considerado como exceso.

➤ Intervalo Tratamiento (IT):

Espacio de tiempo que acontece desde que se toma la primera muestra hasta que se inicia el tratamiento. Intervalos mayores a un día fueron considerados como demora.

➤ Intervalo Clínica (IC):

Es la suma de los intervalos de diagnóstico y de tratamiento. Espacios de tiempo mayores a dos días fueron considerados como demora.

➤ Intervalo Aplicable al Sistema Médico (IASM):

Es la suma del intervalo referencial, diagnóstico y tratamiento. Tiempos mayores a cuatro días fueron interpretados como demora.

➤ Intervalo Diagnóstico Total (IDT):

Tiempo transcurrido desde el inicio de los síntomas hasta el diagnóstico de la enfermedad e inicio del tratamiento, es decir es la suma del intervalo aplicable al paciente y del intervalo Doctor. La demora en este intervalo fue considerada en periodos mayores a 34 días.

2.3 DESCRIPCIÓN DE FACTORES: OTRAS VARIABLES RELACIONADAS CON EL ESTUDIO.

Datos Generales del paciente.

- Ⓞ Género.
- Ⓞ Edad.
- Ⓞ Cobertura por Seguro Médico.
- Ⓞ Nivel de Instrucción.
- Ⓞ Grupo étnico.
- Ⓞ Nivel Socio-económico
- Ⓞ Área de residencia
- Ⓞ IMC (Índice de Masa Corporal).
- Ⓞ Historia de Contacto.
- Ⓞ Hospitalizado para manejo tuberculoso.
- Ⓞ Enfermedad Coexistente.
- Ⓞ Antecedentes de Atopia.

Hábitos Tóxicos.

- Ⓞ Consumo de Alcohol.
- Ⓞ Fumador Activo.
- Ⓞ Fumador Pasivo.

Sintomatología.

- Ⓞ Tos.
- Ⓞ Expectoración.
- Ⓞ Tipo de Expectoración.
- Ⓞ Fiebre.
- Ⓞ Pérdida de peso.
- Ⓞ Hemoptisis.
- Ⓞ Disnea
- Ⓞ Cefalea.
- Ⓞ Mialgia.
- Ⓞ Dolor Torácico.
- Ⓞ Anomalía en la auscultación pulmonar.

Diagnóstico radiológico.

- Ⓞ Forma de Representación Radiológica.
- Ⓞ Extensión radiológica.

3. HERRAMIENTAS ESTADÍSTICAS UTILIZADAS EN EL ANÁLISIS DE LOS DATOS.

Para el análisis univariado de los datos se han empleado los procedimientos de la estadística descriptiva referente al uso de **tablas de frecuencia y cálculo de medidas de posición, tendencia central y de dispersión**. Todos los datos son expresados en **media y desviación estándar** para las variables continuas y en **porcentajes** para las variables categóricas. En el análisis bivariado de los datos se determinaron las distribuciones

conjuntas de las variables y las **tablas de contingencia** que por medio de la **prueba Chi-cuadrado** permite determinar independencia entre un par de variables. Asimismo se realizó un análisis estadístico multivariante mediante el procedimiento de **Regresión Logística Binomial**.

3.1 DESCRIPCIÓN DE LAS TÉCNICAS ESTADÍSTICAS.

Distribuciones Conjuntas: Tablas Bivariadas.

Las tablas bivariadas permiten realizar un análisis en conjunto entre algunos pares de las variables. Una tabla bivariada es un arreglo ordenado de r filas y c columnas, donde las filas indican los valores que toma una variable aleatoria X y las columnas determinan de la misma manera los valores que toma la variable Y. El objetivo principal de esta técnica es determinar la distribución conjunta entre cada par de valores que toman ambas variables aleatorias, es decir: $f(x_i, x_j) = P(X_i=x_i, X_j=x_j)$

Análisis de Contingencia

El análisis es presentado por medio de una tabla de contingencia, que es un arreglo bidimensional en el que se detalla los factores a ser analizados con igual o diferentes niveles de información.

Sea A el primer factor con r niveles de información y B el segundo factor con c niveles de información, se define el modelo de tabla de contingencia en el Grafico 2.

GRÁFICO 2

FACTOR A \ FACTOR B	FACTOR B				X_i
	Nivel 1	Nivel 2	...	Nivel c	
Nivel 1	X_{11} E_{11}	X_{12} E_{12}	...	X_{1c} E_{1c}	$X_{1.}$
Nivel 2	X_{21} E_{21}	X_{22} E_{22}	...	X_{2c} E_{2c}	$X_{2.}$
⋮	⋮	⋮	⋮	⋮	⋮
Nivel r	X_{r1} E_{r1}	X_{r2} E_{r2}	...	X_{rc} E_{rc}	$X_{r.}$
$X_{.1}$	$X_{.1}$	$X_{.2}$...	$X_{.c}$	$X_{..}$

"Esquema de una tabla de contingencia"

Donde:

X_{ij} es el número de valores observados que poseen simultáneamente la i-ésima característica del factor A y la j-ésima característica del factor B.

E_{ij} es el número de observaciones esperadas con la i-ésima característica del factor A y la j-ésima característica del factor B, si H_0 es verdadera y se lo obtiene de la siguiente manera:

$$E_{ij} = \frac{X_{i.} * X_{.j}}{n} = \frac{\sum_{i=1}^r X_{ij} * \sum_{j=1}^c X_{ij}}{n}$$

$X_{i.}$ es el número de observaciones que poseen la característica i-ésima del factor B.

$X_{.j}$ es el número de observaciones que poseen la característica j-ésima del factor A.

$X_{..}$ es el número total de observaciones.

Luego de obtener la Tabla de Contingencia se realiza el siguiente contraste de hipótesis:

H_0 : Los factores A y B son independientes
vs.

H_1 : No es verdad H_0

Se puede probar que el estadístico:

$$\chi^2 = \sum_{i=1}^h \sum_{j=1}^k (X_{ij} - E_{ij})^2 / E_{ij}$$

tiene una distribución Ji cuadrado con $(r-1)(c-1)$ grados de libertad, por lo que se rechaza la hipótesis nula a favor de la hipótesis alternativa con $(1-\alpha)100\%$ de confianza si .

$$\chi^2 > \chi_{\alpha}^2 (r-1)(c-1)$$

Regresión Logística

El objetivo primordial que resuelve esta técnica es el de modelar cómo influye en la probabilidad de aparición de un suceso, habitualmente dicotómico, la presencia o no de diversos factores y el valor o nivel de los mismos. La regresión logística es útil cuando se trata de predecir el valor de una variable respuesta dicotómica Y, que presumiblemente depende de otras m variables explicativas (X_j , $j = 1, \dots, m$) a través del modelo de probabilidad

$$\Pr \{Y_i = 1\} = \frac{1}{1 + \exp(-\beta_0 - \beta_1 x_1 - \dots - \beta_m x_m)}$$

Para simplificar la interpretación del modelo de regresión logística es conveniente llegar a cierto convenio en la codificación de variables.

⊕ En la variable dependiente se codifica como 1 la ocurrencia del evento de interés y como 0 la ausencia.

⊕ Las variables independientes pueden ser varias y cada una de un tipo diferente. A continuación se analiza cada caso:

⊙ Caso Dicotómico: Se codifica como 1 el caso que se cree favorece la ocurrencia del

evento. Se codifica como 0 al caso contrario.

- Caso Categórico: Cuando la variable categórica puede tomar más de 2 valores posibles podemos codificarlas usando variables indicadoras.
- Caso de Variable Numérica: Si creemos que la variable numérica puede afectar la respuesta debemos categorizar la variable.

El modelo de regresión logística es emitido por los programas estadísticos bajo el esquema que se presenta a continuación.

GRÁFICO 3

VARIABLE INDEPENDIENTE	PARÁMETROS					INTERVALO DE CONFIANZA AL 95% PARA EXP(B)	
	B	ERROR TÍP.	WALD	VALOR P	EXP(B)	LÍMITE INFERIOR	LÍMITE SUPERIOR

Esquema de los parámetros estimados en la regresión Logística

- En la primera columna, B, se obtienen los coeficientes estimados del modelo regresión logística.
- En la segunda columna, Error Típico, se obtienen los errores estándar de los coeficientes estimados para cada variable.
- En la tercera columna se obtiene el estadístico de Wald para cada variable, el cual se obtiene de la siguiente manera o forma:

$$\text{Estadístico de Wald} = (B / \text{error típico})^2$$

Si Wald es distinto de cero se dice que las variables son significativas y por el contrario, si es igual a cero las variables no son significativas, por lo que se podrían eliminar del modelo.

- En la cuarta columna, Valor p, se obtienen los niveles de significación de las variables, en donde su interpretación nos dice que si el nivel de significación es igual a cero las variables son significativas, por el contrario si el nivel de significación es distinto de cero, se dice que las variables no son significativas para el modelo.

- En la quinta columna Exp(B) se obtienen los coeficientes estimados exponencialmente, el cual indica la razón entre la probabilidad del

éxito por la probabilidad del fracaso, y si este es igual a 1 se puede señalar que la variable no contribuye en el modelo y la oportunidad de riesgo es igual para los individuos que tienen el factor riesgo como para los que no lo tienen.

- Las últimas columnas contienen un intervalo de confianza para el Exp (B), si en ellos no se contiene al valor uno, es señal de que la variable es de interés en el modelo

4. ANÁLISIS UNIVARIADO DE LOS DATOS.

Variable Género: Observando la tabla I y el Gráfico 4 se aprecia que aproximadamente el 55% de los pacientes pertenecen al género masculino mientras que un 45% fue del género femenino.

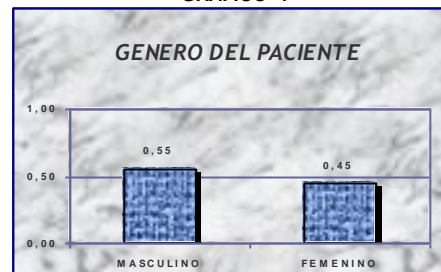
TABLA I

Distribución de frecuencias de "Género"

Género	Frecuencia Absoluta	Frecuencia Relativa
Masculino	60	0,55
Femenino	50	0,45
Total	110	1,00

Fuente: Hospital de División Regional de la Segunda Zona Militar de Guayaquil-Ecuador
Elaborado por: María Luisa Contreras Yagual

GRÁFICO 4



Representación en barras de "Género"

Variable Edad: En la tabla II se observa que la edad promedio, en años. La edad mínima es 15 años y la edad máxima es 93 años.

TABLA II

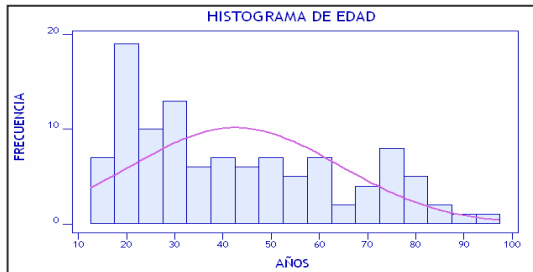
Estadística Descriptiva de "Edad"

MEDIA	42,67
MEDIANA	37,5
DESVIACIÓN ESTANDAR	21,62
ASIMETRIA	0,550
CURTOSIS	-0,934
VALOR MINIMO	15
VALOR MAXIMO	93

Fuente: Hospital de División Regional de la Segunda Zona Militar de Guayaquil-Ecuador
Elaborado por: María Luisa Contreras Yagual

El gráfico 5 muestra que la distribución es sesgada a la derecha, es decir se presenta con menos frecuencia valores después de la media, además la distribución es platicúrtica porque su coeficiente de curtosis es -0,934.

GRÁFICO 5



HISTOGRAMA DE EDAD

En la tabla III se observa que el 68,2% de los pacientes tuvo la cobertura por seguro médico militar. Un 70,9% de la población reportó un nivel de instrucción secundaria. Además, el 59,1% del total de pacientes tuvo un nivel socio-económico medio. También se aprecia que un 51,8% del total de pacientes fue hospitalizado para manejo tuberculoso.

TABLA III

Distribución de frecuencia de otras características del paciente-1

VARIABLES		FRECUENCIA ABSOLUTA	FRECUENCIA RELATIVA
COBERTURA POR SEGURO MEDICO MILITAR	SI	75	0,682
	NO	35	0,318
GRUPO ÉTNICO	INDIGENA	13	0,118
	MESTIZO	81	0,736
	NEGRO	8	0,073
	BLANCO	8	0,073
NIVEL DE INSTRUCCIÓN	PRIMARIA	22	0,200
	SECUNDARIA	78	0,709
NIVEL SOCIO-ECONÓMICO	SUPERIOR	10	0,091
	BAJO	43	0,391
	MEDIO	65	0,591
ÁREA DE RESIDENCIA	ALTO	2	0,018
	URBANA	67	0,609
ÍNDICE DE MASA CORPORAL	RURAL	43	0,391
	DESNUTRIDO	44	0,400
	NORMAL	50	0,455
HISTORIA DE CONTACTO	SOBREPESO	7	0,064
	OBESO	9	0,082
HOSPITALIZADOS PARA MANEJO TUBERCULOSO	SI	13	0,118
	NO	97	0,882
ENFERMEDAD COEXISTENTE	SI	57	0,518
	NO	53	0,482
ANTECEDENTE DE ATOPIA	SI	24	0,218
	NO	86	0,782
	SI	4	0,036
	NO	106	0,964

Fuente: Hospital de División Regional de la Seguridad Zona Militar de Guayaquil-Ecuador.
Elaborado por: María Luisa Contreras Yagala

En la tabla IV se aprecia que el 68,2% de los pacientes tuvo cobertura del seguro médico militar. Un 70,9% reportó un nivel de instrucción secundaria. En cuánto al nivel socio-económico el 59,1% fue representado por la clase media. Además se observa que un 51,8% fue hospitalizado para manejo tuberculoso.

Tabla V
Distribución de frecuencia de otras características del paciente-2

VARIABLES		FRECUENCIA ABSOLUTA	FRECUENCIA RELATIVA
TOS	SI	106	0,964
	NO	4	0,036
EXPECTORACIÓN	SI	81	0,736
	NO	29	0,264
TIPO DE EXPECTORACIÓN	NINGUNA	29	0,264
	BLANQUECINA	10	0,091
	MUCOPURULENTA	57	0,518
	HEMOPTOICA	10	0,091
FIEBRE	MUCOPURULENTA Y HEMOPTOICA	4	0,036
	SI	57	0,518
PERDIDA DE PESO	NO	53	0,482
	SI	42	0,382
HEMOPTISIS	NO	68	0,618
	SI	34	0,309
DISNEA	NO	76	0,691
	SI	41	0,373
CEFALEA	NO	69	0,627
	SI	16	0,145
MIALGIA	NO	94	0,855
	SI	3	0,027
DOLOR TORACICO	NO	107	0,973
	SI	30	0,273
ANOMALÍAS EN LA AUSCULTACIÓN PULMONAR	NO	80	0,727
	SI	73	0,664
	NO	37	0,336

Fuente: Hospital de División Regional de la Seguridad Zona Militar de Guayaquil-Ecuador.
Elaborado por: María Luisa Contreras Yagala

Tal como se muestra en la tabla V del total de pacientes el 96,4% reportó padecer de tos. Un 73,6 tuvo expectoración en su sintomatología, el tipo de expectoración con más frecuencia fue mucopurulenta con una 51,8%. El 30,9% de los pacientes tuvo hemoptisis. En lo que concierne a la auscultación pulmonar el 66,4% de los pacientes registró al menos un sonido anómalo en la exploración