



**ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**

**FACULTAD DE CIENCIAS**

**CARRERA DE INGENIERÍA EN ESTADÍSTICA INFORMÁTICA**

**“ANÁLISIS ESTADÍSTICO IMPLICATIVO PARA LA  
IDENTIFICACIÓN DE FACTORES PRONÓSTICOS EN LA  
MORTALIDAD POR CÁNCER DE MAMA, SOLCA 2017-2018”**

**Trabajo de titulación**

Tipo: Proyecto de Investigación

Presentado para optar el grado académico de:

**INGENIERO EN ESTADÍSTICA INFORMÁTICA**

**AUTOR: WILLIAM GEOVANNY QUINATO MARTÍNEZ**

**DIRECTORA: ING. JOHANNA ENITH AGUILAR REYES**

Riobamba - Ecuador

2020

© 2020, William Geovanny Quinatoa Martínez

Se autoriza la reproducción total y parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el Derecho de Autor.

Yo, WILLIAM GEOVANNY QUINATO A MARTÍNEZ, declaro que el presente trabajo de titulación es de mi autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor asumo la responsabilidad legal y académica de los contenidos de este trabajo de titulación. El patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 15 de agosto de 2020

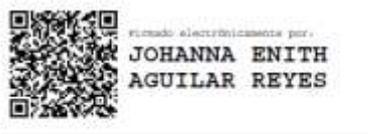
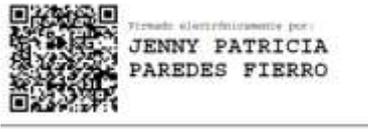


**William Geovanny Quinatoa Martínez**

**060435465-4**

**ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**  
**FACULTAD DE CIENCIAS**  
**CARRERA DE INGENIERÍA EN ESTADÍSTICA INFORMÁTICA**

El Tribunal del trabajo de titulación certifica que: El trabajo de investigación: “ANÁLISIS ESTADÍSTICO IMPLICATIVO PARA LA IDENTIFICACIÓN DE FACTORES PRONÓSTICOS EN LA MORTALIDAD POR CÁNCER DE MAMA, SOLCA 2017-2018”, realizado por el señor **WILLIAM GEOVANNY QUINATO MARTÍNEZ**, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Titulación, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal Autoriza su presentación.

	<b>FIRMA</b>	<b>FECHA</b>
<p>Dra. Jaqueline Elizabeth Balseca Castro</p> <p><b>PRESIDENTE DEL TRIBUNAL</b></p>		2020-08-28
<p>Ing. Johanna Enith Aguilar Reyes</p> <p><b>DIRECTORA DEL TRABAJO DE TITULACIÓN</b></p>		2020-08-28
<p>Ing. Jenny Patricia Paredes Fierro</p> <p><b>MIEMBRO DEL TRIBUNAL</b></p>		2020-08-28

## **DEDICATORIA**

A mis padres, Gabriel y Mariana por su amor, trabajo y sacrificio en todos estos años, gracias a ustedes he logrado llegar hasta aquí y cumplir una de las metas que me he propuesto, es un orgullo y privilegio ser su hijo, son los mejores padres, además a todas las personas que me han apoyado y han hecho que este trabajo se realice con éxito, en especial a aquellos que compartieron sus conocimientos.

William

## **AGRADECIMIENTO**

Primero agradecer a Dios por todas sus bendiciones, a mis padres que me han dado su ejemplo de trabajo, honradez y a nunca darme por vencido a lo largo de mi vida.

A la Escuela Superior Politécnica de Chimborazo, por acogerme en sus aulas y darme la oportunidad de prepararme profesionalmente.

De manera especial a mi tutora de tesis, Ing. Johanna Aguilar por la confianza y paciencia puesta en mí durante la realización de este proyecto de investigación, gracias por guiarme en esta meta propuesta.

A la Ing. Alexandra Marcatoma por haberme guiado, no solo en la elaboración de este trabajo de titulación, sino a lo largo de mi carrera universitaria y haberme brindado el apoyo para desarrollarme profesionalmente.

Agradezco a los todos docentes que, con su sabiduría, conocimiento y apoyo, motivaron a desarrollarme como persona y profesional, y que de una u otra manera ayudaron a mi formación como Ingeniero en Estadística Informática.

Al Hospital Oncológico SOLCA de la provincia de Chimborazo, en especial al Dr. Pablo Espinoza, por facilitar toda la información necesaria para llevar a cabo este trabajo de investigación.

A mi hija Cataleya, aunque llegara recién a mi vida, estoy seguro de que estará conmigo por todo lo que resta de ella.

Finalmente, de manera especial y de corazón deseo agradecer a mi familia por todos los consejos y apoyo en todo momento.

## TABLA DE CONTENIDO

ÍNDICE DE TABLAS.....	xi
ÍNDICE DE FIGURAS.....	xii
ÍNDICE DE GRÁFICOS.....	xiii
ÍNDICE DE ANEXOS .....	xiv
RESUMEN.....	xv
SUMMARY .....	xvi
INTRODUCCIÓN .....	1
<b>CAPÍTULO I</b>	
<b>1</b>	<b>MARCO TEÓRICO REFERENCIAL .....</b> 8
<b>1.1</b>	<b>Definiciones Básicas .....</b> 8
<b>1.2</b>	<b>Factores pronósticos.....</b> 9
<b>1.2.1</b>	<i>Edad.....</i> 10
<b>1.2.2</b>	<i>Paridad.....</i> 10
<b>1.2.3</b>	<i>Afectación a la axila.....</i> 11
<b>1.2.4</b>	<i>Rotura capsular y extensión extracapsular.....</i> 11
<b>1.2.5</b>	<i>Patrón histológico .....</i> 11
<b>1.2.6</b>	<i>Invasión vascular y linfática.....</i> 12
<b>1.2.7</b>	<i>Clasificación TNM.....</i> 12
<b>1.2.8</b>	<i>Tamaño tumoral.....</i> 14
<b>1.2.9</b>	<i>Localización del tumor.....</i> 14
<b>1.2.10</b>	<i>Obesidad .....</i> 14
<b>1.2.11</b>	<i>Hormonodependencia .....</i> 15
<b>1.2.12</b>	<i>Expresión Her-2/neu.....</i> 15
<b>1.2.13</b>	<i>Grado de diferenciación nuclear .....</i> 15
<b>1.2.14</b>	<i>Nectina 4.....</i> 16

<b>1.3</b>	<b>Tipos de cáncer de mama .....</b>	<b>16</b>
<b>1.3.1</b>	<b><i>CDIS: carcinoma ductal in situ .....</i></b>	<b>16</b>
<b>1.3.1.1</b>	<i>Indicios y Síntomas del carcinoma ductal in situ (CDIS) .....</i>	16
<b>1.3.1.2</b>	<i>Diagnóstico del CDIS .....</i>	17
<b>1.3.2</b>	<b><i>CDI: Carcinoma ductal invasivo .....</i></b>	<b>21</b>
<b>1.3.2.1</b>	<i>Indicios y síntomas del CDI .....</i>	22
<b>1.3.2.2</b>	<i>Diagnóstico del CDI .....</i>	22
<b>1.3.3</b>	<b><i>Carcinoma tubular de la mama .....</i></b>	<b>23</b>
<b>1.3.3.1</b>	<i>Síntomas y diagnósticos del carcinoma tubular de la mama .....</i>	23
<b>1.3.4</b>	<b><i>Carcinoma medular de la mama .....</i></b>	<b>24</b>
<b>1.3.4.1</b>	<i>Síntomas y diagnósticos del carcinoma medular de la mama .....</i>	24
<b>1.3.4.2</b>	<i>Tratamiento para el carcinoma medular de la mama .....</i>	25
<b>1.3.5</b>	<b><i>Carcinoma mucinoso de la mama .....</i></b>	<b>25</b>
<b>1.3.5.1</b>	<i>Síntomas y diagnósticos del carcinoma mucinoso de la mama .....</i>	26
<b>1.3.5.2</b>	<i>Tratamiento para el carcinoma mucinoso de la mama .....</i>	27
<b>1.3.6</b>	<b><i>Carcinoma papilar de la mama .....</i></b>	<b>27</b>
<b>1.3.7</b>	<b><i>Carcinoma cribiforme de la mama .....</i></b>	<b>27</b>
<b>1.3.8</b>	<b><i>CLI: Carcinoma lobular invasivo .....</i></b>	<b>27</b>
<b>1.3.8.1</b>	<i>Indicios y síntomas del CLI .....</i>	28
<b>1.3.9</b>	<b><i>CMI: Cáncer de la mama inflamatorio .....</i></b>	<b>28</b>
<b>1.3.9.1</b>	<i>Síntomas del CMI .....</i>	29
<b>1.3.9.2</b>	<i>Diagnóstico y determinación del estadio del CMI .....</i>	30
<b>1.3.10</b>	<b><i>CLIS: Carcinoma Lobular In situ .....</i></b>	<b>31</b>
<b>1.3.10.1</b>	<i>Síntomas y Diagnóstico del CLIS .....</i>	31
<b>1.3.11</b>	<b><i>Cáncer de mama en hombres .....</i></b>	<b>31</b>
<b>1.3.11.1</b>	<i>Factores de riesgo de cáncer de mama en hombres .....</i>	32
<b>1.3.11.2</b>	<i>Síntomas del cáncer de mama en hombres .....</i>	32
<b>1.3.12</b>	<b><i>Enfermedad de Paget en el pezón.....</i></b>	<b>33</b>

1.3.12.1	<i>Síntomas y diagnóstico</i> .....	33
1.3.12.2	<i>Tratamiento para la enfermedad de Paget</i> .....	33
1.3.13	<i>Tumor filoides de la mama</i> .....	34
1.3.14	<i>Cáncer de mama recurrente o metastásico</i> .....	34
1.4	<b>Análisis Estadístico Implicativo</b> .....	35
1.1.1	<i>Tipos de variables</i> .....	35
1.1.2	<i>Análisis clasificadorio</i> .....	36
1.1.3	<i>Árbol de similaridad</i> .....	36
1.1.4	<i>Definición de Análisis Estadístico Implicativo</i> .....	36
1.1.5	<i>Intensidad de la implicación entre variables</i> .....	37
1.4.6	<i>Análisis cohesitivo</i> .....	38
1.4.7	<i>Tipicalidad y contribución</i> .....	38
1.5	<b>Regresión Logística</b> .....	39
1.5.1	<i>Aspectos para el uso de la Regresión Logística</i> .....	40

## CAPÍTULO II

2	<b>MARCO METODOLÓGICO</b> .....	42
2.1	<b>Tipo y diseño de investigación</b> .....	42
2.2	<b>Localización del área de estudio</b> .....	42
2.3	<b>Población de estudio</b> .....	43
2.4	<b>Recolección de información</b> .....	43
2.5	<b>Identificación de variables</b> .....	43
2.6	<b>Operalización de variables</b> .....	44
2.7	<b>Alcances de investigación</b> .....	46
2.8	<b>Análisis de datos</b> .....	46

## CAPÍTULO III

3	<b>RESULTADOS Y DISCUSIÓN</b> .....	47
3.1	<b>Análisis exploratorio de datos cuantitativos</b> .....	47

<b>3.2</b>	<b>Análisis exploratorio de datos cualitativos .....</b>	<b>50</b>
<b>3.3</b>	<b>Asociación entre factores con la variable dependiente .....</b>	<b>55</b>
<b>3.4</b>	<b>Estimación del modelo de Regresión Logística .....</b>	<b>67</b>
<b>3.4.1</b>	<i>Estimación de parámetros.....</i>	<i>67</i>
<b>3.4.2</b>	<i>Validación del modelo.....</i>	<i>74</i>
<b>3.4.3</b>	<i>Evaluación del riesgo de pronóstico .....</i>	<i>74</i>
<b>3.5</b>	<b>Análisis Implicativo .....</b>	<b>75</b>
	<b>CONCLUSIONES.....</b>	<b>82</b>
	<b>RECOMENDACIONES.....</b>	<b>83</b>
	<b>GLOSARIO</b>	
	<b>BIBLIOGRAFÍA</b>	
	<b>ANEXOS</b>	

## ÍNDICE DE TABLAS

<b>Tabla 1-1:</b>	Clasificación del sistema TNM.....	13
<b>Tabla 1-2:</b>	Operalización de variables .....	44
<b>Tabla 1-3:</b>	Análisis descriptivo de los factores peso, talla y edad.....	49
<b>Tabla 2-3:</b>	Análisis de los factores pronóstico.....	49
<b>Tabla 3-3:</b>	Análisis de factores sociales en cáncer de mama.....	53
<b>Tabla 4-3:</b>	Prueba $X^2$ Estadio vs Localización de la mama.....	54
<b>Tabla 5-3:</b>	Prueba $X^2$ Estadio vs Birads .....	55
<b>Tabla 6-3:</b>	Prueba $X^2$ Estadio vs Reporte. ....	56
<b>Tabla 7-3:</b>	Prueba $X^2$ Estadio vs Cuadrante.....	57
<b>Tabla 8-3:</b>	Prueba $X^2$ Estadio vs Tamaño 1 .....	58
<b>Tabla 9-3:</b>	Prueba $X^2$ Estadio vs Tamaño 2 .....	58
<b>Tabla 10-3:</b>	Prueba $X^2$ Estadio vs Ganglios I.....	60
<b>Tabla 11-3:</b>	Prueba $X^2$ Estadio vs Ganglios II. ....	60
<b>Tabla 12-3:</b>	Prueba $X^2$ Estadio vs CA 15-3. ....	62
<b>Tabla 13-3:</b>	Prueba $X^2$ Estadio vs Ki 67. ....	63
<b>Tabla 14-3:</b>	Asociación de los factores de riesgo sociales.....	64
<b>Tabla 15-3:</b>	Pruebas no paramétricas en factores pronósticos .....	65
<b>Tabla 16-3:</b>	Estimación del modelo de predicción.....	66
<b>Tabla 17-3:</b>	Pruebas de la razón de verosimilitud.....	71
<b>Tabla 18-3:</b>	Validación del modelo.....	72
<b>Tabla 19-3:</b>	Pseudo R cuadrado .....	72
<b>Tabla 20-3:</b>	Clasificación con el modelo de regresión estimado. ....	73
<b>Tabla 21-3:</b>	Clasificación del árbol jerárquico.....	75

## ÍNDICE DE FIGURAS

<b>Figura 1-1:</b>	Resultados de células en caso de cáncer invasivo.....	18
<b>Figura 2-1:</b>	CDIS Papilar.....	19
<b>Figura 3-1:</b>	CDIS Cribiforme.....	20
<b>Figura 4-1:</b>	CDIS Solido.....	20
<b>Figura 5-1:</b>	CDIS de grado alto o comedo.....	21
<b>Figura 6-1:</b>	Curva de regresión logística.....	39

## ÍNDICE DE GRÁFICOS

<b>Gráfico 1-2:</b>	Ubicación de la provincia de Chimborazo y su cantón Riobamba .....	43
<b>Gráfico 1-3:</b>	Diagrama de caja de peso, talla, edad y tamaño del tumor dimensión 1. ....	47
<b>Gráfico 2-3:</b>	Diagrama de caja del tumor dim2, ganglios I, II y Ca 15-3.....	48
<b>Gráfico 3-3:</b>	Diagrama de caja de Ki 67 (%)......	49
<b>Gráfico 4-3:</b>	Representación de la localización de la mama. ....	50
<b>Gráfico 5-3:</b>	Representación del estadio.....	51
<b>Gráfico 6-3:</b>	Representación del reporte.....	51
<b>Gráfico 7-3:</b>	Representación de Birads.....	52
<b>Gráfico 8-3:</b>	Representación del cuadrante .....	53
<b>Gráfico 9-3:</b>	Representación del tamaño del tumor 1 y 2.....	53
<b>Gráfico 10-3:</b>	Representación de ganglios I y II, CA 15-3 y Ki 67.....	54
<b>Gráfico 11-3:</b>	Asociación entre Estadio y localización de la mama. ....	56
<b>Gráfico 12-3:</b>	Asociación entre Estadio y Birads. ....	57
<b>Gráfico 13-3:</b>	Asociación entre Estadio y reporte .....	58
<b>Gráfico 14-3:</b>	Asociación entre Estadio y cuadrante. ....	59
<b>Gráfico 15-3:</b>	Asociación entre Estadio y dimensión del tumor 1 .....	61
<b>Gráfico 16-3:</b>	Asociación entre estadio y dimensión del tumor 2. ....	61
<b>Gráfico 17-3:</b>	Asociación entre estadio y ganglios aislados nivel I.....	63
<b>Gráfico 18-3:</b>	Asociación entre Estadio y ganglios aislados nivel II.....	63
<b>Gráfico 19-3:</b>	Asociación entre Estadio y CA 15-3. ....	64
<b>Gráfico 20-3:</b>	Asociación entre Estadio y Ki 67 (%) .....	65
<b>Gráfico 21-3:</b>	Árbol de similaridad en R-CHIC.....	76
<b>Gráfico 22-3:</b>	Árbol jerárquico en R-CHIC. ....	79
<b>Gráfico 23-3:</b>	Gráfico implicativo en R-CHIC. ....	80

## **ÍNDICE DE ANEXOS**

**ANEXO A: CÓDIGO PARA DIAGRAMAS DE CAJA**

**ANEXO B: CÓDIGO EN R PARA EL ANÁLISIS IMPLICATIVO**

## RESUMEN

El presente trabajo de investigación tuvo por objeto analizar estadísticamente los factores pronósticos de la mortalidad de cáncer de mama registrados en la Sociedad de Lucha contra el Cáncer de la ciudad de Riobamba, periodo 2017-2018, se estudió las variables registradas en el sistema de dicha entidad las cuales fueron resultados de las pruebas realizadas por los pacientes que sufrían la afección de cáncer de mama, en el estudio se utilizó técnicas como: Análisis implicativo, Regresión Logística (RL) y Tablas de Contingencia con el apoyo de softwares como R versión 3.5.2, SPSS 24 y la hoja de cálculo de Excel 2018. Mediante el análisis estadístico descriptivo univariante, se determinó que los pacientes de cáncer de mama mayormente vulnerables son de sexo femenino 99%, casadas y de etnia mestiza. Se demostró también mediante técnicas no paramétricas que las variables nivel de instrucción, estado civil y ocupación no se asocian a la mortalidad de cáncer de mama dando a entender que no tienen significancia estadística al momento de contraer esta enfermedad. Su modelo de regresión logística logro alcanzar un r cuadrado de 0.73 mediante el coeficiente de Cox y Snell y asociando de buena manera a la clasificación de esta enfermedad por estadios con las variables dimensión del tumor, localización del cáncer, reporte de la enfermedad y marcador tumoral CA 15.3.

**Palabras clave:** <ESTADÍSTICA>, <REGRESIÓN LOGÍSTICA>, <ANÁLISIS IMPLICATIVO>, <TABLAS DE CONTINGENCIA>, <FACTORES PRONÓSTICOS>, <CÁNCER DE MAMA>, <RIOBAMBA (CANTÓN)>

## SUMMARY

The current research work aimed to statistically analyze the prognostic factors of breast cancer mortality registered by the Society for the Fight Against Cancer of Riobamba, period 2017-2018, the variables registered in the system of mentioned entity which were results of the tests carried out on patients suffering from breast cancer were studied, the study used techniques such as: Implicative Analysis, Logistic Regression (RL) and Contingency Tables with the support of software such as R version 3.5.2, SPSS 24 and Microsoft Excel 2018 spreadsheet. Through univariate descriptive statistical analysis, it was determined that the most vulnerable breast cancer patients are 99% female, married and of mixed ethnicity. It was also demonstrated by non-parametric techniques that the variables level of education, marital status and occupation are not associated with mortality from breast cancer, suggesting that they have no statistical significance at the time of contracting this disease. Its logistic regression model achieved a squared r of 0.73 using the Cox and Snell coefficient and associating in a good way the classification of this disease by stages with the variables dimension of the tumor, location of the cancer, report of the disease and tumor marker CA 15.3.

**Keywords:** <STATISTICS>, <LOGISTICS REGRESSION>, <IMPLICATIVE ANALYSIS>, <CONTINGENCY TABLES>, <FORECAST FACTORS>, <BREAST CANCER>, <RIOBAMBA (CANTON)>

## INTRODUCCIÓN

El cáncer de mama es el proceso tumoral maligno con más frecuente en las mujeres y que sigue constituyendo una verdadera amenaza para la salud en el mundo occidental, además no es de extrañarse, pues es la segunda causa de muerte por cáncer en la mujer y la primera en una franja de edad entre los 40 y los 44 años (González et al., 2019, p. 130-138).

Para concebir un programa terapéutico apropiado en el cáncer de mama se debe identificar un conjunto de factores pronósticos, hacer una valoración individual de cada paciente y conocer como evoluciona cada uno de ellos ya que existen diversos tipos de patologías en la mama. Los factores pronósticos se deben diferenciar bien de los factores predictivos. Un factor pronóstico es cualquier medida utilizable en el momento de la cirugía que se correlaciona con el intervalo libre de enfermedad o la supervivencia global en ausencia de un método o tratamiento adyuvante sistémico y como resultado es capaz de correlacionarse con los antecedentes del padecimiento, en contraste, un factor predictivo es cualquier cálculo asociada con la respuesta a un tratamiento dado con referencia al cáncer de mama (González et al., 2019, p. 130-138).

La presente investigación tiene por objetivo determinar los principales factores pronósticos en los pacientes que han atendidos en el hospital oncológico SOLCA e ilustrar la importancia de conocer estos factores y su influencia en la enfermedad ya que a través de ello se podrá conocer un diagnóstico anticipado reduciendo la mortalidad por esta patología especialmente en el colectivo femenino y junto con ello la creación de una cultura de prevención en la sociedad.

Hasta la fecha, el diagnóstico precoz es la principal arma contra esta enfermedad mortal, pero los conocimientos actuales sobre esta enfermedad y las nuevas técnicas histopatológicas han permitido definir diversos patrones de pronóstico en el cáncer de mama, entre los factores pronóstico más importantes se encuentran la edad, el tamaño del tumor, el tiempo de evolución, la paridad, la obesidad, la localización, el tratamiento de reemplazo hormonal, la presencia de ganglios axilares clínicamente y la clasificación tumoral, por lo que las técnicas estadísticas que se utilizan para el estudio de los factores pronósticos son las técnicas multivariadas, regresión logística y el análisis implicativo (Rivera et al., 2019, p. 308-322).

En el capítulo II se describe el marco metodológico, en él se define la zona de estudio que es el hospital oncológico SOLCA ubicado en la provincia Chimborazo, el diseño de investigación, población de estudio, operacionalización de la variable, y el análisis de datos.

En el capítulo III se muestra los resultados obtenidos del análisis descriptivo de las variables cuantitativas y mudables, no se realizó un relleno en la base de datos ya que solo se utilizó a los pacientes que poseían estudios completos en el diagnóstico del cáncer de mama. Se realiza una comparación entre los factores pronósticos mediante la prueba no paramétrica Chi cuadrado de manera univariante ya que esta permite si tienen significancia estadística para un valor  $p < 0.05$ . La obtención del modelo se lo realizo con un modelo de regresión logística utilizando todos los factores pronósticos obtenidos en la recolección de datos, su validación se dio mediante el coeficiente de Cox y Snell. El análisis implicativo se lo realizado en R-Studio con el paquete R-CHIC, este paquete permite realizar el árbol de similaridad, árbol jerárquico y el gráfico implicativo, asociando así las variables de buena manera con la mortalidad de cáncer de mama.

## Antecedentes

En el transcurso de la humanidad esta ha tenido relación con las enfermedades oncológicas durante todo su paso; así como en el estudio de restos en el Egipto antiguo de la necrópolis situada cerca de Gizeh, demostró que hace muchos años, aproximadamente unos 2700 a.C se identificaron tumores óseos. Demostrando que los tumores son mencionados en la medicina egipcia, en el Papiro de Ebers (1500 a.C) y en el Papiro de Edwin Smith, el antecedente más antiguo conocido en cirugía. Etimológicamente el término “cangrejo” (*karkinos*) fue aplicado a los tumores malignos por los antiguos griegos, adoptado por los romanos y aceptado en la actual nomenclatura de los tumores hasta el momento (Lazcano et al., 1996, p. 140).

En el Ecuador, la incidencia de Cáncer de mama según estadísticas de Globocan 2018, hubo 28.058 casos nuevos de cáncer, en promedio existen 165 casos de cáncer en todas sus variedades por cada 100.000 mujeres y 150 casos por cada 100.000 hombres. Los cánceres en hombre más comunes según su incidencia son: próstata 3322 (26%), estómago 1364 (10%), colorrectal 902 (7,1%), linfoma 770 (6%) y leucemia 655 (5,1%). Mujeres: los cánceres más comunes según su incidencia son: mama 2787 (18,2%), cuello uterino 1612 (10,6%), tiroides 1374 (9%), estómago 1225 (8%) y colorrectal 1123 (7,4%) (MSP, 2018, p. 1).

Esta enfermedad oncológica es una enfermedad que desarrolla células malignas en los tejidos de la mama y si esta es agresiva se expande a otras partes del cuerpo. La glándula mamaria se compone por lóbulos y lobulillos unidos mediante conductos, y esta enfermedad afecta a una o ambas de las estructuras mencionadas, pero esta enfermedad se vuelve maligna con el pasar del tiempo (Silva, 2018, p. 20).

En algunos casos, tener antecedentes familiares de cáncer de mama está ligado a tener un gen anormal asociado con un riesgo elevado de padecer esta enfermedad, como el gen BRCA1 o BRCA2. En otros casos, una anomalía del gen CHEK2 puede desempeñar un papel en el desarrollo de la patología. Las mujeres con parientes cercanos que han sido diagnosticados con cáncer de mama tienen un riesgo mayor de desarrollar la enfermedad, el riesgo se duplica si el familiar es de primer grado tales como la madre, hermana o hija; además si dos familiares de primer grado han sido diagnosticados el riesgo es 5 veces superior. Un elemento adicional aparece cuando se conoce si el padre o hermano fueron diagnosticados con esta enfermedad ya que el riesgo del individuo en proceso de diagnóstico es mayor, aunque los investigadores no han aclarado cuánto más alto.

El riesgo más importante para sufrir este tipo de cáncer es ser mujer, la relación de este tipo de

enfermedad entre mujeres y hombres es aproximadamente de 100 a 1. La incidencia de cáncer de mama se incrementa con la edad, duplicándose aproximadamente cada 10 años hasta la menopausia, etapa en que el ritmo de crecimiento disminuye.

En la revista Ciencias Médicas de Santiago de Cuba, en la publicación denominada “Análisis estadístico implicativo para la identificación de factores pronósticos de la mortalidad por cáncer de mama”, hace referencia a un estudio observacional y analítico en el cual las técnicas que se utilizaron fueron técnicas multivariadas como la regresión logística o regresión de Cox, en el cual el análisis estadístico implicativo resultó ser una técnica apropiada que complementa a la regresión logística en la identificación de factores pronósticos, con lo cual se logró una mejor comprensión de la causalidad e interpretación de estos factores (Moraga et al., 2017, p. 1).

En la investigación denominada “Modelo predictivo de regresión logística en función a los factores de riesgo para contraer cáncer de mama”, cuyo centro de investigación fue en Perú, se analizó los registros de cáncer de Lima Metropolitana, entre los años 1990-1993, el cual dio como resultado que en mujeres el cáncer de mama ocupa el primer lugar con una tasa de incidencia de 24.28x10000 habitantes, seguido del cáncer cérvix con una incidencia de 20.24x100000 habitantes. Otro estudio realizado en 1998, se analizaron 1574 casos consecutivos de cáncer de mama atendidos en el INEM entre los años 1990-1992, encontrándose que el 48.28% tenían cáncer inoperable (estadios III y IV) (Silva, 2018, p. 11).

En la ciudad de México se evaluaron los factores pronósticos junto con la ocurrencia de metástasis cerebral como primer sitio de recurrencia en pacientes con cáncer de mama y metástasis al sistema nervioso central (SNC), en el estudio mencionado la muestra estaba formada por 125 pacientes con CM atendidas en el Instituto Nacional de Cancerología durante 2017-2015, se utilizó el modelo de Cox el cual fue utilizado para determinar los factores pronósticos para el cáncer de mama, el cual dio como resultado una mediana global de supervivencia de 14.2 meses (IC 95% 11.8-26.9), los pacientes clasificados por inmunohistoquímica como triple negativo presentaron tiempos de supervivencia más cortos con un  $p < 0.004$  y con un doble más de riesgo de no sobrevivir, en comparación con otros inmunofenotipos (HR=2.76; IC95% 1.17-6.51), llegando a concluir que la metástasis al sistema nervioso central continua siendo un factor de mal pronóstico que reduce la supervivencia y afecta la calidad de vida de las personas que tienen esta enfermedad (Mohar et al., 2018, p. 1-2).

El investigador Gonzalo Arturo Medina en su estudio realizado en Arequipa Perú, demostró que las características clínicas y pronósticas de los carcinomas de mama cuya base de datos fueron 280 mujeres con cáncer infiltrante de mama unilateral, del 2009-2012, el cual clasifico por

técnicas de inmunohistoquímica de 4 subtipos: luminal A, luminal B, HER2 y triple negativo el cual para determinar la influencia del tipo histológica y del subtipo molecular en la sobrevida global, se utilizó el método de Kaplan Meier y encontró que los carcinomas de mama fueron: luminal A con 105 (37.5%); para luminal B con 88 (31.4%); en el carcinoma HER2 con 46 (16.14%), y triple negativo con 41 (14.6%), llegando a concluir que los carcinomas de tipo luminal fueron con mayor frecuencia tumores bien diferenciados, con ganglios axilares negativos, tamaño tumoral y estadio semejante; mientras que los tumores HER2 y triple negativo presentaron mayor proporción de tumores pobremente diferenciados comprometiendo el ganglionar axilar, y obteniendo una menor supervivencia global (Medina , 2017, p. 1).

En la investigación “Early detection of breast cancer in Brazil: data from the National Health Survey, 2013”, se analizó las pacientes para la detección temprana del cáncer de mama, dando como resultado en las mujeres, el 66.7% tenía una solicitud médica para mamografía (59.4% usuarios del SUS). Los factores como tener un seguro de salud privada, un nivel de educación superior y ser de etnia blanca se asociaron positivamente con tener una solicitud médica. Solo el 5.4% (IC 95% 4.8-6.0) de las mujeres que recibieron una solicitud medica no se pudieron realizar mamografías: el 7.6% eran beneficiarias del SUS y el 1.7% tenían seguro médico. Las razones para no poder realizar el examen fueron que no pensaron que era necesario hacerlo o que tenían una prueba proyectada, pero aún no se la realizaba. El 70% de las mujeres recibió el resultado con menos de un mes desde su realización (Azevedo , 2016, p. 1).

### **Planteamiento del problema**

El cáncer de mama es la enfermedad más habitual en la población de mujeres, ya sea en los países desarrollados como en los países en vías de desarrollo; no obstante, el sexo masculino también puede adquirir esta enfermedad pero en menores proporciones, ya que también pueden desarrollar estas patologías, esto se debe al aumento de la urbanización y la adopción de estilos de vida no convenientes en las personas, A pesar de los avances en la detección y de los métodos de tratamiento existentes de esta enfermedad, su pronóstico sigue dependiendo principalmente de la extensión de la patología en el momento que se la descubre, por lo que un diagnóstico en una etapa temprana continúa siendo la mejor vía para mejorar las posibilidades de curación en las personas que sufren cáncer de mama de cualquier tipo.

En el Ecuador las estadísticas dictan que 30 de cada 100.000 personas son diagnosticadas con esta patología cada año y casi la mitad de ellas en etapas muy avanzadas, reduciendo las posibilidades de curación y supervivencia de las personas, durante el año 2011, se realizaron en el Ministerio de Salud Pública 448.350 consultas de detección oportuna, aun así, este padecimiento representa

la tercera causa de muerte por cáncer en las mujeres ecuatorianas.

¿Cuáles son los factores pronósticos de la mortalidad por cáncer de mama en SOLCA 2017-2018?

### **Justificación**

El análisis estadístico implicativo es una poderosa herramienta de investigación que permite determinar cuasi-implicaciones entre variables y clases de variables, este procedimiento se desenvuelve en correspondencia con los problemas referente al cáncer de mama, su principal objetivo contempla la estructuración de datos, interrelacionados a sujetos y variables, y la extracción de reglas inductivas entre las mismas (Naranjo, 2018, p. 17). Además, permite conocer posibles relaciones de similitud, implicación y cohesión entre los factores pronósticos que estén relacionados con el cáncer de mama; mientras que los modelos de regresión logística aplicados a este problema de salud permiten el estudio de los resultados en términos explicativos y predictivos, ayudando a descubrir la fuerza de asociación mediante las razones de ventaja (ODDS ) de los factores de riesgo con el efecto estudiado de una manera autónoma para lograr conocer el valor predictivo de cada uno de estos (NEFROLOGIA, 2000, p. 477-565).

Las herramientas disponibles en la actualidad para la prevención y el diagnóstico precoz de cáncer de mama son especialmente importantes para las mujeres quienes presentan mayor riesgo de desarrollar cáncer de mama y para quienes el inicio temprano del tratamiento se traduce en una reducción de la mortalidad, la identificación de factores pronósticos no solo contribuye a la detección temprana de la patología sino también a la posibilidad de desarrollar nuevos tratamientos que actúen en función del estilo de vida y características epidemiológicas halladas en los pacientes.

Las herramientas estadísticas que se ocupan en el campo de la medicina son métodos y procedimientos que ayudan a recoger, clasificar, resumir y analizar los datos para la realización de inferencias a partir de ellos, además son necesarios ya que su aplicación tiene como finalidad ayudar a tomar decisiones y, en su caso, formular predicciones. Al respecto se utilizan las técnicas estadísticas multivariadas, como el análisis implicativo y la regresión logística.

## **Objetivos**

### **Objetivo General**

Determinar los factores pronósticos asociados a la mortalidad por cáncer de mama en SOLCA 2017-2018.

### **Objetivos Específicos**

- Establecer los factores de riesgo sociales en la mortalidad de cáncer de mama.
- Establecer los factores de riesgo maternos culturales en la mortalidad de cáncer de mama.
- Estandarizar las tasas brutas de mortalidad según el sexo de los pacientes.
- Comparar los factores pronósticos de la mortalidad por cáncer de mama a través de técnicas no paramétricas.

## CAPÍTULO I

### 1 MARCO TEÓRICO REFERENCIAL

#### 1.1 Definiciones básicas

**Cáncer:** El término “cáncer” es genérico y designa un amplio grupo de enfermedades que pueden afectar a cualquier parte del cuerpo. Esta enfermedad se caracteriza por tener un crecimiento tisular producido por la proliferación continua de células anormales con capacidad de invasión y destrucción de otros tejidos del ser afectado. Esta puede originarse a partir de cualquier tipo de célula en cualquier tejido corporal, no es una enfermedad única, sino un conjunto de enfermedades que se clasifican en función del tejido y de la célula de origen con capacidad de expandirse si no es tratada a su debido tiempo. (Gandur. et al., 2009, p. 7).

**Cáncer de mama:** El cáncer de mama es el crecimiento desordenado de células malignas del tejido mamario, existen 2 tipos frecuentes de cáncer de mama: el carcinoma ductal y el carcinoma lobulillar (Peña. et al 2017, p. 284).

**Factor pronóstico:** Un factor pronóstico es cualquier medición utilizable en el momento de la cirugía que correlaciona con el intervalo libre de enfermedad o supervivencia global en ausencia de tratamiento y como resultado es capaz de correlacionar con la historia natural de la enfermedad (Gonzáles. et al., 2011, p. 130).

**Antecedentes familiares:** Registro de las relaciones entre los miembros de una familia junto con sus antecedentes médicos, esto abarca las enfermedades actuales y pasadas de todos sus familiares ya seas tíos abuelos, etc. En estos antecedentes a veces se observa la distribución de ciertas enfermedades en una familia. También se llama historia médica familiar (Diccionario del Instituto Nacional del Cáncer, 2019).

**Historia clínica:** Es la recopilación de la información médica de un paciente, una historia clínica digital incluye información sobre los antecedentes de salud de un paciente ya sea este como diagnósticos, medicamentos, pruebas, alergias, vacunaciones y planes de tratamiento. Facilitando a los proveedores de atención de la salud a cargo de un paciente que pueden ver y usar las historias clínicas digitales o físicas para ayudarlos a dar recomendaciones acerca de la atención del paciente (Diccionario del Instituto Nacional del Cáncer, 2019).

**Factor de Riesgo:** Un factor de riesgo es cualquier rasgo, particularidad o exposición de un

individuo que aumente su probabilidad de sufrir una enfermedad o lesión. Entre los elementos particulares de los factores de riesgo más importantes cabe citar la insuficiencia ponderal, las prácticas sexuales de riesgo, la hipertensión, el consumo de tabaco y alcohol, el agua insalubre, las deficiencias del saneamiento, falta de saneamiento y la falta de higiene (OMS, 2019).

**Factor de protección:** Se refiere a los elementos que pueden disminuir la probabilidad de contraer cierta enfermedad, algunos ejemplos de factores de protección del cáncer son realizar actividad física, mantener un peso saludable y seguir un régimen de alimentación sano (Diccionario del Instituto Nacional del Cáncer, 2019).

**Regresión logística:** La regresión logística es una técnica adecuada cuando la variable de respuesta Y es politómica ya que admite varias categorías de respuesta, tales como mejora mucho, empeora, se mantiene, mejora, mejora mucho, pero es especialmente útil en particular cuando solo hay dos posibles respuestas es decir cuando la variable de respuesta es dicotómica, que es el caso más común. La RL es una de las técnicas estadístico-inferenciales más empleadas en la producción científica de hoy en día y surge en la década de los 60 (De la Fuente, 2011, pp. 1).

**Variable cualitativa en el modelo logístico:** Como la metodología empleada para la estimación del modelo logístico se basa en la utilización de variables mudables, al igual que en cualquier otro procedimiento de predicción, es incorrecto que en él intervengan variables cualitativas, ya sean nominales u ordinales. La solución a este problema es crear tantas variables dicotómicas como número de respuestas, estas nuevas variables, artificialmente creadas, reciben el nombre de dummy, traducándose con diferentes designaciones como pueden ser variables internas, indicadoras, o variables diseño. Si una variable acopia datos del tabaco con las respuestas (Nunca fumó, Exfumador, fuma 20 o más cigarrillos al día), hay 4 posibles respuestas por lo que se construyen  $(4-1=3)$  variables dummy dicotómicas (valores 0, 1), existiendo diferentes posibilidades de clasificación, que conducen a interpretaciones diferentes en un trabajo de investigación (De la Fuente, 2011, pp. 3).

## 1.2 Factores pronósticos

Los factores pronósticos son datos que se obtienen de los múltiples estudios clínicos realizados sobre una enfermedad determinada, en este caso sobre el cáncer de mama. En estos estudios se obtienen variables llamadas factores pronóstico que el médico consigue a través de la sintomatología y de las pruebas clínicas realizadas a los pacientes, según estos factores pronóstico se puede establecer un tipo de procedimiento para lograr su recuperación (Martínez, 2007).

Entre los factores pronósticos más importantes del cáncer mamario son:

- Edad
- Paridad
- Afectación de la axila
- Rotura capsular y la extensión extracapsular
- Patrón histológico
- Invasión vascular y linfática
- Clasificación TNM
- Tamaño tumoral
- Localización del tumor
- Obesidad,
- Hormonodependencia
- Expresión de *HER-2/neu*
- Nectina 4

### **1.2.1 *Edad***

Es el proceso de envejecimiento con mayor factor de riesgo para el cáncer de mama, esto se debe a que, como vivimos más, hay más oportunidades para el daño genético (mutaciones) en el organismo. Y a medida que envejecemos, nuestros cuerpos son menos capaces de remediar el daño genético (González, 2011, pp.132-133).

Como en muchas enfermedades, el riesgo de contraer esta enfermedad aumenta con la edad, según la Sociedad Americana contra el Cáncer, aproximadamente 1 de cada 8 tipos de cáncer de mama invasivos se desenvuelven en mujeres menores de 45 años y aproximadamente 2 de 3 cánceres de mama invasivos se encuentran en mujeres de 55 años o mayores.

Según lo reportado por Van der Leest de Holanda, el riesgo de recurrencia local en las pacientes menores de 40 años sometidas a tratamiento conservador de la mama se reduce en más de un 50 % cuando se utiliza un método sistémico adyuvante (González, 2011, pp.132-133).

### **1.2.2 *Paridad***

Se refiere al número total de gestaciones que ha tenido una mujer, incluyendo abortos, también se pueden manejar con las siglas TPAL del inglés: term, preterm, abortion, living. (nacidos a

término, nacidos prematuros, abortos, vivos). Un modelo, TPAL: 2,0,1,2 quiere decir que la mujer ha tenido dos embarazos a término, ningún nacido prematuro, 1 aborto y tiene 2 hijos vivos.

Según *Anderson* y otros en un análisis multivariado manifestaron que, en los pacientes mayores de 60 años con cáncer de mama, la paridad fue el predictor independiente con más significancia de disminución de las metástasis a distancia y del progreso de la sobrevida global (González, 2011, p.133).

### **1.2.3 Afectación a la axila**

En la actualidad el factor pronóstico más significativo relacionado con el control local, la supervivencia libre de enfermedad y la sobrevida global lo constituye la metástasis a ganglios linfáticos axilares (González, 2011, p.133).

Los ganglios linfáticos son órganos pequeños en forma de frijol que actúan como coladores a través de los canales de líquido linfático, debido a que el líquido linfático sale de la mama y luego retoma el torrente sanguíneo, los ganglios linfáticos sirven para atrapar e inmovilizar las células cancerígenas antes de que lleguen a otras partes del organismo (González, 2011, p.133).

Existen 3 expresiones diferentes de ganglios linfáticos axilares positivos:

- El número de ganglios linfáticos axilares positivos 1 a 3 ganglios, de 4 a 9 ganglios y de 10 o más.
- La afectación ganglionar en los diferentes niveles de Berg (I II y III nivel).
- La proporción de ganglios linfáticos positivos.

La presencia de células cancerígenas en los ganglios linfáticos de la axila indica un riesgo más alto para la propagación de la enfermedad.

### **1.2.4 Rotura capsular y extensión extracapsular**

En la Escuela de Medicina de Graz, se analizó a 301 pacientes con metástasis ganglionar y extensión extracapsular donde se logró demostrar un índice de recaídas de un 9,3 %, la extensión de las metástasis a través de la cápsula de los ganglios afectados ocurre con mayor frecuencia en mujeres con 4 o más ganglios metastásicos (aproximadamente el 70 %) (González, 2011, p.133).

### **1.2.5 Patrón histológico**

El tipo histológico del cáncer de mama influye de manera fuerte en el pronóstico, estos cánceres de mama no invasivos representan aproximadamente el 10 % de todos los cánceres mamarios, tienen generalmente un mejor pronóstico (González, 2011, p.133).

El patrón histológico se relaciona estrechamente con la anatomía celular, pues su estudio no se detiene en los tejidos, sino que va más allá, observando también las células interiores y otros corpúsculos, relacionándose con la bioquímica y la citología para este tipo de padecimientos como lo es el cáncer de mama.

Los tipos histológicos con un mejor pronóstico, asociados con un índice de supervivencia global a los 5 años mayor del 85 % se encuentran el carcinoma tubular, el cribiforme, el coloide o mucinoso y el papilar, entre los carcinomas menos favorables se encuentran el carcinoma medular, cuyo pronóstico es intermedio entre el cáncer lobular invasivo y el carcinoma ductal infiltrante (González, 2011, p.134).

### **1.2.6 Invasión vascular y linfática**

La invasión del sistema vascular o del sistema linfático se origina cuando las células del cáncer de mama entran en los vasos sanguíneos o en los canales linfáticos aumentando el riesgo de que el cáncer se extienda afuera de la mama o regrese en el futuro.

La invasión vascular, linfática y perineural constituyen factores de mal pronóstico, mayor recurrencia y mortalidad, tanto para las pacientes con un proceso de cirugía conservadora de la mama, como para las pacientes tratadas con mastectomía (extirpación de la mama o parte de ella) (González, 2011, p.134).

No obstante, lo expresado con anterioridad, no podemos olvidar el enunciado de Goldman, hecha en 1897, de que no todas las células tumorales circulantes necesariamente originan metástasis (González, 2011, p.134).

### **1.2.7 Clasificación TNM**

El sistema TNM es el intento más aceptado para la codificación y estadificación de los tumores, el periodo clínico es un elemento que puede guiar la conducta terapéutica, además de ser un factor

pronóstico para la paciente con esta enfermedad, ya que no se comportan de igual manera los pacientes en estadios diferentes (González, 2011, p.134).

Los médicos combinan los resultados de T, N y M y otros factores específicos del cáncer para establecer el estadio del cáncer para cada persona, dando como resultado la siguiente clasificación:

- **Estadio 0:** Los cánceres de estadio 0 aún se ubican en el lugar en el que se iniciaron y no se han dispersado a los tejidos adyacentes, este estadio del cáncer a menudo es curable, generalmente extirpando el tumor entero mediante extirpación o cirugía.
- **Estadio I:** Este estadio generalmente es un cáncer o tumor pequeño que no ha crecido profundamente en los tejidos contiguos. Además, no se ha dispersado a los ganglios linfáticos ni a otras partes del cuerpo, a menudo se lo designa como cáncer en estadio temprano (American Society of Clinical Oncology, 2018).
- **Estadio II y Estadio III:** En general, estos 2 estadios indican cánceres o tumores más grandes que han crecido más profundo en los tejidos vecinos también es viable que se hayan dispersado a los ganglios linfáticos, pero no a otras partes del cuerpo (American Society of Clinical Oncology, 2018).
- **Estadio IV:** Este estadio significa que el cáncer se ha esparcido a otras partes del cuerpo, también se lo puede denominar cáncer avanzado o metastásico.

**Tabla 1-1:** Clasificación del sistema TNM

Estadificación (anatómica)	TNM
<b>0 (<i>in situ</i>)</b>	Tis N0 M0
<b>I (estadio local)</b>	
IA	T1 N0 M0
IB	T0 N1mi M0 T1 N1mi M0
<b>II (estadio regional temprano)</b>	
IIA	T0 N1 M0 T1 N1 M0 T2 N0 M0
IIB	T2 N1 M0

	T3 N0 M0
<b>III (estadio regional tardío)</b>	
IIIA	T0 N2 M0 T1 N2 M0 T2 N2 M0 T3 N1 M0 T3 N2 M0
IIIB	T4 N0-2 M0
IIIC	T1-4 N3 M0
<b>IV (estadio generalizado)</b>	T1-4 N1-3 M1

Fuente: Medicina Basada en Evidencia, 2020.

### 1.2.8 *Tamaño tumoral*

No es posible determinar el instante en que aparece la primera célula maligna, después de lo cual el tumor progresa por división celular y cuando la rapidez de desarrollo tumoral depende de la velocidad de duplicación. *Allgood* y otros encontraron en su trabajo de investigación que a mayor tamaño tumoral más frecuencia de invasión ganglionar axilar; con tumores menores de 3 cm se obtuvo diseminación axilar en un tercio de los casos, y con tumores de más de 3 cm aparecen metástasis axilares en más del 50% de los pacientes (González, 2011, p.134).

### 1.2.9 *Localización del tumor*

La localización del tumor en el cuadrante inferointerno (CII) está agrupada con el empeoramiento de la supervivencia en las mujeres con cáncer de mama en estadios tempranos además existe una evidencia cada vez más creciente que los tumores de los cuadrantes internos, estos metastizan con más frecuencia a la cadena mamaria interna (González, 2011, p.134).

Es muy importante estar al tanto la localización del tumor ya que este nos ayudara en la toma de decisiones terapéuticas.

### 1.2.10 *Obesidad*

El factor obesidad afecta de manera desfavorable a la salud de la mujer y del hombre, el riesgo de cáncer de mama en la mujer posmenopáusica se desarrolla con la obesidad, estableciendo que la obesidad está asociada con un pronóstico pobre de esta enfermedad oncológica (González, 2011, p.135).

Entre los posibles mecanismos por los cuales la obesidad causa un pronóstico malo del CM se hallan:

- La actividad estrogénica.
- Una enfermedad más avanzada o agresiva en el momento del diagnóstico.
- Elevada posibilidad de fracaso al tratamiento local y sistémico.

#### **1.2.11 *Hormonodependencia***

Los descubrimientos de Jensen y cols. en los años 70s, fueron la base científica de la hormonodependencia tumoral, después de estos estudios se llega a la conclusión de que la presencia de una proteína estrogénica, conocida como *receptor de estrógeno*, ayuda a mejorar cuantitativamente el pronóstico y hace que los CM respondan de una manera propicia a la terapia endocrina (González, 2011, p.135).

#### **1.2.12 *Expresión de HER-2/neu***

En el cáncer de mama el gen *HER-2/neu* está amplificado en el 20-30 % de los casos, el mecanismo por el cual la amplificación/sobreexpresión ocurre es aún desconocido. El gen del receptor 2 del factor de crecimiento epidérmico humano (*HER-2/neu* o *c-erb-B2*) es un protooncogen mapeado en el cromosoma 17. Se ha confirmado que los niveles elevados de *HER-2/neu* están asociados con un pronóstico clínico pobre y una supervivencia rebajada en las pacientes con cáncer de mama en estadios tempranos, y ello podría ayudar en el método de tratamiento para esta enfermedad (González, 2011, p.135).

#### **1.2.13 *Grado de diferenciación nuclear***

La modificación Nottingham del índice de Scarff-Bloom-Richardson incluye 3 clases:

- Tumores bien diferenciados (grado I).
- Tumores moderadamente diferenciados (grado II).
- Tumores pobremente diferenciados (grado III).

El cáncer de mama con un grado histológico alto (III) es altamente invasivo y metastásico. En lo que si están de acuerdo muchos científicos es en que el grado de diferenciación nuclear tendría correspondencia con la positividad de los estudios por citometría de flujo, el estado de

aneuploidía, la elevada fase S de división celular e incluso los valores altos de catepsina D, valores todos unidos empeorarían el pronóstico de la paciente (González, 2011, p.136).

#### **1.2.14 Nectina 4**

Es un marcador de progresión de la enfermedad y sus niveles se correlacionan con el número de metástasis, las nectinas son miembros de la super familia inmunoglobulina y son componentes de las E-caderin o distinguidas como moléculas de adhesión intercelular en células epiteliales (González, 2011, p.136).

Se han descrito hasta este momento 4 nectinas, de ellas, la nectina 4 es un nuevo antígeno asociado al tumor y un biomarcador utilizable para el CM. En el estudio de Fabre-Lafay y otros se manifestó que la nectina 4 no se expresa en células estándar y que dicho biomarcador es predominantemente expresado en el ovario y por arriba de todo en las líneas celulares del tumor mamario (carcinoma ductal mamario) (González, 2011, p.136).

### **1.3 Tipos de cáncer de mama**

El cáncer de mama puede comenzar en distintas áreas de la mama: los conductos, los lobulillos o, en algunos casos, el tejido intermedio, entre las diferentes clasificaciones para esta enfermedad estan incluidos los casos de cáncer de mama invasivo, no invasivo, recurrente y metastático, esta enfermedad también se puede encontrar en hombres, pero con muy poca frecuencia.

#### **1.3.1 CDIS: carcinoma ductal in situ**

El carcinoma ductal in situ (CDIS) es el tipo más frecuente de cáncer de mama no invasivo, ductal quiere decir que el cáncer comienza dentro de los conductos lácteos, la palabra carcinoma se refiere a cualquier cáncer que inicia en la piel u otros tejidos (incluyendo tejido mamario) que cubren o revisten los órganos internos, y la frase in situ significa “en su lugar original”. El CDIS se denomina “no invasivo” debido a que no trasciende fuera del conducto lácteo hacia otros tejidos mamarios circundantes normales.

El CDIS no pone en riesgo la vida, pero padecer CDIS puede incrementar el riesgo de desarrollar un cáncer de mama invasivo más adelante.

##### **1.3.1.1 Indicios y Síntomas del carcinoma ductal in situ (CDIS)**

El CDIS por lo regular no presenta indicios ni síntomas, una pequeña cantidad de pacientes pueden tener un bulto en la mama o secreción del pezón. Según el Instituto Nacional del Cáncer, aproximadamente el 80 % de casos de CDIS se descubren por mamografía.

### **1.3.1.2** *Diagnóstico del CDIS*

Para diagnosticar un CDIS se suele utilizar una combinación de procedimientos:

**Exploración física de las mamas:** Es posible que el médico palpe un pequeño bulto en la mama durante la exploración física, si bien la aparición de un bulto es rara en un CDIS, es posible que en algunos casos no sea posible palpar el CDIS en una exploración física, generalmente puede detectarse a través de una mamografía.

**Mamografía:** El CDIS por lo general se detecta con un examen mamográfico. A medida que las células cancerosas viejas sucumben y se acumulan, se forman pequeños almacenes de calcio (llamados “calcificaciones” o “microcalcificaciones”) dentro de las células descompuestas, la mamografía muestra las células cancerosas dentro de los conductos en forma de agrupación de estas microcalcificaciones, que aparecen como manchas blancas o tonos oscuros como sombras.

**Biopsia:** Si la mamografía muestra una rareza, es posible que el galeno indique una biopsia. Existen dos tipos de biopsia que se realizan mediante una cirugía menor (las biopsias más invasivas son poco frecuentes el caso CDIS):

**Biopsia de aspiración con aguja fina:** Para este proceso se inserta en la mama una aguja hueca muy pequeña y se extrae una muestra de las células y se las analiza con un microscopio, este procedimiento no deja cicatrices.

**Biopsia con aguja gruesa:** En este proceso se introduce una aguja de mayor tamaño para obtener varias muestras de tejido de la zona afectada, Para que la aguja gruesa atraviese la piel, el cirujano debe hacer un pequeño corte, cabe recalcar que esta técnica deja una pequeña cicatriz que apenas se puede percibir en el lapso de algunas semanas.

Si no es posible extraer células o tejidos mediante en una biopsia con aguja, o bien la biopsia no genera derivaciones exitosas, puede ser necesario aplicar una técnica de biopsia más compleja, dando la apariencia de una cirugía normal.

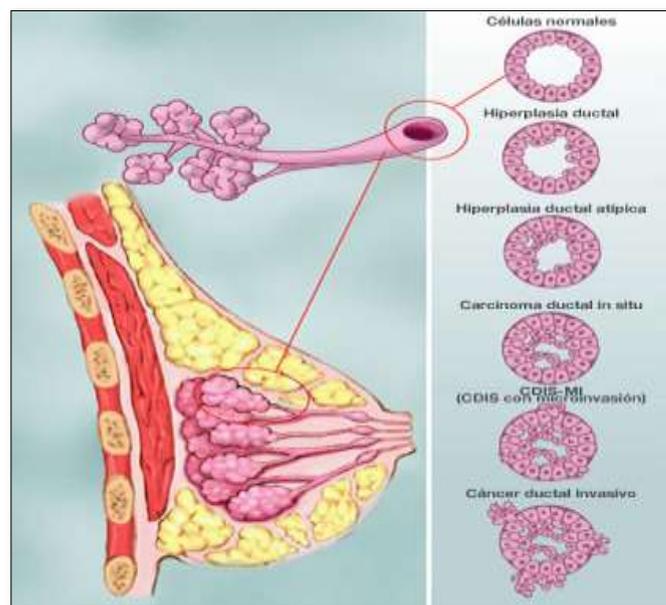
**Biopsia incisional:** Durante una biopsia incisional, se extrae una pequeña porción de tejido para examinarlo.

**Biopsia excisional:** en una biopsia excisional, se pretende extraer de la mama la totalidad del bulto con tejido presuntamente afectado.

Las biopsias se realizan únicamente para llegar a un análisis. Si este diagnóstica un CDIS, es necesario realizar una cirugía posterior para avalar la extirpación total del cáncer con “márgenes libres”, es decir, se elimina también un borde de tejido sano cerca del cáncer. Esto involucra someterse a una lumpectomía o, en algunos casos una mastectomía, que es la extirpación de la mama o parte de ella.

Todos los CDIS se analizan como cáncer de mama de estadio 0, el estadio más temprano que existe, el “estadio” indica en qué medida se ha extendido la enfermedad más allá de su ubicación original, aunque se dice que el CDIS pertenece siempre al estadio 0, este puede ser de cualquier tamaño y ubicarse en varias zonas dentro de la mama.

La imagen muestra el rango de resultados posibles, desde células normales hasta casos de cáncer ductal invasivo.



**Figura 1-1:** Resultados de células en casos de Cáncer invasivo

**Fuente:** Breast Cancer;2019.

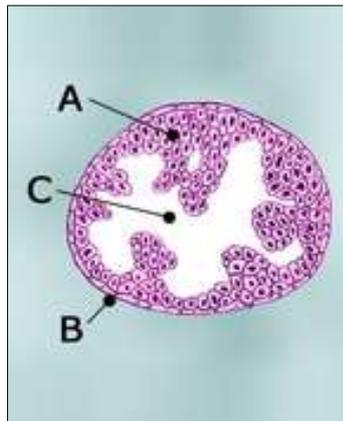
Los tres grados del CDIS son: bajo o grado I; moderado o grado II; y alto o grado III, cuanto más bajo es el grado, estas células se parecen más a las células cancerosas a las células mamarias normales y más lento es su crecimiento con frecuencia resulta difícil ubicar las células en un rango que va de normal a anormal, si bien las células se ubican entre grados o en el límite de éstos, suelen designarlas como “ambiguas”.

### **Grado I (grado bajo) o grado II (grado moderado)**

Las células de un CDIS de grado bajo o grado I son muy similares a las células normales o las células de una hiperplasia ductal atípica. Las células de un CDIS de grado moderado o grado II se reproducen más rápido que las células normales y no son tan similares, los CDIS de grado I y grado II suelen crecer lentamente y a menudo se los describe como CDIS “no comedo” lo que significa que no hay excesivas células cancerosas muertas en el tumor indicando que la enfermedad crece muy lentamente, ya que hay alimento suficiente para todas las células. Cuando un tumor crece muy rápido, algunas de sus células comienzan a morir.

Las personas con CDIS de grado bajo poseen un mayor riesgo de desarrollar cáncer de mama invasivo en el futuro con un riesgo más alto que las personas sin CDIS, no obstante, las personas con CDIS de grado bajo tienen menos probabilidades de recurrencia o de desarrollo de un nuevo cáncer, en cotejo con aquellos pacientes que poseen CDIS de grado alto.

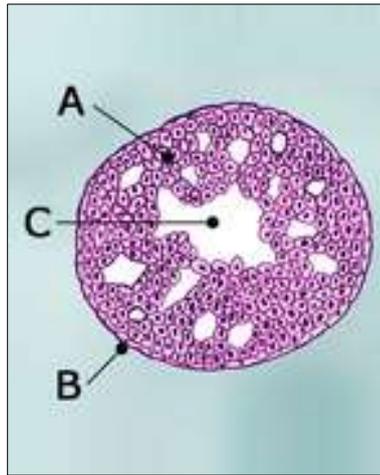
El CDIS papilar tiene sus células cancerosas dentro de sus canales, siguiendo un patrón en forma de dedos, si las células son diminutas, se las denomina micro papilares, entendiéndose en el gráfico “A” las células cancerosas, “B” la membrana basal y “C” el lumen.



**Figura 2-1: CDIS Papilar**

**Fuente:** Breast Cancer;2019.

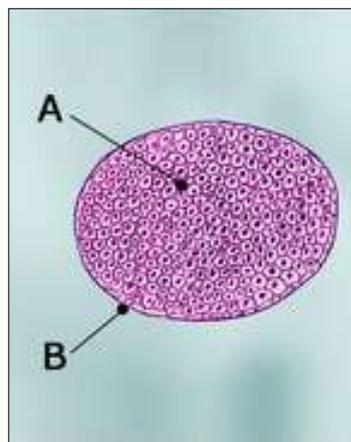
En el CDIS cribiforme hay espacios desocupados entre las células cancerosas dentro de los conductos mamarios con cáncer de mama (similar a la disposición de los orificios del queso suizo) entendiéndose en el gráfico “A” las células cancerosas, “B” la membrana basal y “C” el lumen (dentro de conducto).



**Figura 3-1: CDIS Cribiforme**

**Fuente:** Breast Cancer;2019

En el CDIS sólido las células cancerosas ocupan todo el conducto mamario infectado, hallando en el gráfico “A” las células cancerosas y “B” la membrana basal.



**Figura 4-1: CDIS Sólido**

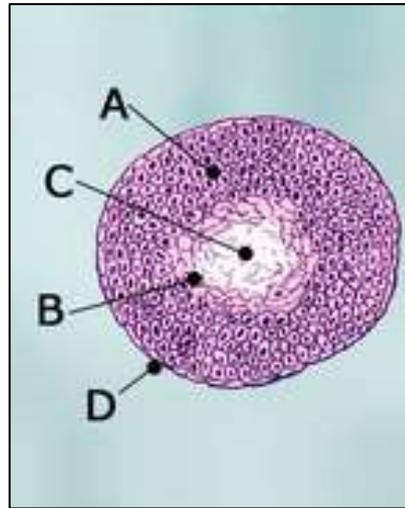
**Fuente:** Breast Cancer;2019

### **CDIS de grado III (grado alto)**

En un patrón de grado alto, las células del CDIS suelen desarrollarse más rápido y tienen un semblante muy diferente a las células mamarias normales y sanas, las personas con CDIS de grado alto tienen un riesgo más alto de poseer un cáncer invasivo, ya sea cuando se diagnóstica el CDIS o en algún otro momento más adelante. También poseen un mayor riesgo de que la enfermedad regrese de forma temprana, dentro de los primeros 5 años y no después de 5 años.

Al CDIS de grado alto suele nombrarse como “comedo” o “comedonecrosis”. Esta hace referencia

a las zonas de células cancerosas muertas conocidas como necróticas que se forman dentro del tumor, cuando las células cancerosas crecen rápidamente generando que otras células no reciban alimento suficiente, estas células privadas de nutrientes suelen morir y dejar zonas de necrosis.



**Figura 5-1:** CDIS de grado alto o comedo

Fuente: Breast Cancer;2019

### 1.3.2 *CDI: Carcinoma ductal invasivo*

El carcinoma ductal invasivo (CDI), denominado carcinoma ductal infiltrante es el tipo común de cáncer de mama, alrededor del 80 % de todos los casos de cáncer de mama son carcinomas ductales invasivos.

El termino invasivo significa que el cáncer ha “invadido” o se ha propagado hacia los tejidos mamarios cercanos a ella. Ductal, significa que el cáncer inicio en los conductos lácteos, que son las “tuberías” que trasladan la leche desde los lobulillos productores de leche a la mama. Carcinoma, hace referencia a todo cáncer que se origina en la piel o en otros tejidos que revisten órganos internos, como el tejido mamario. Dando a entender que el termino, “carcinoma ductal invasivo” hace referencia al cáncer que ha atravesado la pared del conducto lácteo y ha comenzado a irrumpir los tejidos de la mama, con el pasar del tiempo, el carcinoma ductal invasivo puede propagarse hacia los ganglios linfáticos y posiblemente a otras áreas del cuerpo.

Aunque este carcinoma puede afectar a mujeres de cualquier edad, resulta más común a medida que la mujer envejece. Según la Sociedad Americana del Cáncer, aproximadamente 2/3 de las mujeres que son diagnosticadas con cáncer de mama invasivo tienen 55 años o más, este tipo de cáncer también afecta a los hombres.

### **1.3.2.1** *Indicios y síntomas del CDI*

Al principio el carcinoma ductal invasivo no ocasiona ningún síntoma, con frecuencia aparece un área anómala en una mamografía de detección, lo que lleva a más pruebas que se deben realizar.

En algunos casos el primer indicio de cáncer de mama es un cuerpo o masa reciente en la mama que palpan uno mismo o el galeno, según la Asociación Americana del Cáncer, cualquiera de los siguientes cambios inusitados en la mama puede ser una sospecha de cáncer de mama, incluido el carcinoma ductal invasivo, algunos síntomas son la inflamación de la mama o parte de ella, irritación cutánea o formación de hoyos, dolencia de la mama, dolor en el pezón o inversión del pezón, enrojecimiento, descamación o engrosamiento del pezón o la piel de la mama, una secreción del pezón que no sea leche o un bulto en las axilas.

### **1.3.2.2** *Diagnóstico del CDI*

#### **Exploración física de las mamas**

Se puede notar un pequeño bulto en la mama durante la exploración física. Y también se puede palpar los ganglios linfáticos bajo la axila y sobre la clavícula para determinar si hay inflamación u otras alteraciones atípicas.

#### **Mamografía**

Por lo general el carcinoma ductal invasivo se detecta con una mamografía, una prueba que obtiene imágenes de rayos X de la mama. Las mamografías se realizan en mujeres aparentemente sanas para detectar indicios prematuros de cáncer de mama, una característica clave de un cáncer de mama invasivo es la aparición de márgenes espiculados, que significa que en la película de la mamografía el médico observa una anomalía que presenta proyecciones salientes en forma de dedos, lo que revela la “invasión” del cáncer a otros tejidos. Si se detecta un área sospechosa, por lo general se realizan mamografías adicionales para reunir más información acerca del área en cuestión, y si es necesario la mamografía se realizara en ambas mamas.

#### **Ecografía**

Este método emite ondas sonoras desde la mama para obtener imágenes adicionales del tejido. Este estudio puede emplearse para complementar la mamografía.

## **IRM de mama**

La IRM o resonancia magnética usa campos magnéticos, ondas de radio y una computadora para generar imágenes de los tejidos internos del cuerpo. En algunos casos, el médico puede utilizar la IRM de mama para reunir más datos acerca de un área posiblemente afectada por la enfermedad.

## **Biopsia**

Cuando los resultados de la mamografía o de otro estudio de diagnóstico por imágenes sugieren una anomalía, es muy probable que el médico desee hacer una biopsia. Una biopsia implica extirpar parte o la totalidad del tejido de aspecto anormal para que un patólogo (un médico capacitado para diagnosticar cáncer a partir de muestras de una biopsia) pueda examinarlo con un microscopio.

### **1.3.3 *Carcinoma tubular de la mama***

El carcinoma tubular de la mama es un subtipo poco común de carcinoma ductal invasivo (cáncer que comienza dentro del conducto lácteo y se propaga fuera de él). Este subtipo de cáncer denominado carcinoma tubular representa cerca del 1-2 % de todos los casos de cáncer de mama. El tumor generalmente es pequeño y está formado por células con forma de tubo que son de bajo grado, entendiéndose como “Bajo grado” a que tienen una apariencia similar a las células normales sanas y a que suelen crecer lentamente.

El carcinoma tubular de la mama presenta una menor probabilidad de propagarse fuera de la mama que otros tipos de cáncer de mama.

Las investigaciones indican que la edad promedio para el diagnóstico de carcinoma tubular está entre los 45 y los 65-70 años.

#### **1.3.3.1 *Síntomas y diagnóstico del carcinoma tubular de la mama***

### **Indicios y síntomas**

Como otros tipos de cáncer de mama, el carcinoma tubular de la mama puede o no causar algún síntoma al principio. Con el transcurso del tiempo, es posible que un bulto o cuerpo alcance el tamaño suficiente para ser palpado durante una autoexploración o un examen que realice el galeno, pero por lo general, los carcinomas tubulares son pequeños (1 cm o menos de diámetro) y se sienten firmes o duros al tacto.

## Diagnóstico

La mayoría de los carcinomas tubulares se descubren por primera vez durante un examen de mama o mediante una mamografía de detección, en este examen aparece como una pequeña masa de forma irregular por lo que se necesita más pruebas adicionales para determinar con seguridad si se trata de un carcinoma tubular.

El diagnóstico de carcinoma tubular generalmente implica una serie de pasos:

- **Una exploración física de las mamas**, es aquí donde el médico puede tocar el bulto en la mama o puedes tocarlo uno mismo durante una autoexploración.
- **Una mamografía** puede ayudar a localizar el tumor y buscar indicios de cáncer en otras áreas de la mama.
- **Una IRM, ecografía, o ambas**: se la usa para obtener más imágenes de las mamas y detectar la presencia de cáncer en otras áreas que pueden posiblemente estar afectadas.
- **Una biopsia**, consiste en extraer el tumor entero o parte de él para examinarlo con el microscopio, esta muestra puede extraerse introduciendo una aguja especial o haciendo una incisión pequeña. Este método es clave para un diagnóstico exacto, ya que las pruebas de diagnóstico por imágenes solas no distinguen el carcinoma tubular de otros tipos de cáncer de mama y de otras afecciones benignas de la mama.

### 1.3.4 *Carcinoma medular de la mama*

El carcinoma medular de la mama es un subtipo muy poco común del carcinoma ductal invasivo (cáncer que comienza dentro del conducto lácteo y se propaga fuera de él) lo que representa cerca del 3 al 5 % de todos los casos de esta enfermedad. Se denomina carcinoma “medular” porque el tumor es una masa suave y pulposa que se encuentra en una parte del cerebro conocida como bulbo raquídeo o médula.

El carcinoma medular puede aparecer en cualquier edad, pero por lo general afecta a mujeres de entre 45 y 55 años. El carcinoma medular afecta con más repetición a mujeres que tienen una mutación del gen *BRCA1*. Las investigaciones también indican que el carcinoma medular es más común en Japón que en los Estados Unidos.

#### 1.3.4.1 *Síntomas y diagnóstico del carcinoma medular de la mama*

Este tipo de cáncer denominado carcinoma medular de la mama puede no causar ningún síntoma al inicio de la enfermedad por lo que, con el paso del tiempo, puede formarse un bulto, que puede

ser suave y carnoso o algo firme al tacto, la mayoría de los carcinomas medulares son pequeños es decir menos de 2 cm de tamaño. El carcinoma medular también puede causar dolencia, inflamación, enrojecimiento o sensibilidad en la mama.

### **Diagnóstico**

Las pruebas obtenidas por imágenes del tejido dentro de la mama, como la mamografía y la ecografía, a veces pueden encontrar un carcinoma medular, en estas pruebas, el carcinoma medular surge como un bulto pequeño y bien determinado, sin embargo, algunas investigaciones sugieren que la mamografía no es confiable para detectar carcinomas medulares. En un estudio se reveló que existe más probabilidad que estos tipos de cáncer se detecten durante una autoexploración o durante una exploración de las mamas realizada por un médico. El diagnóstico de carcinoma medular generalmente involucra una serie de pasos: una exploración física de las mamas, una mamografía, una ecografía o una biopsia.

#### **1.3.4.2** *Tratamiento para el carcinoma medular de la mama*

Con respecto al tratamiento algunos médicos consideran que un carcinoma medular puro puede requerir únicamente cirugía sin ningún otro tratamiento adicional como quimioterapia o terapia de radiación, esto se debe a que el carcinoma medular por lo general es un tumor pequeño no agresivo que rara vez se expande a los ganglios linfáticos.

Otros médicos creen que todos los casos de carcinoma medular deben tratarse como a un carcinoma ductal invasivo. Este tipo de cáncer es poco frecuente y muy difícil de diagnosticar con certeza, y algunos galenos posiblemente quieran asegurarse de no estar administrando el tratamiento correcto para el cáncer. El carcinoma medular atípico debe tomar los mismos tratamientos que el carcinoma ductal invasivo.

#### **1.3.5** *Carcinoma mucinoso de la mama*

El carcinoma mucinoso de la mama, a veces llamado carcinoma coloideo, es una forma poco habitual de carcinoma ductal invasivo (cáncer que comienza dentro del conducto lácteo y se irradia fuera de él). El carcinoma mucinoso de la mama representa cerca del 2-3 % de todos los casos de cáncer de mama, el tumor se forma a partir de células anómalas que “flotan” en acopios de mucina, uno de los principales componentes de la sustancia viscosa y escurridiza conocida como moco.

Usualmente, esta mucosidad cubre la mayoría de las superficies internas del cuerpo humano, entre ellas, el tubo digestivo, los pulmones, el hígado y otros órganos vitales. Muchos tipos de células cancerosas (entre ellas, la mayoría de las células de cáncer de mama) producen esta mucosidad, sin embargo, en el carcinoma mucinoso, la mucosidad se transforma en una parte principal del tumor y rodea las células del cáncer de mama. Este cáncer suele afectar a las mujeres postmenopáusicas, algunos estudios indican que la edad promedio al momento del diagnóstico es de 60 años o más.

#### **1.3.5.1 Síntomas y diagnóstico del carcinoma mucinoso de la mama**

##### **Indicios y síntomas**

El carcinoma mucinoso de la mama como en otros casos no puede causar algún síntoma al inicio de la enfermedad. Con el transcurso del tiempo, es posible que un bulto o cuerpo alcance el tamaño significativo para ser palpado durante una autoexploración o un examen que realice el médico, estos tumores por lo general miden entre 1 cm y 5 cm.

##### **Diagnóstico**

El análisis del carcinoma mucinoso por lo general implica una serie de pasos como una exploración física de las mamas, una mamografía, una ecografía, una IRM o una biopsia

El carcinoma mucinoso puro presenta otras características clave:

- **Receptores de hormonas positivos:** las investigaciones indican que el carcinoma mucinoso puro presenta receptores de estrógeno positivos en el 90-100 % de los casos y receptores de progesterona positivos en el 50-68 % de los pacientes.
- **HER2 negativo:** este tipo de cáncer generalmente muestra receptores de la proteína HER2/neu negativos.
- **Ganglios linfáticos negativos:** el carcinoma mucinoso puro raras veces se propaga a los ganglios linfáticos, especialmente si el tumor tiene 1-2 cm o menos. En el caso de tumores de mayor tamaño, es viable que haya propagación a los ganglios linfáticos, en algunas ocasiones, el cáncer en los ganglios linfáticos indica que el tumor es en realidad un carcinoma mucinoso mixto, con apariencia de células de carcinoma ductal invasivo.

### **1.3.5.2 Tratamiento para el carcinoma mucinoso de la mama**

El carcinoma mucinoso puro es más fácil de tratar que el carcinoma ductal invasivo, de manera que no requiere tanto tratamiento, en especial si el tumor es diminuto y el cáncer no se ha extendido a los ganglios linfáticos.

Dentro del plan de tratamiento puede estar la cirugía para extirpar el cáncer y los ganglios linfáticos afectados y terapia adyuvante.

### **1.3.6 Carcinoma papilar de la mama**

Los carcinomas papilares invasivos de la mama son poco habituales y representan menos del 1-2 % de los casos de cáncer de mama invasivo, en la mayoría de los pacientes, estos tipos de tumor se diagnostican en mujeres de avanzada edad postmenopáusicas. El carcinoma papilar invasivo por lo general tiene un borde bien definido y está conformado por prolongaciones pequeñas en forma de dedos. Con frecuencia, es de grado 2 es decir un grado moderado, en una escala de 1 a 3, en que el grado 1 describe células cancerosas que se ven y se desarrollan en forma muy similar a las células sanas normales y el grado 3 describe células muy anormales de rápido desarrollo. En la mayoría de los casos de carcinoma papilar invasivo, también está presente un carcinoma ductal in situ (CDIS) por lo que su método de tratamiento es muy similar.

### **1.3.7 Carcinoma cribiforme de la mama**

En el carcinoma cribiforme invasivo, las células cancerosas irrumpen el estroma (tejidos conjuntivos de la mama) en formaciones que se parecen a un nido entre los canales y los lobulillos. Dentro del tumor hay orificios individuales entre las células cancerosas, por lo que este se parece a un queso suizo. Por lo general, el carcinoma cribiforme invasivo es de bajo grado, lo que significa que las células se ven y se comportan de forma similar a las células de la mama sanas y normales. En un estudio relacionado a este tipo de cáncer se concluyó que cerca del 5-6 % de los casos de cáncer de mama invasivo, puede considerarse cribiforme. Con frecuencia, también hay indicios de carcinoma ductal in situ (CDIS) de tipo cribiforme por lo que su tratamiento puede ser muy análogo.

### **1.3.8 CLI: Carcinoma lobular invasivo**

El carcinoma lobular invasivo (CLI), denominado como carcinoma lobular infiltrante, es el segundo tipo de cáncer de mama más frecuente después del carcinoma ductal invasivo (cáncer que comienza en los conductos lácteos y se propaga más allá de ellos), según la Sociedad Americana del Cáncer, más de 180.000 mujeres en los Estados Unidos descubren cada año que padecen un cáncer de mama invasivo, alrededor del 10 % de estos casos de cáncer son carcinomas

lobulares invasivos y alrededor del 80 % son carcinomas ductales invasivos).

Los carcinomas lobulares invasivos pueden afectar a mujeres de cualquier edad, son más comunes en mujeres de edad avanzada, según la Sociedad Americana del Cáncer, aproximadamente dos tercios de las mujeres que son establecidas con cáncer de mama invasivo con edades entre los 55 años o más, este tipo de cáncer tiende a aparecer en edades más avanzadas que los carcinomas ductales invasivos.

En muchas investigaciones se sugiere que el uso de terapias de reemplazo hormonal durante y después de la menopausia puede incrementar el riesgo de desarrollar un CLI.

#### **1.3.8.1** *Indicios y síntomas del CLI*

Los carcinomas lobulares invasivos son más dificultosos de visualizar en las mamografías que los carcinomas ductales invasivos, esto se debe a que, en lugar de formar un bulto o cuerpo, las células cancerosas casi siempre estas se propagan al tejido conjuntivo circundante en una formación en línea.

El primer indicio de CLI es el engrosamiento o endurecimiento de la mama que se puede sentir al contacto con la mama, más que un bulto bien definido, otros síntomas posibles contienen un área agrandada o inflamada, dando como resultado un cambio en la textura de la piel e invasión del pezón.

Según la Asociación Americana del Cáncer, cualquiera de los siguientes cambios inusitados en la mama puede ser un primer indicio de este tipo de cáncer, incluido el carcinoma lobular invasivo:

- Inflamación de la mama o parte de ella
- Irritación cutánea o formación de hoyos
- Dolor de mama
- Dolor en el pezón
- Enrojecimiento, descamación o engrosamiento del pezón o la piel de la mama
- Una secreción del pezón que no sea leche
- Un bulto en las axilas

### **1.3.9 CMI: Cáncer de mama inflamatorio**

El cáncer de mama de tipo inflamatorio (CMI) es una forma de cáncer poco frecuente y agresivo, según datos del Instituto Nacional del Cáncer, alrededor del 1 al 5 % de todos los casos de cáncer de mama en los Estados Unidos son de cáncer de mama inflamatorio.

El cáncer de mama inflamatorio por lo general comienza con el enrojecimiento e inflamación de la mama, en lugar del bulto distintivo, esta enfermedad suele reproducirse y propagarse velozmente, y sus síntomas desmejoran en apenas unos días e incluso horas, por lo que resulta fundamental reconocer los síntomas y procurar un método de curación inmediata. Si bien el diagnóstico de cáncer de mama inflamatorio es grave, se debe tener en cuenta que los tratamientos actuales controlan la enfermedad mejor que en el pasado.

La edad promedio al instante del diagnóstico de cáncer de mama inflamatorio en los EE.UU es de 57 años para mujeres de etnia blanca y de 52 años para mujeres afroamericanas, esto representa alrededor de 5 años menos que la edad promedio de diagnóstico de otras formas de cáncer de mama, de acuerdo con la Sociedad Americana del Cáncer, este tipo de enfermedad es más común en las mujeres afroamericanas. En un estudio realizado en 2008 se reveló que el sobrepeso aumenta las probabilidades de que una persona desarrolle CMI, Al igual que otras formas de cáncer de mama.

#### **1.3.9.1 Síntomas del CMI**

Los síntomas más frecuentes de CMI son:

- **Enrojecimiento de la mama:** esta puede ser en la totalidad o en parte de la mama y es el principal indicio de cáncer de mama inflamatorio, aunque veces el enrojecimiento aparece y desaparece.
- **Inflamación de la mama:** puede suceder en la totalidad o parte de la mama puede además de agrandarse y endurecerse.
- **Calor:** puedes sentir calor en la mama.
- **Apariencia de cáscara de naranja:** la mama puede inflamarse y lograr la apariencia de una cáscara de naranja de ombligo.
- **Otros cambios en la piel de la mama** está en su coloración ya que puede lograr un color rosado o verse amoratada o presentar especies de surcos, ronchas o urticaria.

- Inflamación de los ganglios linfáticos debajo del brazo o sobre la clavícula pueden ver notorias con el tacto.
- Pezón invertido o plano.
- Dolor o ardor

Algunos de estos síntomas son similares a los causados por la mastitis que es una infección de la mama que pueden padecer las mujeres que están dando de lactar.

### 1.3.9.2 *Diagnóstico y determinación del estadio del CMI*

Debido a que el CMI se desarrolla con mayor rapidez a otros tipos de cáncer, normalmente se lo detecta en un estadio localmente avanzado, lo que significa que las células cancerosas se han desarrollado en el tejido mamario y propagado a los ganglios linfáticos circundantes, casi todas las personas a quienes se les diagnostica CMI presentan indicios de cáncer en los ganglios linfáticos. Aproximadamente 1 de cada 3 personas con CMI, su cáncer se ha propagado de la mama a otras partes del cuerpo.

Dado el caso que este cáncer de mama inflamatorio no comienza normalmente como un bulto definido sino en forma de cambios en la piel, generalmente para el diagnóstico se utiliza una biopsia de piel.

Una vez que se diagnosticó CMI, se realizan análisis adicionales para determinar si el cáncer se ha extendido fuera de la mama a otros órganos, como los pulmones, los huesos o el hígado, este proceso se denomina determinación de estadio, los análisis que pueden incluirse son radiografía de tórax, tomografía computarizada del tórax, el abdomen y la pelvis, exploración ósea y análisis de la función hepática.

Después de finalizar los estudios se determina la clasificación del estadio, el médico describirá el cáncer de mama inflamatorio como estadio IIIB, estadio IIIC o estadio IV

- **Estadio III B**, esto significa que el cáncer se ha extendido a los tejidos próximos a la mama, como la piel de la mama o la pared torácica, incluyendo las costillas y los músculos del pecho, este cáncer se puede haber propagado a los ganglios linfáticos de la mama o debajo del brazo.
- **Estadio III C**, significa que el cáncer se ha extendido a los ganglios linfáticos que se hallan debajo de la clavícula y cerca del cuello. El cáncer también se pudo haber propagado a los

ganglios linfáticos situados en la mama o debajo del brazo, así como a los tejidos circundantes a la mama.

- **Estadio IV**, significa que el cáncer se ha extendido a otros órganos, estos incluyen los huesos, los pulmones, el hígado y el cerebro, además de los ganglios linfáticos del cuello.

### **1.3.10 CLIS: Carcinoma Lobular In situ**

El carcinoma lobular in situ (CLIS) es un área o varias áreas de crecimiento celular anómalo que desarrollan el riesgo de que una persona contraiga cáncer de mama invasivo en el futuro, a pesar de incluir el término “carcinoma” en su nombre, el CLIS no es un verdadero cáncer de mama. En realidad, el CLIS es una señal de que la persona presenta un riesgo mayor al promedio de padecer cáncer de mama en el futuro, es por esto que algunos expertos prefieren el término “neoplasia lobular” en lugar de “carcinoma lobular”.

El CLIS se determina generalmente antes de la menopausia, más repetidamente entre los 40 y 50 años, menos del 10 % de las mujeres diagnosticadas con CLIS ya han atravesado la menopausia. Esta enfermedad es muy poco frecuente en hombres.

#### **1.3.10.1 Síntomas y Diagnóstico del CLIS**

En general, el CLIS no causa sospechas ni síntomas como bultos ni otros cambios visibles en la mama y no siempre es visible en una mamografía de detección. El principal motivo es que con frecuencia carece de microcalcificaciones, los pequeños depósitos de calcio que se forman dentro de otros tipos de células de cáncer de mama. En una mamografía, las microcalcificaciones surgen como manchas blancas por lo que se cree que muchos casos de CLIS no son diagnosticados y puede que nunca causen ningún problema.

Generalmente el CLIS se diagnostica después de una biopsia realizada en la mama por algún diferente a este, como un resultado anormal de una mamografía o la presencia de un bulto mamario sospechoso.

### **1.3.11 Cáncer de mama en hombres**

El cáncer de mama en hombres es una enfermedad muy poco frecuente, menos del 1 % de todos los casos de cáncer de mama se producen en hombres. En una investigación realizada en 2017,

252.710 mujeres fueron diagnosticadas con cáncer de mama en los Estados Unidos, en el cual fueron 2.470 los hombres que recibieron ese diagnóstico.

En la sociedad se piensa de forma errónea que los hombres no tienen mamas, la verdad es que tanto niños como niñas, hombres como mujeres, tienen tejido mamario. Las distintas hormonas en el cuerpo de las niñas y mujeres estimulan el tejido mamario para desarrollar plenamente las mamas. Regularmente, el cuerpo de los niños y los hombres utiliza muy poco las hormonas que estimulan las mamas.

#### **1.3.11.1** *Factores de riesgo de cáncer de mama en hombres*

Existe una serie de factores que pueden aumentar el riesgo de que los hombres desarrollen cáncer de mama:

- **Envejecimiento:** este es el factor más importante para contraer esta enfermedad, tal como sucede con las mujeres, el riesgo aumenta a medida que pasan los años, la edad promedio de los hombres que son diagnosticados con cáncer de mama es alrededor de los 67 años por lo que esto significa que la mitad de ellos tiene más de 67 años.
- **Niveles altos de estrógeno:** la presencia de estrógeno ayuda a la multiplicación celular mamaria, tanto la normal como la anormal. En los hombres pueden tener altos niveles de estrógeno a causa de tomar medicinas hormonales, tener obesidad, estar expuestos a estrógenos en el ambiente, como por ejemplo en químicos o comidas para engordar el ganado, consumir grandes cantidades de alcohol, lo cual puede afectar la función hepática que regula los niveles de estrógeno en sangre.
- **Síndrome de Klinefelter:** los hombres con síndrome de Klinefelter poseen niveles bajos de andrógenos conocidas como hormonas masculinas y niveles altos de estrógeno llamadas hormonas femeninas. Por lo que se poseen un mayor riesgo de contraer una ginecomastia y cáncer de mama.
- **Antecedentes marcados de cáncer de mama en la familia o alteraciones genéticas:** los antecedentes familiares pueden incrementar el riesgo en los hombres para desarrollar cáncer de mama, especialmente si hay otros hombres en la familia que hayan sido fijados con la enfermedad.
- **Exposición a la radiación:** someterse a terapia de radiación en el tórax antes de los 30 años es un factor que ayuda a contraer la enfermedad, y en especial durante la adolescencia, esto se ha visto en jóvenes que recibieron radiación para tratar la enfermedad de Hodgkin

### **1.3.11.2** *Síntomas del cáncer de mama en hombres*

Un estudio reveló que el cáncer de mama en hombres está en subida ya que en el período de 25 años comprendido entre 1973 y 1988, aumentó un 25 %, o queda claro si dicho aumento significa que la enfermedad progresivamente se hace cada vez más frecuente o si los hombres entienden mejor los síntomas y los comunican, lo que permite realizar diagnósticos más tempranos de esta enfermedad.

Si se observa cambios constantes en las mamas, se debe consultar a un médico especialista ya que estos son algunos indicios a los que debes proporcionar atención, algunos indicios son, un bulto en la mama, dolor en el pezón, secreciones en el pezón (transparentes o con sangre), un pezón invertido, úlceras en el pezón y la areola (el pequeño anillo de color alrededor del centro del pezón) o ganglios linfáticos dilatados debajo del brazo.

### **1.3.12** *Enfermedades de Paget en el pezón*

La enfermedad de Paget en el pezón es una forma poco habitual de cáncer de mama en que se acopian células cancerosas en el pezón o a su proximidad, este cáncer generalmente afecta en primer lugar a los conductos del pezón (pequeños canales que transportan leche) y luego se propaga hasta la superficie del pezón y la areola que es el círculo oscuro de piel que rodea el pezón. El pezón y la areola por lo general se vuelven de forma escamosa, se enrojecen, producen comezón y se irritan.

Es importante reconocer los síntomas, ya que más del 97 % de las personas que sufren la enfermedad de Paget también tienen cáncer, ya sea CDIS (cáncer ductal in situ) o cáncer invasivo, en alguna parte de la o las mamas.

#### **1.3.12.1** *Síntomas y diagnóstico*

La enfermedad de Paget incita enrojecimiento, dolor y descamación en la piel que cubre y rodea el pezón, al inicio estos síntomas tienden a aparecer y desaparecer.

Con el pasar del tiempo, esos síntomas empeoran y pueden incluir: comezón, cosquilleo, sensación de ardor, dolor, sensibilidad, descamación o engrosamiento de la piel, aplanamiento del pezón y secreción amarillenta o sanguinolenta del pezón.

Para determinar la enfermedad de Paget, regularmente se necesitan los siguientes exámenes:

exploración física de las mamas, mamografía, ecografía y biopsia.

#### **1.3.12.2** *Tratamiento para la enfermedad de Paget*

El tratamiento estándar recomendado contra la enfermedad de Paget es la cirugía. El tipo de cirugía que se necesita dependerá de qué extensión de tu mama esté afectada por el cáncer.

El tipo de cirugía necesaria para tratar la enfermedad de Paget depende de lo adelantado que esta enfermedad, además de suprimir el pezón y la areola, el cirujano deberá quitar todo el tejido canceroso que haya dentro de la mama y si bien la mastectomía ha sido tradicionalmente el principal tratamiento para la enfermedad de Paget, en algunos casos es viable hacer una cirugía más limitada conocido a veces como “cirugía de conservación de la mama”, este procedimiento busca eliminar el cáncer preservando el tejido sano de la mama.

#### **1.3.13** *Tumor filoides de la mama*

Los tumores filoides de las mamas son muy poco comunes y representan menos del 1 % de todos los tumores mamarios que se diagnostiquen. El término “filoides”, significa “similar a una hoja”, y hace referencia al hecho de que las células del tumor se multiplican siguiendo un patrón en forma de hoja, otro nombre para designar este tipo de tumor es “cistosarcoma filoides”, estos tumores suelen crecer de una manera rápida.

Si bien la mayoría de los tumores filoides son benignos es decir no cancerosos, algunos son malignos o cancerosos y otros, ambiguos que es un punto intermedio entre cancerosos y no cancerosos. Los tres tipos de tumores filoides tienden a crecer con prisa y exigen cirugía para reducir el riesgo de recurrencia en la mama.

Los tumores filoides pueden surgir a cualquier edad de la mujer, pero tienden a formarse posteriormente de los 40 años. Los tumores filoides benignos por lo general se diagnostican a edades más tempranas que los malignos, estos tumores filoides son extremadamente poco comunes entre los hombres.

#### **1.3.14** *Cáncer de mama recurrente o metastásico*

Una recurrencia o el cáncer de mama recurrente es un cáncer de mama que vuelve a surgir después de un determinado curso de tiempo en el que ya no fue detectado.

El cáncer puede volver en la misma mama del diagnóstico original o en el peor de los casos surgir en la otra mama o en la pared torácica.

Una metástasis o un cáncer de mama metastásico es un cáncer que se ha propagado a otra parte del cuerpo sus células cancerosas pueden desprenderse del tumor original de la mama y albergarse en otras partes del cuerpo usando el torrente sanguíneo o el sistema linfático que es una gran red de ganglios y vasos que ayudan a eliminar bacterias, virus y desechos celulares.

El cáncer de mama puede volver en la zona de las mamas o en otra parte del cuerpo meses o incluso años después del diagnóstico y el tratamiento original, se calcula que casi el 30 % de las mujeres diagnosticadas con cáncer de mama en etapa temprana desarrollarán un cáncer metastásico.

#### **1.4 Análisis Estadístico Implicativo**

Es un método de análisis no simétrico de datos que permite analizar datos, a partir de un conjunto de factores que se interrelacionan en una población de sujetos u objetos con un conjunto de variables, la extracción y estructuración del conocimiento en forma de reglas y normas generalizadas a partir de la contingencia de estas reglas, la definición en consecuencia es una determinada previsión en distintas ramas del saber. Su origen es la modelización estadística de la cuasi-implicación: cuando la variable o el conjunto de variables “a” es observada en la población, entonces por lo general la variable “b” lo es también. (Zamora et al., 2002)

##### **1.4.1 Tipos de variables**

###### **Binarias**

Los factores cualitativos aparecen a menudo bajo la forma de información binaria (dicotómica). Ejemplos como, un individuo ha participado o no en un programa de formación profesional, una empresa ofrece o no un plan de pensiones a sus trabajadores, etc. Estas variables mudables se representan mediante variables binarias que toman los valores cero y uno, y se denominan variables ficticias o, utilizando el término en inglés, variables dummy (Greene, 2011).

###### **Modal**

Ayuda a responder preguntas como: ¿En una escala de 1 a 5, ¿cómo evaluarías tu gusto por la música ROCK?, donde los extremos de la escala representan carencia total de gusto por la música ROCK y gusto máximo por ésta, respectivamente. En este caso estamos en presencia de una variable mudable que puede tomar los valores de 1 a 5.

## **Frecuencial**

Deposita valores de preguntas tales como, al día, ¿cuántos horas escuchas la música ROCK?, se debe responder con un valor numérico que representa la periodicidad diaria con que escucha este tipo de música, por lo que esta variable será una variable frecuencial, que puede tomar, en este ejemplo en particular, los valores desde 0 hasta 7. Se puede observar que a pesar de tomar valores similares las variables modales y frecuenciales, el significado de estos es distinto.

## **Intervalo**

Responde preguntas como, ¿Cuál es la edad promedio que tienen los estudiantes de la ESPOCH, a los cuales les gusta la música ROCK? A esta pregunta se le debe responder con un valor numérico, que, en este caso particular, por referirse a la edad de estudiantes de la ESPOCH, debe pertenecer al intervalo [18, 23] años.

### **1.4.2 Análisis Clasificadorio**

El Análisis Clasificadorio tiene, como objetivo esencial, agrupar aquellos objetos que reúnan características similares, es decir, es una técnica de análisis exploratorio delineada para revelar las agrupaciones o clúster naturales dentro de una población.

Su objetivo es aglomerar objetos similares, se necesita alguna medida para evaluar las diferencias y similitudes entre objetos. La definición de similaridad es importante, esta es una medida de correspondencia o semejanza entre los objetos que van a ser agrupados. La estrategia más común consiste en medir la equivalencia en términos de la distancia entre los pares de objetos. Los objetos con distancias reducidas entre ellos son más parecidos entre sí, que aquellos que tienen distancias mayores y se agruparán, por tanto, dentro del mismo cluster. De esta manera, cualquier objeto puede compararse con cualquier otro, a través de la medida de similaridad.

Para realizar este análisis, se parte de considerar un conjunto **I** formado por **n** individuos y un conjunto **A** formado por **p** características,  $A = \{a_1, a_2, \dots, a_p\}$ , se supone además que:

$$A_i = \{x \in I: a_i(x) = 1\}, \text{Card}(I) = n \text{ y } \text{Card}(A_i) = n_{a_i} \quad (1.1)$$

### **1.4.3 Árbol de Similaridad**

Para cada pareja de variables ( $a_i, a_j$ ) se calculan los Índices de Proximidad o Similaridad:

$$s(a_i, a_j) = \Pr[\text{Card}(X_i, X_j) \leq K] \quad (2.1)$$

Donde  $K = \text{Card}(A_i \cap A_j)$ , es decir, el valor observado de la cantidad de co-presencias

entre  $a_i$  y  $a_j$ .

#### 1.4.4 Definición de Análisis Estadístico Implicativo

Al igual que en el análisis clasificatorio, se parte de considerar un conjunto  $I$  formado por  $n$  individuos y un conjunto  $A$  formado por  $n$  características,  $A = \{a_1, a_2, \dots, a_p\}$ , se supone también que:

$$A_i = \{x \in I: a_i(x) = 1\}, \text{Card}(I) = n, \text{Card}(A_i) = n_{a_i} \text{ y } \text{Card}(\bar{A}_i) = n_{\bar{a}_i} \quad (3.1)$$

El análisis se desarrolla a través del cálculo de los índices de implicación y de cohesión, en el estudio cohesivo se pueden determinar la tipicalidad y contribución de sujetos y variables, así como los nodos significativos para la formación de las clases. Los índices de implicación y de cohesión se definen en términos de una probabilidad, la diferencia está en la variable aleatoria de interés, que ahora es  $\text{Card}(X_i \cap \bar{X}_j)$ , es decir, la cuantía de individuos que poseen la característica  $a_i$ , pero no poseen la característica  $a_j$ . Todo lo explicado anteriormente para la variable aleatoria  $\text{Card}(X_i \cap X_j)$  es también válido para  $\text{Card}(X_i \cap \bar{X}_j)$ ,

#### 1.4.5 Intensidad de la implicación entre variables

En lógica matemática, la regla " $a_i \Rightarrow a_j$ " es verdadera si para todo  $x$ ,  $a_j(x)$  sólo es nulo cuando  $a_i(x)$  lo sea también; es decir si el conjunto  $A$  de los  $x$  por los cuales  $a_i(x) = 1$  está contenido en el conjunto  $B$  de los  $x$  para los cuales  $a_j(x) = 1$ .

Sin embargo, en esta fijación estricta se observa excepcionalmente en la realidad. En el caso de un cuestionario de conocimientos se podrían observar a algunos raros pacientes que tengan una enfermedad  $a_i$  y no consiguen recuperarse el  $a_j$ , sin que se contradiga la tendencia a tener  $a_j$  cuando se tiene  $a_i$ . La formalización matemática de esta situación se expresa a través de la cuasi-implicación  $i \rightarrow a$  (Gras, 1999).

La cuasi-implicación  $a_i \rightarrow a_j$  significa que cuando  $a_i$  está presente entonces generalmente  $a_j$  está también presente.

Los nodos internos del árbol dirigido, los cuales representan la jerarquía dirigida, describen relaciones implicativas complejas, llamadas R-reglas, entre los atributos de  $A$ .

- Cuando  $R \rightarrow a_i$ ,  $a_i$  puede ser interpretada como una consecuencia de  $R$ .
- La R-regla  $a_i \rightarrow R$ , significa que una R-regla  $R$  puede ser deducida de la observación de  $a_i$ .

- La R-regla  $R' \rightarrow R''$ , significa que la propiedad  $R''$  es el corolario de una propiedad previamente definida  $R'$ .

El objetivo de la Intensidad Implicativa es expresar la inverosimilitud (“asombro”) del número de contraejemplos que invalidan la regla  $a_i \rightarrow a_j$ , para lo cual se coteja el número de contraejemplos observados con el número de contraejemplos esperado bajo la hipótesis de ausencia de relación.

#### 1.4.6 *Análisis cohesivo*

La cohesión surge como una medida de la calidad implicativa de la R-regla. El objetivo es descubrir R-reglas  $R' \rightarrow R''$  con una fuerte relación implicativa entre las componentes de  $R'$  y las de  $R''$ . Por ejemplo, es natural formar la regla  $(a_1 \rightarrow a_2) \rightarrow (a_3 \rightarrow a_4)$  si las relaciones implicativas  $a_1 \rightarrow a_3$ ,  $a_1 \rightarrow a_4$ ,  $a_2 \rightarrow a_3$  y  $a_2 \rightarrow a_4$  son lo suficientemente significativas.

La **Cohesión de una R-regla**  $a_i \rightarrow a_j$  de grado 1 es:

$$Coh(a_i, a_j) = \begin{cases} \sqrt{1 - E^2} & \text{si } p \geq 0.5 \\ 0 & \text{en caso contrario} \end{cases} \quad (4.1)$$

#### 1.4.7 *Tipicalidad y contribución*

El cálculo de la tipicalidad y/o contribución indica cuál es el individuo o variable más típica o contributiva, y también qué otras categorías pueden conducir a los mismos valores de tipicalidad y/o contribución. La noción de tipicalidad se concreta por el hecho de que ciertos individuos son “típicos” del comportamiento de la población, es decir con una intensidad de implicación similar a la de la regla. La noción de contribución se define para determinar los individuos que contribuyen bien a la creación de la regla. Estos individuos son más comprometidos que otros en la formación de la regla.

El método para la obtención de la tipicalidad y contribución de los individuos a las clases que se forman en el árbol de similaridad y en el cohesivo es el mismo, los métodos se diferencian en el hecho de que en el primero se utiliza el índice de similaridad y en el segundo el índice de intensidad implicativa entre variables. Un sujeto típico, es aquél que verifica todas las implicaciones que poseen mayor intensidad de implicación en la formación de las clases.

Se define la distancia de un individuo  $x$  a la clase  $C$  dada la siguiente fórmula:

$$d^2(x, C) = \frac{1}{g} \sum_{i=1}^g \frac{\varphi_i - \varphi_{xi}}{1 - \varphi_i} \quad (5.1)$$

Donde  $i$  denota una regla de  $C$ .

La Tipicalidad del individuo  $x$  se define como:

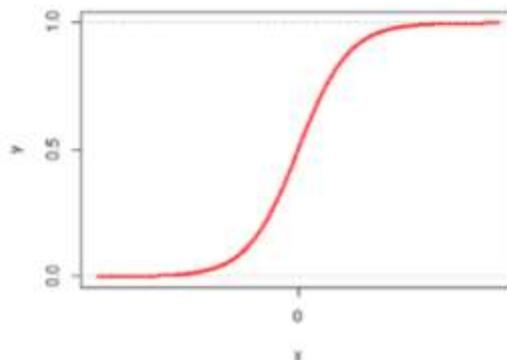
$$\gamma(x, C) = 1 - \frac{d(x, C)}{\text{Max } d(y, C)} \quad (6.1)$$

## 1.5 Regresión Logística

La regresión logística es una técnica estadística multivariable propuesta para el análisis de una relación de dependencia entre una variable dependiente y un conjunto de variables independientes, de forma similar a como actúa el análisis de regresión lineal clásico.

El análisis de regresión logística tiene dos modalidades: la regresión logística binaria que es cuando se pretende explicar una particularidad o suceso dicotómico, por ejemplo, estar desempleado o no, abstenerse en las elecciones o no y la regresión logística multinomial que es el caso más general de querer explicar una variable cualitativa politómica (López et al., 2016).

El modelo se formaliza a partir de la función logística, donde se diseña una relación funcional, aquí la función es la logística, una función con forma de curva sigmoideal:



**Figura 6-1:** Curva de la Regresión Logística

Realizado por: López et al;2016.

El modelo logístico obedece a la siguiente relación entre la probabilidad de que ocurra el suceso, dado que el individuo presenta los valores  $X_1=x_1, X_2=x_2, \dots, X_p=x_p$ .

$$\Pr(Y = 1 | x_1, x_2, \dots, x_p) = \frac{1}{1 + \exp(-\alpha - \beta_1 x_1 - \beta_2 x_2 - \dots - \beta_p x_p)} \quad (7.1)$$

Otra forma de presentar la relación es:

$$\text{logit}(\Pr(Y = 1|x)) = \log\left(\frac{\Pr(Y = 1|x)}{1 - \Pr(Y = 1|x)}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p \quad (8.1)$$

Donde denotamos con  $\Pr(Y = 1|x)$  la probabilidad condicional  $\Pr(Y = 1|x_1, x_2, \dots, x_p)$

Para una mejor interpretación de los coeficientes  $\beta_i$  (cualquiera de los coeficientes  $\beta_i$ ) es obligatorio referirnos al concepto de riesgo relativo, ya que esta se define como la razón entre la probabilidad de que dicho suceso ocurra ( $P$ ) y la probabilidad de que este no ocurra ( $1 - P$ ). El exponencial de los  $\beta_i$  se corresponde con el riesgo relativo, o sea, es una medida del dominio de la variable  $X_i$  sobre el riesgo de que ocurra ese hecho y suponiendo que el resto de las variables del modelo permanezcan constantes.

Un intervalo de confianza para el exponencial de  $\beta$  que contenga al 1 indica que la variable no tiene una influencia significativa en la ocurrencia del suceso y, por el contrario, valores más alejados de este indican una mayor influencia de la variable.

### 1.5.1 Aspectos para el uso de la Regresión Logística

#### **Tamaño de muestra y número de variables independientes**

Una de las ventajas de la regresión logística es que accede a el uso de múltiples variables con relativamente pocos casos, sin embargo, hay que tener en cuenta algunas moderaciones, se ha sugerido que el número de sujetos para poder usar esta técnica estadística sin problemas debe ser superior a 10 ( $k+1$ ) donde  $k$  es el número de variables explicativas; por tanto, si se introducen interacciones o variables mudables, el número de individuos en la muestra debe aumentar. Además, se ha propuesto que, si una de las variables dicotómicas no tiene al menos 10 casos en cada uno de sus 2 valores posibles, entonces las estimaciones no son fiables (Silva, 2018).

#### **Efecto sobre el riesgo de que ocurra el evento**

Es necesario tener en cuenta el efecto sobre el riesgo de que ocurra el evento, de los cambios de las variables explicativas cuando son cuantitativas (continuas), en ocasiones es necesario categorizarlas, ya que los cambios que se producen de una unidad a otra pueden resultar intrascendentes o no ser constantes a lo largo del rango de valores de la variable.

#### **Correlación**

Cuando algunas de las variables independientes analizadas están altamente correlacionadas, los resultados que se consiguen pueden no ser satisfactorios o provechosos, por esta razón debe realizarse un análisis previo univariado entre las distintas variables o factores explicativos.

### **Relación entre las variables**

Para que la regresión logística tenga un sentido muy claro y preciso, tiene que haber una relación monótona entre las variables explicativas y la de respuesta, esto quiere decir que el aumento de las unas se acompañe del aumento o la disminución aproximadamente constante de la otra variable, para todo el rango de valores estudiados.

## **CAPÍTULO II**

### **2 MARCO METODOLÓGICO**

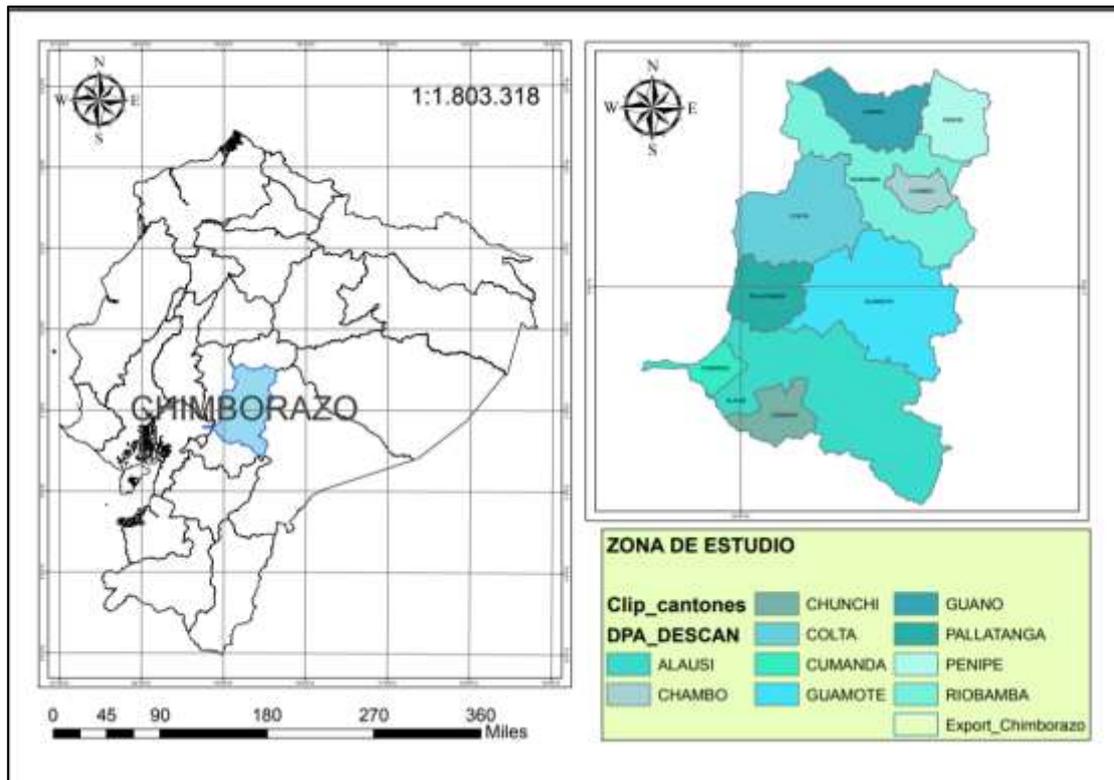
#### **2.1 Tipo y diseño de investigación**

La investigación según:

- El grado de abstracción es de tipo aplicada ya que la investigación se centra en la solución de un problema en el campo de medicina.
- La naturaleza de estudio es exploratorio puesto que se ha logrado un primer acercamiento a la problemática mediante la búsqueda de artículos relacionados, y predicativa porque mediante modelación estadística se pretende predecir fenómenos futuros.
- La naturaleza de los datos es cualitativa puesto que las variables a analizar son mudables estadísticas.
- El periodo temporal es transversal dado que no se siguió al individuo en el lapso de estudio.
- El método de estudio es inductivo ya que se busca conocer los factores influyentes en la patología a través de la información concentrada en SOLCA núcleo de Chimborazo para derivar los resultados a las diferentes casas de salud en el Ecuador.
- El nivel de profundización en el objeto de estudio es explicativo ya que se busca determinar las causas que involucran el desarrollo del cáncer de mama.

#### **2.2 Localización del área de estudio**

El presente estudio se ejecutó en el Hospital Oncológico SOLCA núcleo de la provincia de Chimborazo ubicado en la ciudad de Riobamba entre las calles Esmeraldas y Duchicela, con coordenadas latitud -1.668748 y longitud -78.663062.



**Gráfico 1-2:** Ubicación de la provincia de Chimborazo y su cantón Riobamba

Fuente: (INEC, 2010).

Realizado por: Quinatoa M., William G.,2020.

### 2.3 Población de estudio

El estudio se realizará a pacientes con cáncer de mama atendidos en la unidad oncológica SOLCA núcleo de Chimborazo durante el periodo 2017-2018.

### 2.4 Recolección de información

Los datos fueron obtenidos de las historias clínicas y ordenadores que posee el área de estadística en el Hospital Oncológico SOLCA núcleo de Chimborazo para la realización del presente estudio.

### 2.5 Identificación de variables

En el presente trabajo de investigación, las variables de interés son: peso, talla, estado civil, etnia, edad, localización de la mama, catálogo, reporte de la enfermedad, estadio, cuadrante, tamaño del tumor 1, tamaño del tumor 2, número de ganglios aislados nivel I, número de ganglios aislados de nivel II, CA 15-3 y Ki 67%. Las mismas que ayudaran al desarrollo de la investigación en beneficio de la colectividad.

## 2.6 Operacionalización de variables

A continuación, se presenta las variables proporcionadas por el Hospital Oncológico SOLCA, las cuales cumplen con los requerimientos para el desarrollo de los objetivos planteados.

**Tabla 1-2:** Operacionalización de las variables.

<b>Variable</b>	<b>Unidad de medida</b>	<b>Tipo</b>	<b>Escala</b>	<b>Descripción</b>
<b>Peso</b>	kg	Cuantitativa Continua	Intervalo	Medida que se toma con una balanza a una persona.
<b>Talla</b>	cm	Cuantitativa Continua	Intervalo	La estatura (o talla humana) designa la altura de un individuo.
<b>Estado civil</b>		Cualitativo Nominal	Nominal	Condición de una persona según el registro civil en función de si tiene o no pareja y su situación legal respecto a esto
<b>Etnia</b>		Cualitativa Nominal	Nominal	Se trata de una comunidad humana que comparte una afinidad cultural.
<b>Edad</b>	años	Cuantitativa Continua	Razón	Tiempo que ha vivido una persona u otro ser vivo contando desde su nacimiento.
<b>Localización de la mama</b>		Cualitativo Nominal	Nominal	Es la denominación del lugar en donde se encuentra la mama, esta puede ser izquierda o derecha
<b>Birads</b>		Cualitativo Nominal	Nominal	Sistema de reporte y base de datos de imágenes mamarias. Es un método que usan los radiólogos para interpretar y comunicar los resultados de la

mamografía.

<b>Reporte</b>		Cualitativo Nominal	Nominal	Es el diagnóstico de un cáncer y las características del tumor
<b>Estadio</b>		Cualitativo Nominal	Nominal	Es el sistema de estadificación, el TNM se usa para describir la mayoría de los tipos de cáncer
<b>Cuadrante</b>		Cualitativo Nominal	Nominal	Ayuda a localizar las lesiones en las mamas, se divide virtualmente en cuatro cuadrantes
<b>Tamaño del tumor 1</b>	cm	Cuantitativa Continua	Intervalo	Primera dimensión del tumor
<b>Tamaño del tumor 2</b>	cm	Cuantitativa Continua	Intervalo	Segunda dimensión del tumor
<b>Número de ganglios aislados de nivel I</b>		Cuantitativa Discreta	Razón	Número de ganglios linfáticos regionales aislados de nivel I
<b>Número de ganglios aislados de nivel II</b>		Cuantitativa Discreta	Razón	Número de ganglios linfáticos regionales aislados de nivel II
<b>CA 15-3</b>	U/ml	Cuantitativa Continua	Intervalo	Prueba de laboratorio que mide el nivel de proteína que se encuentra en las células epiteliales.

**Ki 67%**

%

Cuantitativo  
Continua

Intervalo

El Ki-67 es una manera de medir qué tan rápido crecen y se dividen las células cancerosas. Valores altos (más del 30%) para Ki-67 significan que muchas células se están dividiendo, por lo que el cáncer es probable que crezca y se propague más rápidamente.

---

**Realizado por:** Quinatoa M., William G.,2020.

## **2.7 Alcance de la investigación**

Mediante el problema planteado, la investigación tiene un alcance descriptivo y estimativo, donde el primero refleja el comportamiento de las variables cualitativas siendo las principales los factores de riesgo, factores sociales y factores pronósticos, y por último se tiene el modelo de regresión logística y el análisis implicativo.

## **2.8 Análisis de datos**

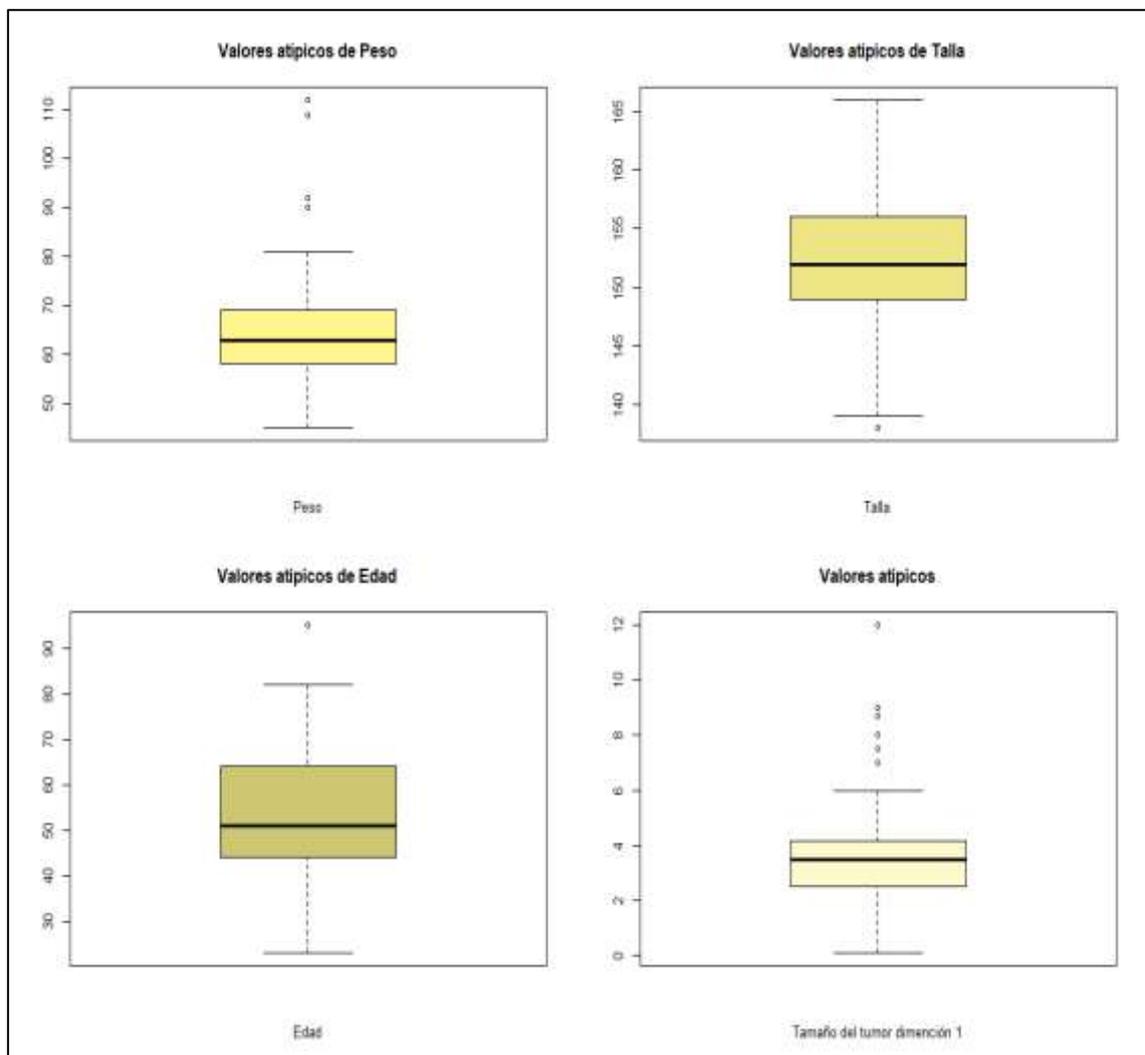
Para el presente estudio, en primera instancia se realizó un análisis exploratorio, para luego proceder con la separación de pacientes que tienen sus historias clínicas incompletas y pruebas de laboratorio no realizadas. Previo a la obtención del modelo de regresión logística se obtuvo un análisis descriptivo de cada factor y el nivel de asociación de cada uno con variable mudable clasificación del estadio. Se ejecutó un modelo de predicción a través de la regresión logística el cual para validarlo se obtuvo el estadístico de Cox y Snell y la clasificación que realiza este modelo. Para el análisis implicativo se realizó la programación correspondiente en el programa estadístico R-studio, logrando realizar el árbol de similaridad, árbol jerárquico y el diagrama implicativo.

## CAPÍTULO III

### 3 RESULTADOS Y DISCUSIÓN

#### 3.1 Análisis exploratorio de datos cuantitativos

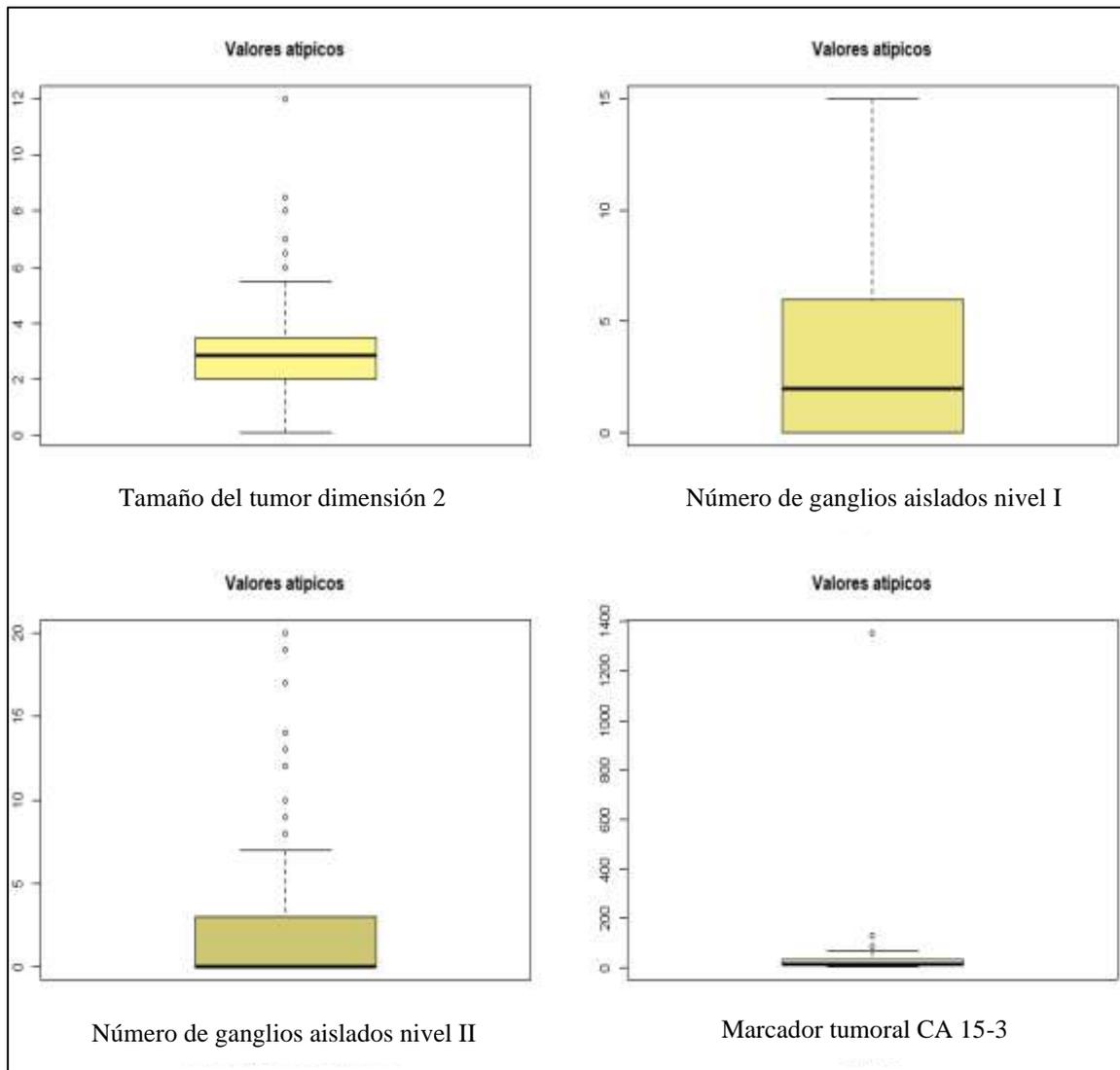
El primer paso en cualquier trabajo de investigación implica un análisis exploratorio de datos atípicos, faltantes o ausentes (NA's), ya que a priori se necesita conocer el estado de la base de datos con la que se realizará el estudio. Cabe recalcar que para este estudio se utilizó solo pacientes que se realizaron todos los exámenes necesarios para el diagnóstico de cáncer de mama.



**Gráfico 1-3:** Diagramas de caja de peso, talla, edad y tamaño del tumor dimensión 1

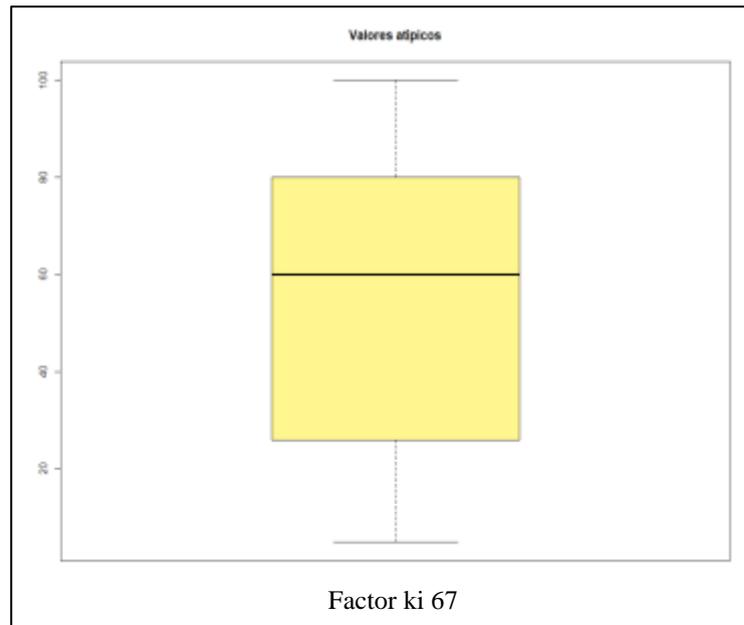
Realizado por: Quinatoa M., William G.,2020.

Para la identificación de datos atípicos se realizó diagramas de caja, en donde se observa la existencia de datos atípicos en el peso, dimensión del tumor 1, dimensión del tumor 2 y en el número de ganglios aislados de nivel II, estos datos sospechosos no fueron eliminados ya que los valores altos corresponden a pacientes con un alto grado respecto al cáncer de mama, en el gráfico 1-3 se muestra los valores atípicos de dichos pacientes.



**Gráfico 2-3:** Diagramas de caja del tumor dim2, ganglios I, II y Ca 15-3

**Realizado por:** Quinatoa M., William G.,2020.



**Gráfico 3-3:** Diagramas de caja de ki 67 (%)

**Realizado por:** Quinatoa M., William G.,2020

Se puede observar en la tabla 1-3 que existen pesos altos con valor promedio de 64 kg, dando a entender que el sobrepeso es un factor de riesgo para los pacientes que sufren esta enfermedad. Con respecto a la variable edad se obtuvo un valor promedio de 53.6 años junto con una edad mínima de 23 años y una edad máxima de 95 años.

**Tabla 1-3:** Análisis descriptivo de los factores peso, talla y edad.

	<b>Peso (kg)</b>	<b>Talla (cm)</b>	<b>Edad</b>
<b>Promedio</b>	64.0	152.1	53.6
<b>Moda</b>	60.00	155.00	48.00
<b>Mínimo</b>	45.00	138.00	23.00
<b>Máximo</b>	112.00	166.00	95.00

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

En la tabla 2-3 se observa que existen dimensiones del tumor que en promedio rondan los 3.74 cm y 3.08 cm, además de obtener valor de ganglios aislados muy altos, siendo estos 15 para el nivel I y 20 para el nivel II. Para el análisis descriptivo de las variables de laboratorio que son el CA 15-3 y el Ki 67 se obtuvieron valores muy dispersos ya que en la base de datos existen pacientes cuya enfermedad está empezando y otros que ya tienen un cáncer de mama muy

avanzado dando como resultado la obtención de valores muy pequeños y valores muy grandes respectivamente para estas variables.

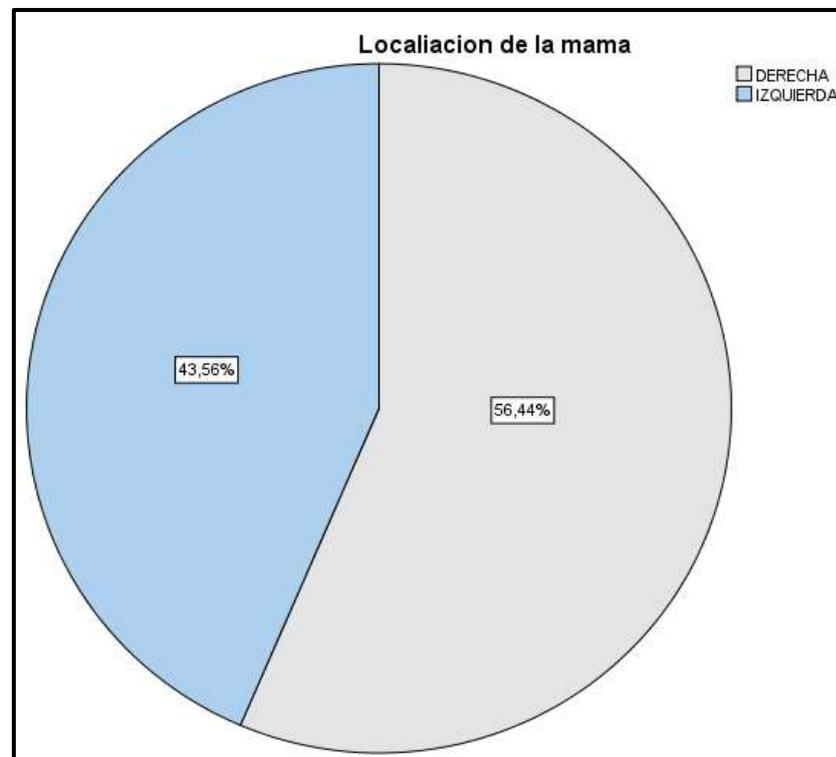
**Tabla 2-3:** Análisis de los factores pronóstico.

	Tamaño tumor dimensión 1	Tamaño tumor dimensión 2	Número de ganglios aislados Nivel I	Número de ganglios aislados Nivel II	CA 15-3 U/ml (0-26.4)	ki 67 (%)
<b>Promedio</b>	3.74	3.08	3.06	2.40	35.73	56.49
<b>Moda</b>	3.00	2.00	0.00	0.00	35.72	80.00
<b>Mínimo</b>	0.10	0.10	0.00	0.00	5.59	5.00
<b>Máximo</b>	12.00	12.00	15.00	20.00	1352.00	100.00

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

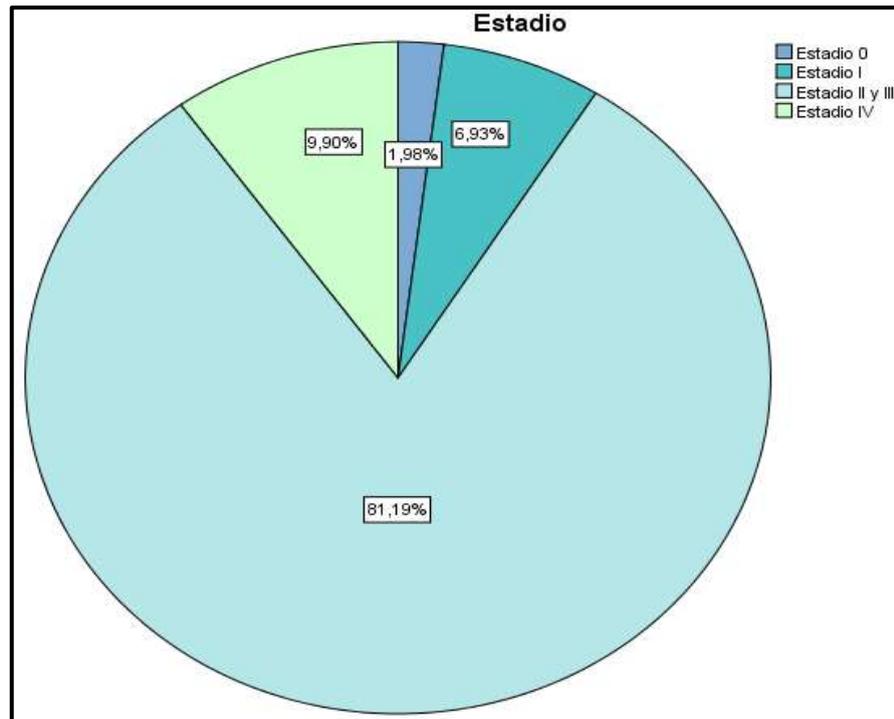
**Realizado por:** Quinatoa M., William G.,2020

### 3.2 Análisis exploratorio de datos cualitativos



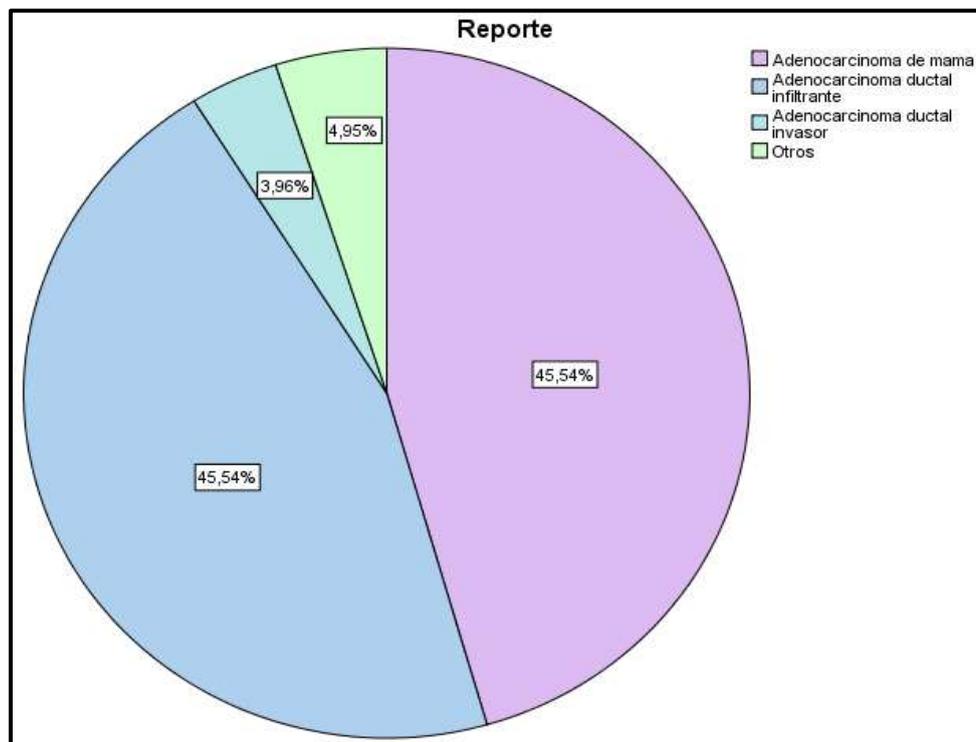
**Gráfico 4-3:** Representación de la localización de la mama

**Realizado por:** Quinatoa M., William G.,2020



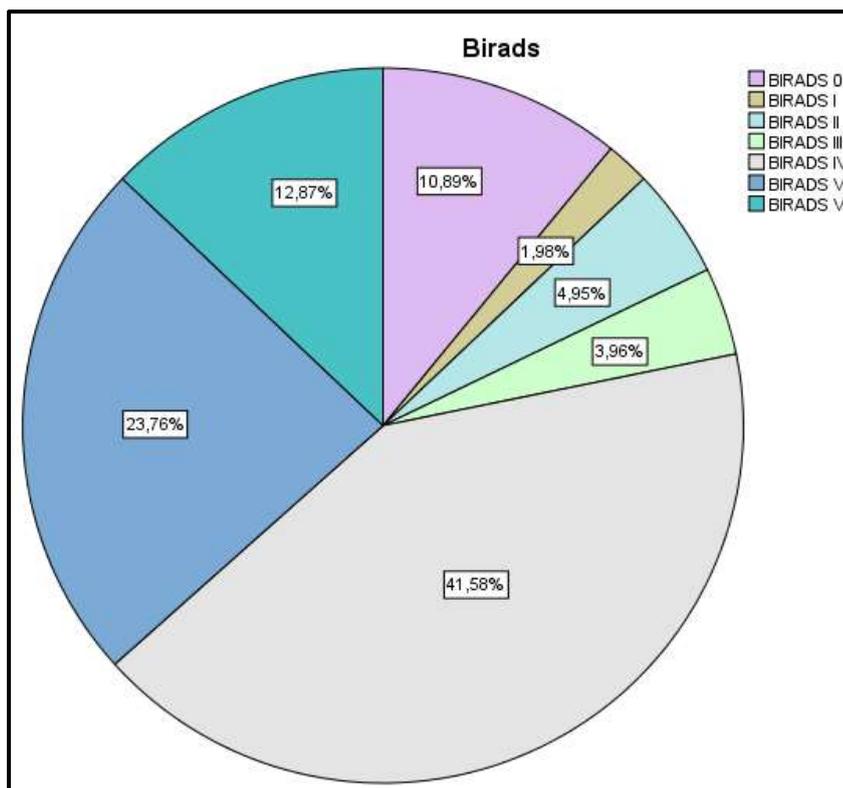
**Gráfico 5-3:** Representación del estadio

Realizado por: Quinatoa M., William G.,2020



**Gráfico 6-3:** Representación del reporte de tipo de cáncer

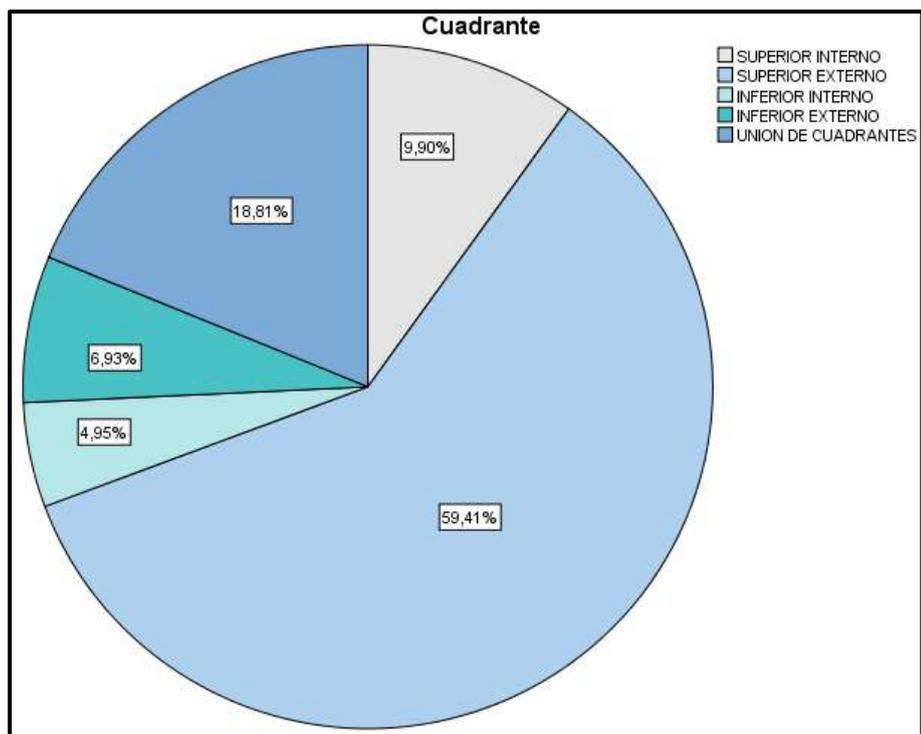
Realizado por: Quinatoa M., William G.,2020



**Gráfico 7-3:** Representación de la variable Birads

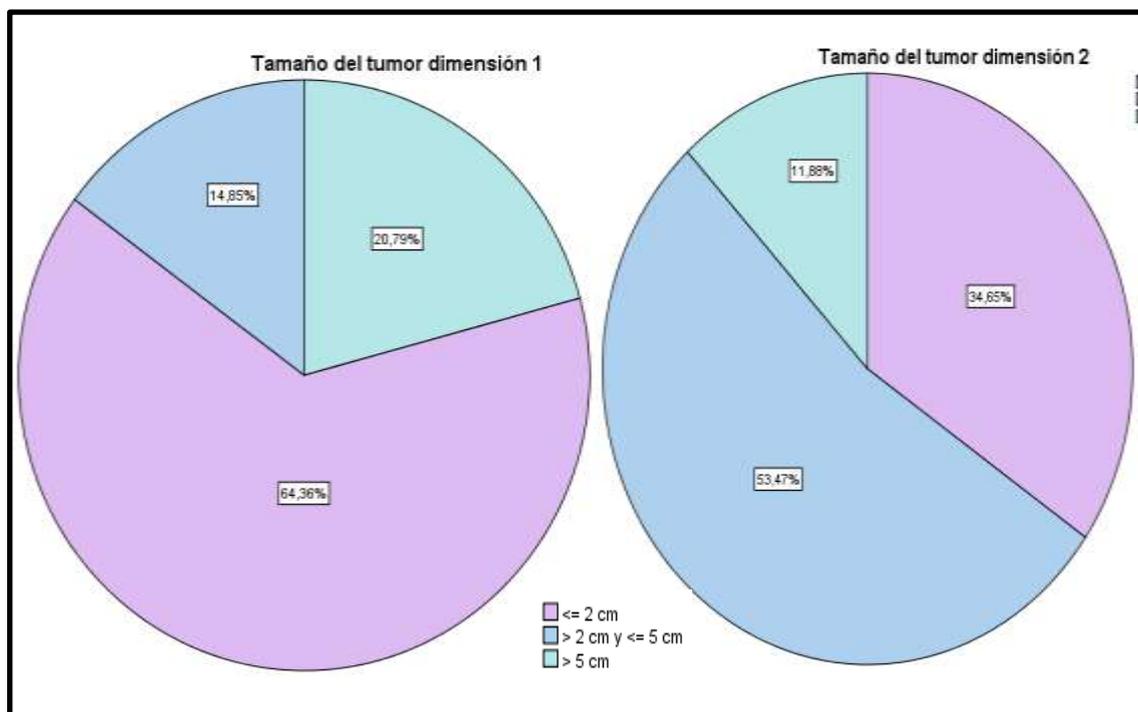
**Realizado por:** Quinatoa M., William G.,2020

De acuerdo con el gráfico 4-3 se observa que la localización del tumor tiene una ligera tendencia a ubicarse en la mama derecha con una aparición del 56.44% con respecto a los pacientes en estudio, en la gráfica 7-3 se muestra que a la variable BIRADS IV se diagnostica con mayor frecuencia obteniendo una aparición del 41.58% seguido del BIRADS V con un 23.76%. Con respecto a la gráfica 6-3 se muestran que los 2 reportes de cáncer con mayor porcentaje de aparición en los pacientes con cáncer de mama son el Adenocarcinoma de mama y el Adenocarcinoma de mama infiltrante con un 45.54% para cada una, y con respecto a la gráfica 5-3 que comprende al estadio del tumor se puede observar que existe un mayor número de casos con Estadio II y III obteniendo el 81.19%, dando a entender que los tumores son grades y han crecido más profundamente en los tejidos adyacentes.



**Gráfico 8-3:** Representación de la variable cuadrante

Realizado por: Quinatoa M., William G.,2020

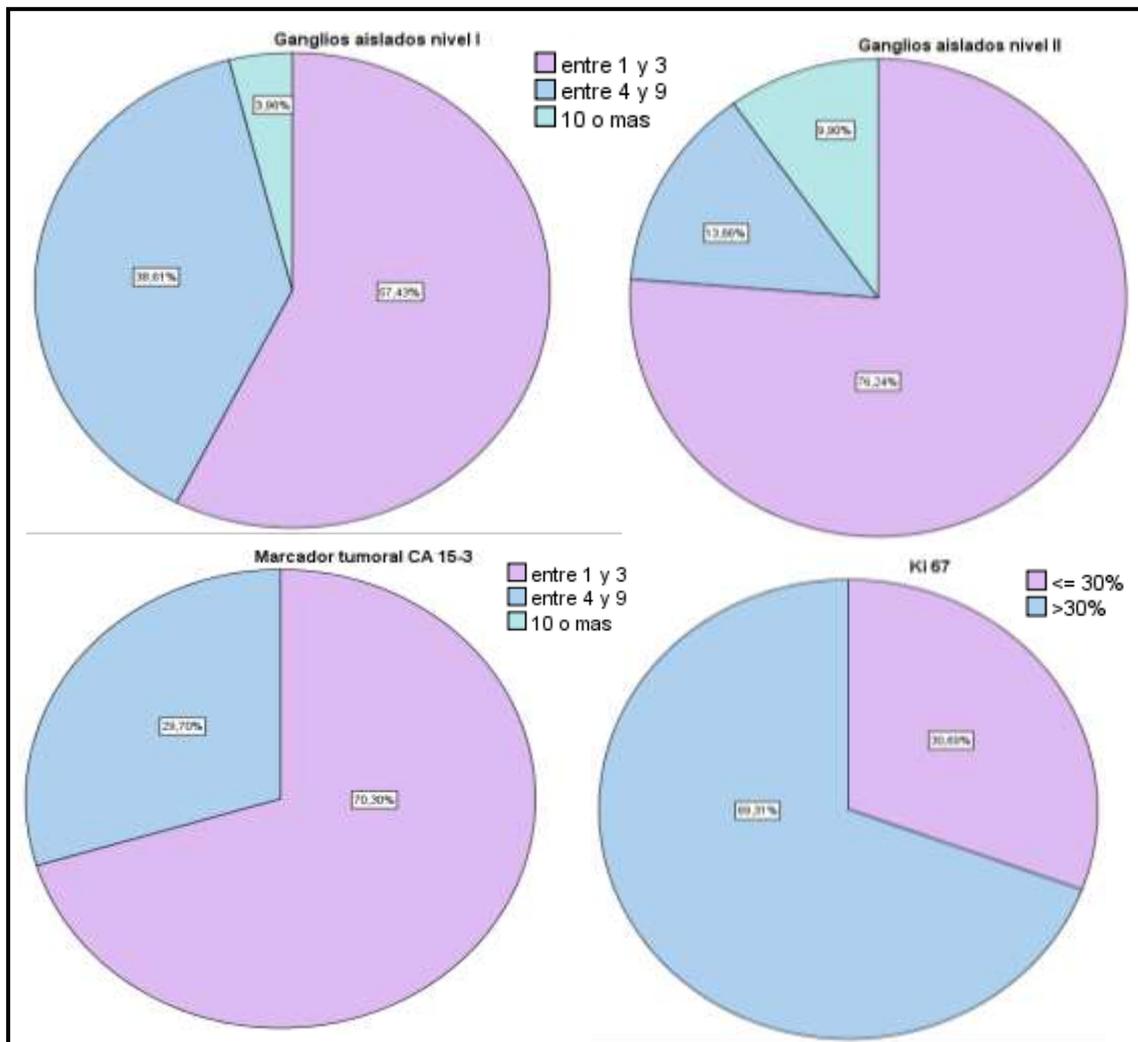


**Gráfico 9-3:** Representación de la variable tamaño del tumor

Realizado por: Quinatoa M., William G.,2020

De acuerdo con el gráfico 8-3 se muestra que el cuadro superior externo es el que aloja con mayor frecuencia al tumor con 59.41% de presencia en esta zona, y respecto al gráfico 9-3

correspondiente a las variables dimensión del tamaño del tumor 1 y 2 se observa que en el primero las medidas con menores a 2cm, con un porcentaje de valores encontrados al 64.36%, contrario a la segunda variable que tiene medidas que oscilan entre valores mayores a 2 cm pero menores o iguales a 5 cm, esto con un porcentaje de valores encontrados del 53.47%.



**Gráfico 10-3:** Representación de ganglios I y II, CA 15-3 y Ki 67

**Realizado por:** Quinatoa M., William G., 2020

De acuerdo al gráfico 10-3 se observa que el número de ganglios aislados oscilan entre 1 y 3 tanto para el nivel I y nivel II con el 57.43% y 76.24% de casos para cada uno de los niveles respectivamente, con respecto al marcador tumoral CA 15-3 se puede apreciar que de los 101 pacientes en estudio el 70.3% tienen un valor normal y solo el 29.7% lo tienen elevado, caso contrario para la variable Ki 67 que solo el 30.69% de los pacientes tienen valores menores o iguales a 30 (rango normal) y un 69.31% de individuos con valores mayores a 30.

Los factores de riesgo sociales son: estado civil, nivel de instrucción y ocupación, obteniendo el conteo de cada uno en la tabla 3-3.

**Tabla 3-3:** Análisis descriptivo de los factores sociales

<b>Estado civil</b>		<b>Ocupación</b>	
Casada	46	Abogada	1
Divorciada	14	Agricultura	2
Soltera	22	Ama de casa	57
Unión libre	4	Auxiliar de contabilidad	1
Viuda	6	Comerciante	10
<b>Instrucción</b>		Diseño de modas	1
Bachillerato	2	Docente	3
Básica	3	Empleada publica	3
Inicial	1	Enfermera	2
Ninguna	6	Estilista	1
Primaria	21	Estudiante	1
Secundaria en curso	1	Jubilada	6
Secundaria	29	Profesora	3
Superior	29	Secretaria	1

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

En la tabla 3-3 se observa que existe una frecuencia elevada de pacientes con estado civil casada, una instrucción con cantidad igual en el nivel de secundaria y superior, con respecto a la ocupación de los pacientes, el término ama de casa es el más frecuente con 57 casos.

### 3.3 Asociación entre los factores con la variable dependiente

En la tabla 4-3, se muestra la realización de la prueba Chi-cuadrado para medir el grado de asociación de la localización de la mama con el estadio, dando como resultado que esta variable no está asociada con el estadio, con un valor  $p > 0.05$ , aceptando  $H_0$ , para la siguiente hipótesis:  $H_0$ : La clasificación de estadio es independiente de la localización de la mama

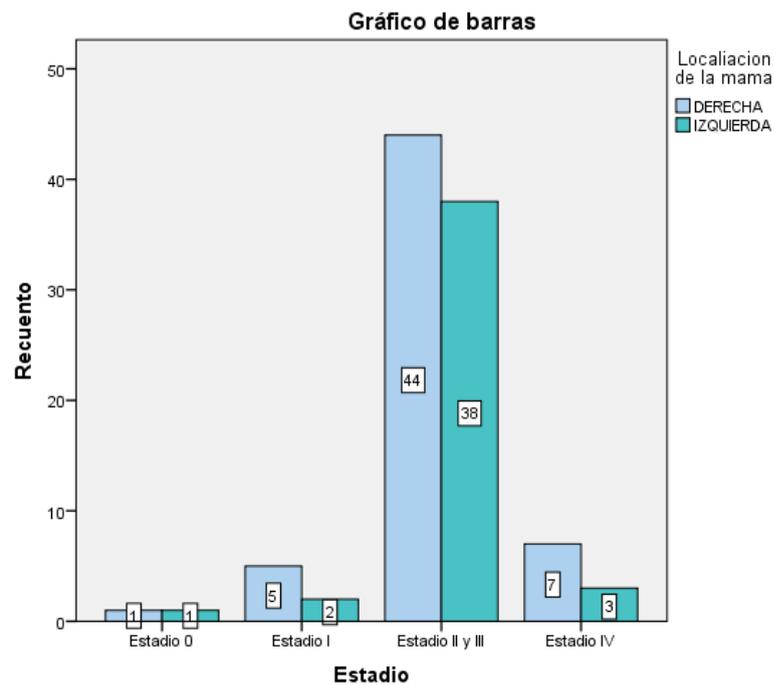
H1: La clasificación de estadio es dependiente a la localización de la mama

**Tabla 4-3:** Prueba  $\chi^2$  Estadio vs Localización de la mama

Pruebas de chi-cuadrado			
	Valor	df	Significación asintótica (bilateral)
Chi-cuadrado de Pearson	1,679	3,000	0,642
Razón de verosimilitud	1,735	3,000	0,629
Asociación lineal por lineal	0,051	1,000	0,821

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020



**Gráfico 11-3:** Asociación entre Estadio y localización de la mama

**Realizado por:** Quinatoa M., William G.,2020

De acuerdo al gráfico 11-3 se muestran una mayor frecuencia de aparición del estadio II y III en las mamas derecha e izquierda mientras que los otros tipos de estadio están con menor frecuencia en la mama izquierda.

En la tabla 5-3, se muestra la realización de la prueba Chi-cuadrado para medir el grado de

asociación del factor BIRADS, dando como resultado que esta variable si está asociada con el estadio con un valor  $p < 0.05$ , rechazando  $H_0$ , para la siguiente hipótesis:

$H_0$ : La clasificación de estadio es independiente al factor Birads

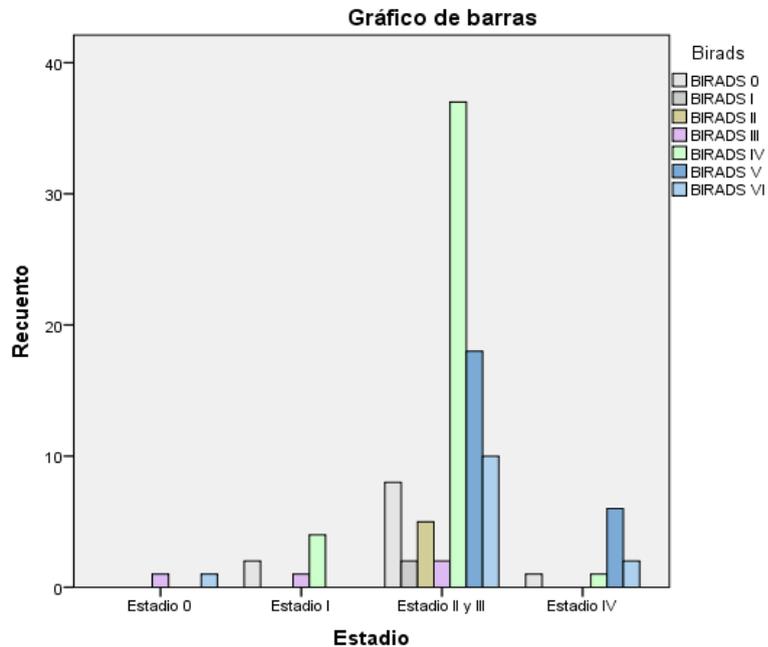
$H_1$ : La clasificación de estadio es dependiente al factor Birads

**Tabla 5-3:** Prueba  $X^2$  Estadio vs Birads

Pruebas de chi-cuadrado			
	Valor	df	Significación asintótica (bilateral)
Chi-cuadrado de Pearson	32,547	18.000	0.019
Razón de verosimilitud	27.935	18.000	0.063
Asociación lineal por lineal	2.266	1.000	0.132

Fuente: Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

Realizado por: Quinatoa M., William G.,2020



**Gráfico 12-3:** Asociación entre Estadio y Birads

Realizado por: Quinatoa M., William G.,2020

De acuerdo con el gráfico 12-3 se muestran, que en los estadios II y III hay una mayor presencia de BIRADS IV y BIRADS V, mientras que en otros estadios la presencia del BIRADS IV y V es muy poco o casi nula.

En la tabla 6-3, se muestra la realización de la prueba Chi-cuadrado para medir el grado de asociación del factor reporte, dando como resultado que esta variable no está asociada con el estadio con un valor  $p > 0.05$ , aceptando  $H_0$ , para la siguiente hipótesis:

$H_0$ : La clasificación de estadio es independiente al reporte de tipo de cáncer

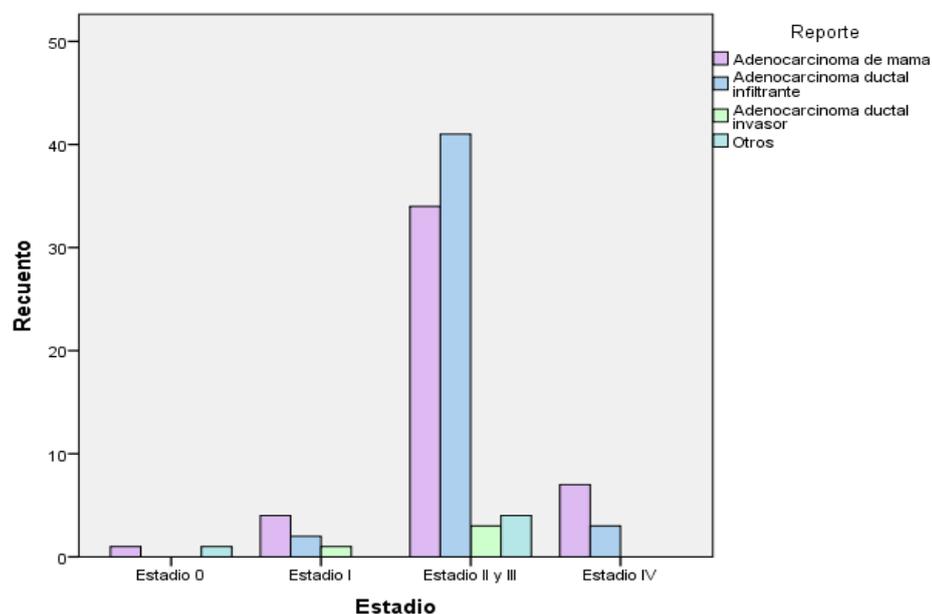
$H_1$ : La clasificación de estadio es dependiente al reporte de tipo de cáncer

**Tabla 6-3:** Prueba  $X^2$  Estadio vs Reporte

Pruebas de chi-cuadrado			
	Valor	df	Significación asintótica (bilateral)
<b>Chi-cuadrado de Pearson</b>	15,490	9.000	0.078
<b>Razón de verosimilitud</b>	11.515	9.000	0.242
<b>Asociación lineal por lineal</b>	2.667	1.000	0.102

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020



**Gráfico 13-3:** Asociación entre Estadio y Reporte

**Realizado por:** Quinatoa M., William G.,2020

De acuerdo al gráfico 13-3 se muestran, que los reportes que más se replican en los pacientes son adenocarcinoma de mama y adenocarcinoma ductal infiltrante para el estadio II y III.

En la tabla 7-3, se muestra la realización de la prueba Chi-cuadrado para medir el grado de asociación del factor reporte, dando como resultado que esta variable no está asociada con el estadio con un valor  $p > 0.05$ , aceptando  $H_0$ , para la siguiente hipótesis:

$H_0$ : La clasificación de estadio es independiente al cuadrante donde se encuentra el tumor

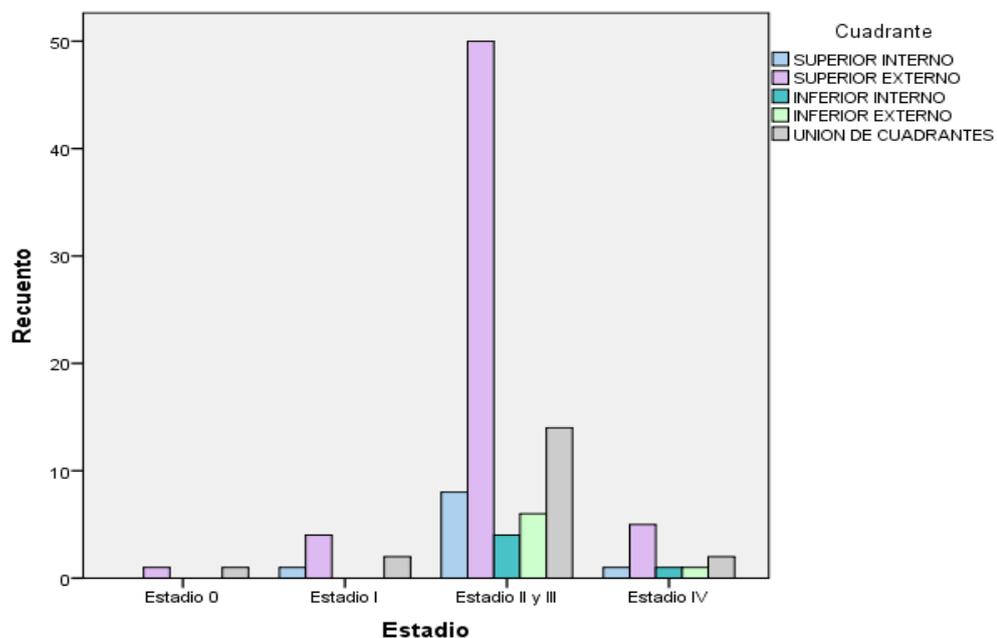
$H_1$ : La clasificación de estadio es dependiente al cuadrante donde se encuentra el tumor

**Tabla 7-3:** Prueba  $X^2$  Estadio vs Cuadrante

Pruebas de chi-cuadrado			
	Valor	df	Significación asintótica (bilateral)
<b>Chi-cuadrado de Pearson</b>	3,821	12.000	0.986
<b>Razón de verosimilitud</b>	4.575	12.000	0.971
<b>Asociación lineal por lineal</b>	0.129	1.000	0.720

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020



**Gráfico 14-3:** Asociación entre Estadio y Cuadrante

**Realizado por:** Quinatoa M., William G.,2020

De acuerdo con el gráfico 14-3 se muestran, que el cuadrante superior externo es el que más frecuencias tiene en los pacientes con cáncer de mama.

En la tabla 8-3 y 9-3, se muestra la realización de la prueba Chi-cuadrado para medir el grado de

asociación entre el factor tamaño del tumor 1 y tamaño del tumor 2, dando como resultado que estas variables no están asociadas con el estadio con un valor  $p > 0.05$ , aceptando  $H_0$ , para la siguiente hipótesis: (esta hipótesis es igual para el tamaño del tumor en dimensión 2):

$H_0$ : La clasificación de estadio es independiente al tamaño del tumor en dimensión 1

$H_1$ : La clasificación de estadio es dependiente al tamaño del tumor en dimensión 1

**Tabla 8-3:** Prueba  $X^2$  Estadio vs Tamaño 1

<b>Pruebas de chi-cuadrado Estadio vs Tamaño del tumor de dimensión 1</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
<b>Chi-cuadrado de Pearson</b>	10,942	6.000	0.090
<b>Razón de verosimilitud</b>	12.391	6.000	0.054
<b>Asociación lineal por lineal</b>	8.974	1.000	0.003

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

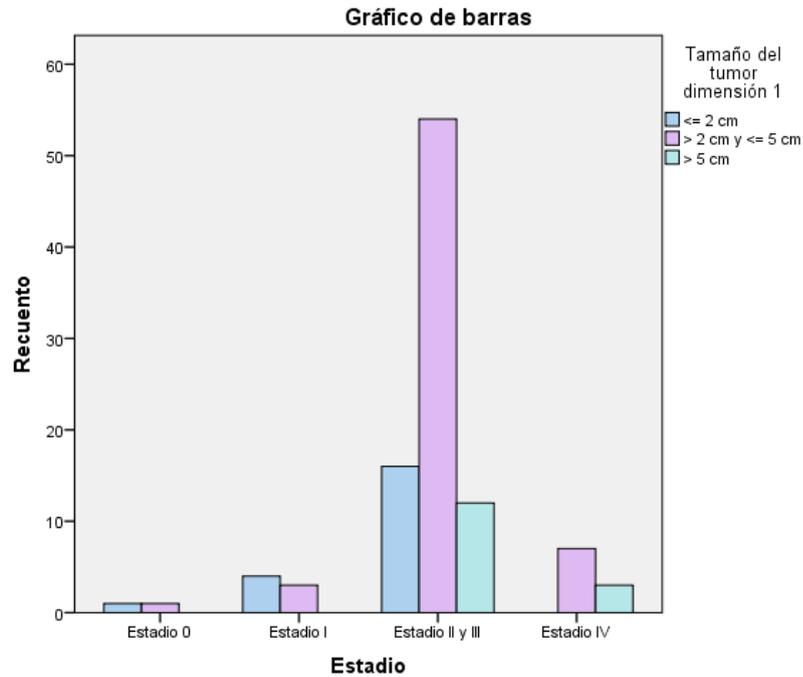
**Realizado por:** Quinatoa M., William G.,2020

**Tabla 9-3:** Prueba  $X^2$  Estadio vs Tamaño 2

<b>Pruebas de chi-cuadrado entre Estadio vs Tamaño del tumor de dimensión 2</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
<b>Chi-cuadrado de Pearson</b>	5,472	6.000	0.485
<b>Razón de verosimilitud</b>	5.879	6.000	0.437
<b>Asociación lineal por lineal</b>	1.363	1.000	0.243

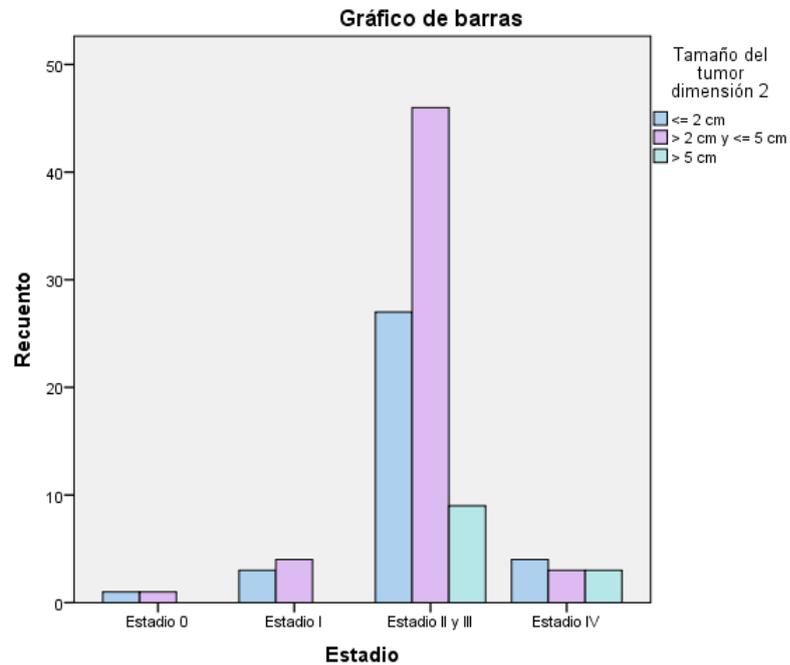
**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020



**Gráfico 15-3:** Asociación entre Estadio y dimensión del tumor 1

Realizado por: Quinatoa M., William G.,2020



**Gráfico 16-3:** Asociación entre Estadio y dimensión del tumor 2

Realizado por: Quinatoa M., William G.,2020

De acuerdo con las gráficas 15-3 y 16-3 se observa que los tumores con dimensiones entre 2 cm y 5 cm son las que más se presentan en los estadios II y III.

En la tabla 15-3 y 16-3, se muestra la realización de la prueba Chi-cuadrado para medir el grado de asociación del factor número de ganglios de nivel I y II 2, dando como resultado que el número de ganglios de nivel I no está asociada con el estadio con un valor  $p > 0.05$ , y el número de ganglios II si está asociada con el estadio con un valor  $p < 0.05$ , aceptando  $H_0$  para el número de ganglios I y Rechazando  $H_0$  para el número de ganglios II, para las siguiente hipótesis:

$H_0$ : La clasificación de estadio es independiente del número de ganglios aislados de nivel I

$H_1$ : La clasificación de estadio es dependiente al número de ganglios aislados de nivel II

**Tabla 10-3: Prueba  $X^2$  Estadio vs Ganglios I**

<b>Pruebas de chi-cuadrado entre Estadio vs Número de ganglios aislado de nivel I</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
<b>Chi-cuadrado de Pearson</b>	3,629	6.000	0.727
<b>Razón de verosimilitud</b>	4.399	6.000	0.623
<b>Asociación lineal por lineal</b>	2.408	1.000	0.121

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

$H_0$ : La clasificación de estadio es independiente del número de ganglios aislados de nivel II

$H_1$ : La clasificación de estadio es dependiente del número de ganglios aislados de nivel II

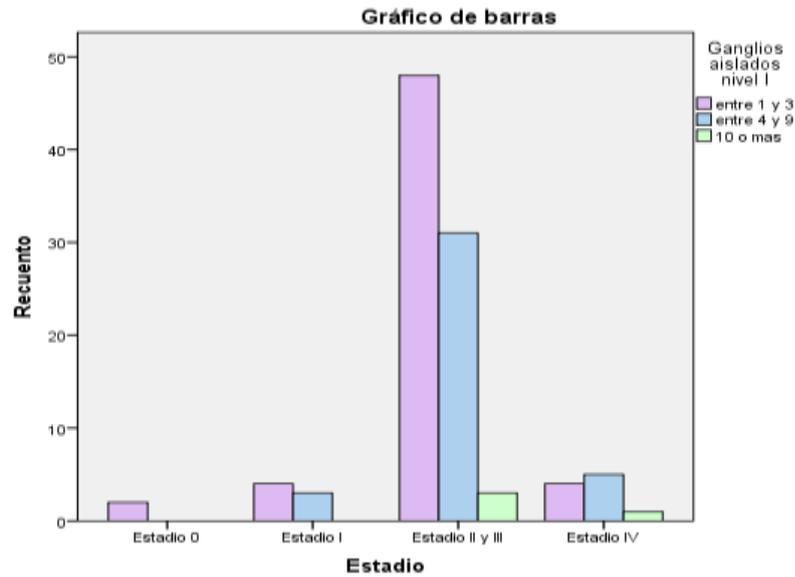
**Tabla 11-3: Prueba  $X^2$  Estadio vs Ganglios II**

<b>Pruebas de chi-cuadrado entre Estadio vs Número de ganglios aislados de nivel II</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
<b>Chi-cuadrado de Pearson</b>	22,343	6.000	0.001
<b>Razón de verosimilitud</b>	16.914	6.000	0.010
<b>Asociación lineal por lineal</b>	12.085	1.000	0.001

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

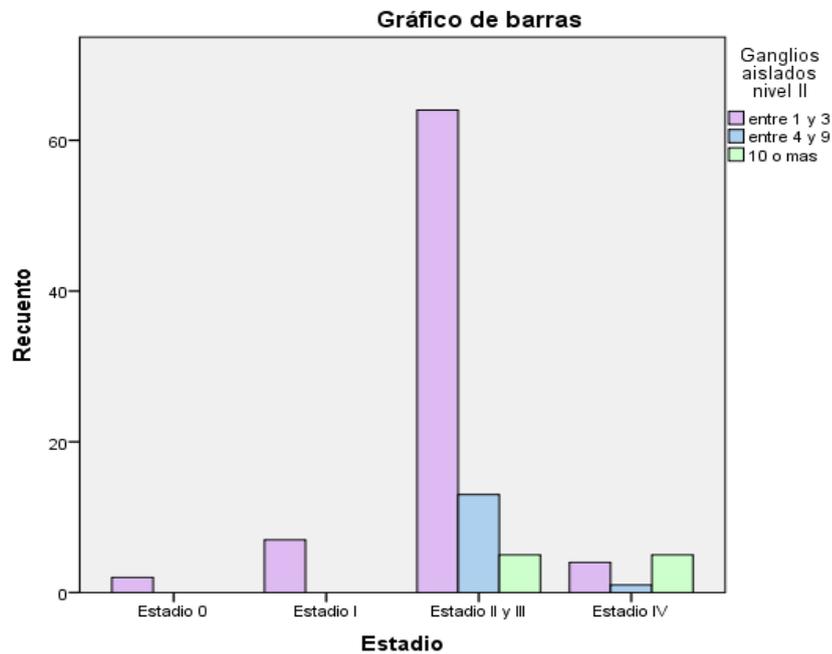
**Realizado por:** Quinatoa M., William G.,2020

De acuerdo con las gráficas 13-3 y 14-3 se muestran, que el número de ganglios aislados son generalmente entre 1 y 3 en los estadios II y III.



**Gráfico 17-3:** Asociación entre Estadio y ganglios aislados nivel I

Realizado por: Quinatoa M., William G.,2020



**Gráfico 18-3:** Asociación entre Estadio y ganglios aislados II

Realizado por: Quinatoa M., William G.,2020

De acuerdo a la tabla 17-3, se muestra la realización de la prueba Chi-cuadrado para medir el grado de asociación del factor marcador tumoral CA 15-3, dando como resultado que esta variable no está asociada con el estadio con un valor  $p > 0.05$ , aceptando  $H_0$ , para la siguiente hipótesis:

H0: La clasificación de estadio es independiente del marcador tumoral CA 15-3

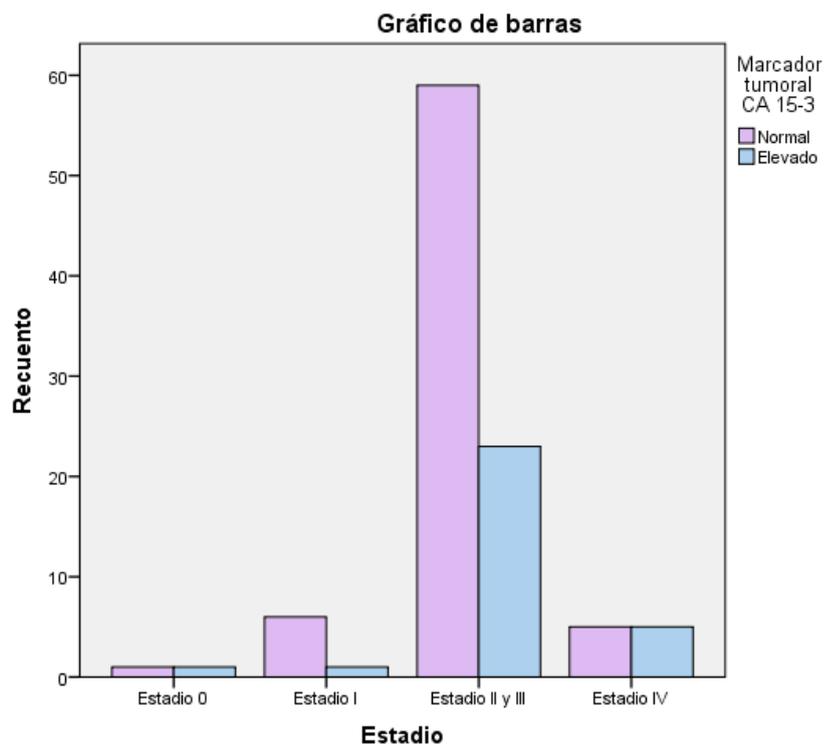
H1: La clasificación de estadio es dependiente del marcador tumoral CA 15-3

**Tabla 12-3:** Prueba  $X^2$  Estadio vs CA 15-3

Pruebas de chi-cuadrado			
	Valor	df	Significación asintótica (bilateral)
<b>Chi-cuadrado de Pearson</b>	3,272	3.000	0.352
<b>Razón de verosimilitud</b>	3.185	3.000	0.364
<b>Asociación lineal por lineal</b>	1.001	1.000	0.317

Fuente: Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

Realizado por: Quinatoa M., William G.,2020



**Gráfico 19-3:** Asociación entre Estadio y CA 15-3

Realizado por: Quinatoa M., William G.,2020

De acuerdo al gráfico 19-3 se observa que el marcador tumoral CA 15-3 con frecuencia se encuentra el rango normal, es decir estos valores se encuentran entre 0 y 26.4 u/ml para los estadios I, II, y III.

En la tabla 13-3, se muestra la realización de la prueba Chi-cuadrado para medir el grado de asociación del factor Ki 67, dando como resultado que esta variable no está asociada con el estadio con un valor  $p > 0.05$ , aceptando H0, para la siguiente hipótesis:

H0: La clasificación de estadio es independiente del Ki 67

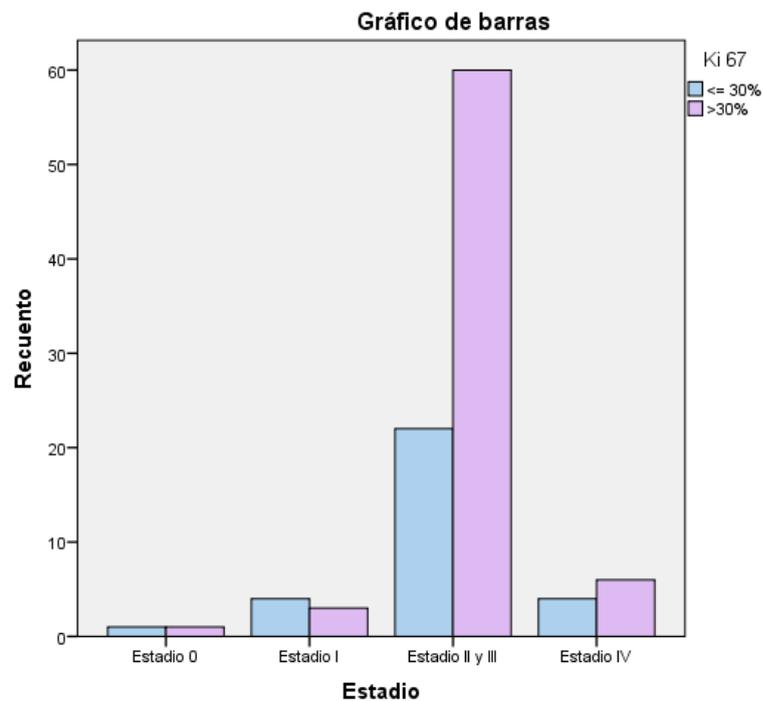
H1: La clasificación de estadio es dependiente del ki 67

**Tabla 13-3:** Prueba  $X^2$  Estadio vs Ki 67

<b>Pruebas de chi-cuadrado</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
<b>Chi-cuadrado de Pearson</b>	3,635	3.000	0.304
<b>Razón de verosimilitud</b>	3.390	3.000	0.335
<b>Asociación lineal por lineal</b>	0.534	1.000	0.465

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020



**Gráfico 20-3:** Asociación entre Estadio y Ki 67

**Realizado por:** Quinatoa M., William G.,2020

De acuerdo al gráfico 16-3 se muestra, que los exámenes de Ki 67 generalmente muestran frecuencia con valores mayores a 30%, y se nota particularmente en los estadios II, III y IV.

**Tabla 14-3:** Asociación de los factores de riesgo sociales

<b>Estadio vs Instrucción</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
<b>Chi-cuadrado de Pearson</b>	5,498	15.000	0.987
<b>Razón de verosimilitud</b>	6.374	15.000	0.973
<b>Asociación lineal por lineal</b>	1.910	1.000	0.167
<b>N de casos válidos</b>	92.000		
<b>Estadio vs Ocupación</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
<b>Chi-cuadrado de Pearson</b>	7.971	6.000	0.240
<b>Razón de verosimilitud</b>	8.214	6.000	0.223
<b>Asociación lineal por lineal</b>	5.350	1.000	0.021
<b>N de casos válidos</b>	92.000		
<b>Estadio vs Estado Civil</b>			
	<b>Valor</b>	<b>df</b>	<b>Significación asintótica (bilateral)</b>
<b>Chi-cuadrado de Pearson</b>	6.651	12.000	0.880
<b>Razón de verosimilitud</b>	7.135	12.000	0.849
<b>Asociación lineal por lineal</b>	0.687	1.000	0.407
<b>N de casos válidos</b>	92		

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

En la tabla 14-3 se observa que los factores de riesgo sociales, y social maternos como la instrucción, ocupación y estado civil no son significativos para la mortalidad de cáncer de mama.

**Tabla 15-3:** Pruebas no paramétricas en factores pronósticos

<b>Estadísticos de prueba</b>						
	<b>Estadio</b>	<b>Localización de la mama</b>	<b>Birads</b>	<b>Reporte</b>	<b>Cuadrante</b>	<b>Tamaño del tumor dimensión 1</b>
<b>Chi-cuadrado</b>	171,356	1,673	84,396	68,228	103,703	44,277
<b>gl</b>	3.000	1.000	6.000	3.000	4.000	2.000
<b>Sig. asintótica</b>	0.000	0.196	0.000	0.000	0.000	0.000
	<b>Tamaño del tumor dimensión 2</b>	<b>Ganglios aislados nivel I</b>	<b>Ganglios aislados nivel II</b>	<b>Marcador tumoral CA 15-3</b>	<b>Ki 67</b>	
<b>Chi-cuadrado</b>	26,277	44,574	83,901	16,644	15,059	
<b>gl</b>	2.000	2.000	2.000	1.000	1.000	
<b>Sig. asintótica</b>	0.000	0.000	0.000	0.000	0.000	

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G., 2020

En la tabla 15-3 se muestra una comparación de los factores pronóstico a través de la prueba no paramétrica Chi-cuadrado, la cual al ser analizada de manera univariante se obtiene que las variables estadio, Birads, reporte de cáncer, cuadrante, tamaño del tumor en dimensiones 1 y 2, número de ganglios aislados de nivel I y II, el marcador tumoral Ca 15-3 y el Ki 67 presentan significancia estadística con  $p < 0.05$ , mientras que la localización de mama no mostro significancia estadística con  $p > 0.05$ .

### 3.4 Estimación del modelo de Regresión Logística

#### 3.4.1 Estimación de parámetros

El modelo fue estimado empleando SPSS y haciendo uso del procedimiento de selección de variables Wald, cuyos resultados se proporcionan en la tabla 13-3, el cual se observan los coeficientes estimados, estadístico de Wald, el valor  $p$ , el Exp(B) y los intervalos de confianza de los Exp(B).

**Tabla 16-3:** Estimación del modelo de predicción

Estimaciones de parámetros para el modelo de Regresión Logística									
Estadio	B	Error estándar	Wald	gl	Sig.	Exp(B)	95% de intervalo de confianza para Exp(B)		
							Límite inferior	Límite superior	
<b>Estadio 0</b>	Intersección	566.452	5752.757	0.010	1.000	0.922			
	n_tam_1	-7.130	302.013	0.001	1.000	0.981	0.001	6.754E-261	9.49E+253
	n_ganglios_1	-25.584	484.518	0.003	1.000	0.958	7.745E-12	0.000	
	n_ganglios_2	5.965	599.887	0.000	1.000	0.992	389.724	0.000	
	n_CA	-0.498	21.551	0.001	1.000	0.982	0.608	2.752E-19	1.34E+18
	n_Ki	-0.243	113.401	0.000	1.000	0.998	0.784	2.331E-97	2.64E+96
	n_tam_2	4.412	173.547	0.001	1.000	0.980	82.462	1.559E-146	4.36E+149
	[Reporte=1.00]	-239.440	6314.741	0.001	1.000	0.970	1.029E-104	0.000	
	[Reporte=2.00]	-156.481	5482.641	0.001	1.000	0.977	1.099E-68	0.000	
	[Reporte=3.00]	5894.899	19686.219	0.090	1.000	0.765		0.000	
	[Reporte=4.00]	0.000			0.000				
	[Localización de la mama=1.00]	-48.377	4117.561	0.000	1.000	0.991	9.778E-22	0.000	
	[Localización de la mama=2.00]	0.000			0.000				

	[Cuadrante=1.00]	-177.892	14270.525	0.000	1.000	0.990	5.528E-78	0.000	
	[Cuadrante=2.00]	125.351	6753.407	0.000	1.000	0.985	2.75E+54	0.000	
	[Cuadrante=3.00]	114.144	12455.852	0.000	1.000	0.993	3.73E+49	0.000	
	[Cuadrante=4.00]	-56.556	8297.269	0.000	1.000	0.995	2.741E-25	0.000	
	[Cuadrante=5.00]	0.000			0.000				
	[Birads=1.00]	21.084	11609.919	0.000	1.000	0.999		0.000	
	[Birads=2.00]	-104.893	0.000		1.000		2.791E-46	2.791E-46	2.791E-46
	[Birads=3.00]	396.486	17247.578	0.001	1.000	0.982	1.56E+172	0.000	
	[Birads=4.00]	-268.934	4625.130	0.003	1.000	0.954	1.598E-117	0.000	
	[Birads=5.00]	154.659	4180.371	0.001	1.000	0.970	1.47E+67	0.000	
	[Birads=6.00]	-45.088	5237.488	0.000	1.000	0.993	2.621E-20	0.000	
	[Birads=7.00]	0.000			0.000				
<b>Estadio I</b>	Intersección	-342.940	7043.835	0.002	1.000	0.961			
	n_tam_1	-13.544	35.152	0.148	1.000	0.700	1.312E-06	1.572E-36	1.09E+24
	n_ganglios_1	-71.414	179.986	0.157	1.000	0.692	9.666E-32	6.036E-185	1.55E+122
	n_ganglios_2	-18.823	87.724	0.046	1.000	0.830	6.688E-09	1.427E-83	3.13E+66
	n_CA	-0.591	2.083	0.080	1.000	0.777	0.554	0.009	32.873

n_Ki	-7.841	22.971	0.117	1.000	0.733	0.000	1.099E-23	1.41E+16
n_tam_2	15.888	42.983	0.137	1.000	0.712	7941620.602	2.055E-30	3.07E+43
[Reporte=1.00]	407.828	6197.851	0.004	1.000	0.948	1.31E+177	0.000	
[Reporte=2.00]	425.038	6134.630	0.005	1.000	0.945	3.91E+184	0.000	
[Reporte=3.00]	6834.815	19725.342	0.120	1.000	0.729		0.000	
[Reporte=4.00]	0.000			0.000				
[Localización de la mama=1.00]	509.308	2065.448	0.061	1.000	0.805	1.55E+221	0.000	
[Localización de la mama=2.00]	0.000			0.000				
[Cuadrante=1.00]	315.516	1568.007	0.040	1.000	0.841	1.06E+137	0.000	
[Cuadrante=2.00]	-239.827	1687.077	0.020	1.000	0.887	6.991E-105	0.000	
[Cuadrante=3.00]	-354.778	3292.227	0.012	1.000	0.914	8.357E-155	0.000	
[Cuadrante=4.00]	-726.876	7158.225	0.010	1.000	0.919	0.000	0.000	
[Cuadrante=5.00]	0.000			0.000				
[Birads=1.00]	523.872	1919.936	0.074	1.000	0.785	3.27E+227	0.000	
[Birads=2.00]	362.006	8413.172	0.002	1.000	0.966	1.65E+157	0.000	

	[Birads=3.00]	-17.391	4112.072	0.000	1.000	0.997	2.799E-08	0.000	
	[Birads=4.00]	54.236	2998.527	0.000	1.000	0.986	#####	0.000	
	[Birads=5.00]	500.976	1769.345	0.080	1.000	0.777	3.73E+217	0.000	
	[Birads=6.00]	-433.985	1778.731	0.060	1.000	0.807	3.331E-189	0.000	
	[Birads=7.00]	0.000			0.000				
<b>Estadio II y III</b>	Intersección	540.667	4336.890	0.016	1.000	0.901			
	n_tam_1	-7.298	23.416	0.097	1.000	0.755	0.001	7.922E-24	5.79E+16
	n_ganglios_1	-23.691	72.925	0.106	1.000	0.745	5.140E-11	4.334E-73	6.10E+51
	n_ganglios_2	5.522	31.361	0.031	1.000	0.860	250.241	5.055E-25	1.24E+29
	n_CA	-0.476	1.439	0.109	1.000	0.741	0.621	0.037	10.423
	n_Ki	0.102	4.878	0.000	1.000	0.983	1.107	7.801E-05	15706.614
	n_tam_2	4.280	15.619	0.075	1.000	0.784	72.253	3.665E-12	1.42E+15
	[Reporte=1.00]	-230.191	3870.084	0.004	1.000	0.953	1.070E-100	0.000	
	[Reporte=2.00]	-116.361	3836.670	0.001	1.000	0.976	2.918E-51	0.000	
	[Reporte=3.00]	5922.060	18569.391	0.102	1.000	0.750		0.000	
	[Reporte=4.00]	0.000			0.000				
	[Localización de la mama=1.00]	-40.958	272.256	0.023	1.000	0.880	1.631E-18	2.937E- 250	9.05E+213

[Localización de la mama=2.00]	0.000			0.000					
[Cuadrante=1.00]	-188.643	855.227	0.049	1.000	0.825	1.184E-82	0.000		
[Cuadrante=2.00]	125.052	790.090	0.025	1.000	0.874	2.04E+54	0.000		
[Cuadrante=3.00]	107.695	836.433	0.017	1.000	0.898	5.91E+46	0.000		
[Cuadrante=4.00]	-65.734	6273.922	0.000	1.000	0.992	2.832E-29	0.000		
[Cuadrante=5.00]	0.000			0.000					
[Birads=1.00]	49.260	160.359	0.094	1.000	0.759	2.47E+21	7.866E-116	7.77E+157	
[Birads=2.00]	-95.927	5639.609	0.000	1.000	0.986	2.185E-42	0.000		
[Birads=3.00]	418.587	2852.622	0.022	1.000	0.883	6.17E+181	0.000		
[Birads=4.00]	-290.223	2557.077	0.013	1.000	0.910	9.074E-127	0.000		
[Birads=5.00]	189.685	642.602	0.087	1.000	0.768	2.40E+82	0.000		
[Birads=6.00]	-15.397	216.658	0.005	1.000	0.943	2.057E-07	7.825E-192	5.40E+177	
[Birads=7.00]	0.000			0.000					

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

**Tabla 17-3:** Pruebas de la razón de verosimilitud

<b>Pruebas de la razón de verosimilitud de los factores para el modelo logístico</b>				
<b>Efecto</b>	<b>Criterios de ajuste de modelo</b>	<b>Pruebas de la razón de verosimilitud</b>		
		<b>Logaritmo de la verosimilitud -2 de modelo reducido</b>	<b>Chi-cuadrado</b>	<b>gl</b>
<b>Intersección</b>	,006 <sup>a</sup>	0.000	0.000	
<b>n_tam_1</b>	18,643	18.637	3.000	0.000
<b>n_ganglios_1</b>	35,817	35.811	3.000	0.000
<b>n_ganglios_2</b>	18,404	18.398	3.000	0.000
<b>n_CA</b>	55,865	55.860	3.000	0.000
<b>n_Ki</b>	21,328	21.322	3.000	0.000
<b>n_tam_2</b>	12,961	12.955	3.000	0.005
<b>Reporte</b>	63,868	63.863	9.000	0.000
<b>Localización de la mama</b>	19,849	19.843	3.000	0.000
<b>Cuadrante</b>	44,494	44.488	12.000	0.000
<b>Birads</b>	50,889	50.883	18.000	0.000

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

En la tabla 16-3, de acuerdo con los odds ratio, se aprecia que actúan como factor de riesgo para la mortalidad del cáncer de mama en estadio II y III el número de ganglios aislados de nivel II, el cuadrante en el que está ubicado específicamente en el superior externo e inferior interno.

En total fueron seleccionados diez factores y estos son: tamaño del tumor dimensión 1 (n\_tam1, X1), número de ganglios aislados de nivel I (n ganglios I, X2), número de ganglios aislados de nivel II (n ganglios II, X3), marcador tumoral CA 15-3 (n CA, X4), factor Ki 67 (n Ki, X5), tamaño del tumor dimensión 2 (n\_tam2, X6), Reporte del tipo de cáncer (Reporte, X7), Localización de la mama (Localización de la mama, X8), Cuadrante afectado por el tumor (Cuadrante, X9), BIRADS (Birads,X10), resultando todas estas variables significativas con un valor  $p > 0.05$  según la tabla 17-3.

### 3.4.2 Validación del modelo

El modelo estimado fue sometido a la Prueba de Bondad de Ajuste de Pearson y Prueba de Desvianza, mostrando los resultados en la tabla 18-3.

**Tabla 18-3:** Validación del modelo

<b>Validación del modelo para la mortalidad del cáncer de mama</b>			
	<b>Chi-cuadrado</b>	<b>gl</b>	<b>Sig.</b>
<b>Pearson</b>	0.003	240.000	1.000
<b>Desvianza</b>	0.006	240.000	1.000

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

Al realizar la prueba de bondad de ajuste de Pearson, cuyo estadístico calculado para el modelo es 0.003 ( $p = 1 > 0.05$ ), por lo tanto, el modelo de regresión logística es adecuado. La prueba de Desvianza, presenta el estadístico calculado para el modelo es ( $p = 0.006 > 0.05$ ), entonces no existe diferencia entre el modelo estimado y saturado, confirmando los hallazgos de la prueba de Pearson.

**Tabla 19-3:** Pseudo R cuadrado

<b>Pseudo R</b>	
Cox y Snell	0.733
Nagelkerke	1.000
McFadden	1.000

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

En la tabla 19-3 se obtiene la comparación pseudo r cuadrados entre Cox y Snell, Nagelkerke y McFadden, dando como resultado que el modelo de regresión logística encontrado explicará una varianza de los datos oscilando entre 73.3% y el 100%.

### 3.4.3 Evaluación del riesgo de pronóstico

Tabla 20-3: Clasificación con modelo de regresión estimado.

Observado	Pronosticado				Porcentaje correcto
	Estadio 0	Estadio I	Estadio II y III	Estadio IV	
Estadio 0	2.0	0.0	0.0	0.0	100.0%
Estadio I	0.0	7.0	0.0	0.0	100.0%
Estadio II y III	0.0	0.0	82.0	0.0	100.0%
Estadio IV	0.0	0.0	0.0	10.0	100.0%
Porcentaje global	2.0%	6.9%	81.2%	9.9%	100.0%

Fuente: Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

Realizado por: Quinatoa M., William G.,2020

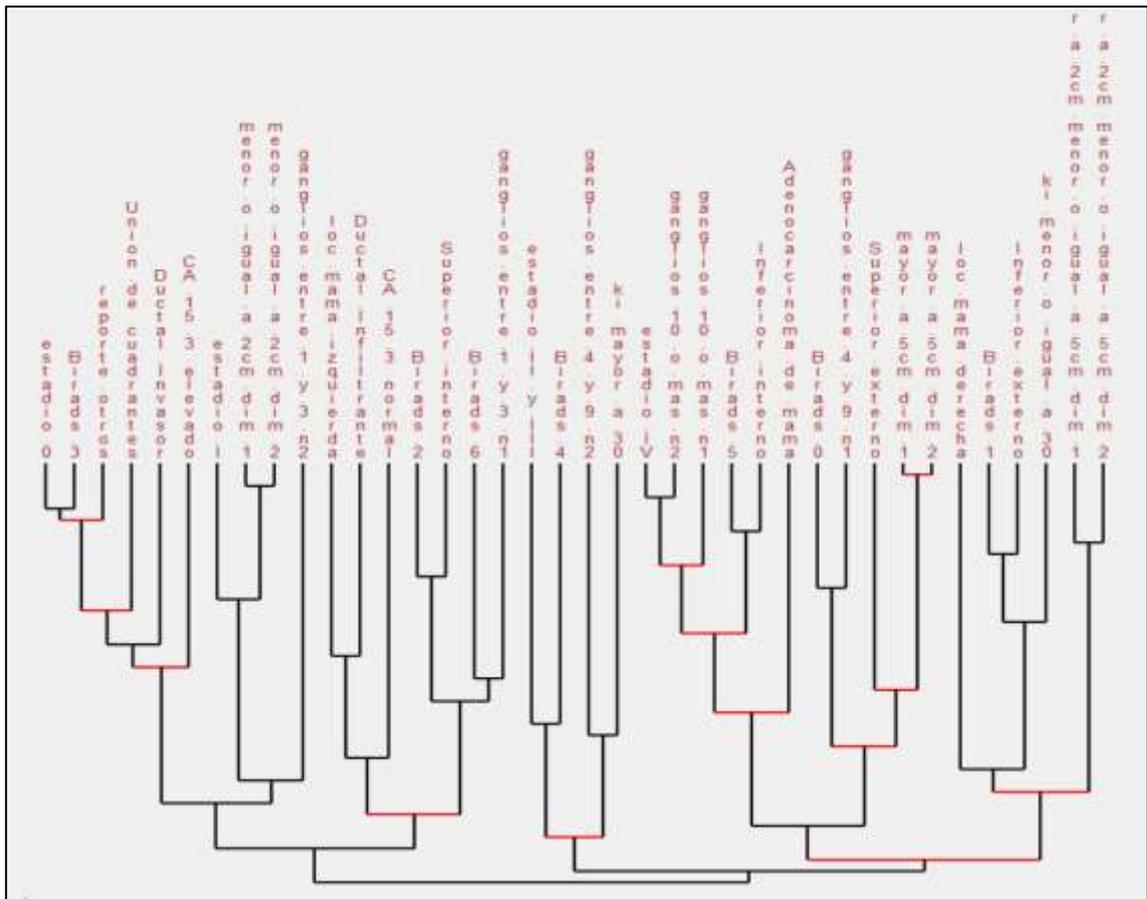
Los resultados de clasificar las observaciones para la mortalidad del cáncer de mama a través del estadio usando el modelo estimado son presentados en la tabla 20-3.

El modelo identifica correctamente todos los estadios referentes al cáncer de mama mostrando un éxito global del 100%.

### 3.5 Análisis Implicativo

Se lo aplicó utilizando el software R-CHIC versión 0.25, con el se obtuvieron árboles de similaridad, grafos implicativos y árboles de jerárquicos que se basaron en los datos en formato plano separado por comas y luego se exportaron en formato JPEG.

De acuerdo al gráfico 21-3 se muestra, el árbol de similaridad en él se puede apreciar la formación de dos grandes grupos o clases que engloban algunas variables y que establecen las conexiones de similaridad entre las diferentes clases de variables.



**Gráfico 21-3:** Árbol de Similaridad en R-CHIC.

**Realizado por:** Quinatoa M., William G.,2020

El grupo de la izquierda está conformado por la variable Estadio 0 asociado a un Birads III que se asocia a un quinto nivel que es el reporte de un Carcinoma distinto a los de más alta frecuencia, este reporte puede ser Papiloma ductal de mama, fibroadenoma de mama, papilomatosis intraductal o sarcoma de mama. En el grupo de la izquierda se relaciona con un nivel 13 conformado por localización del tumor en la unión de cuadrantes en el diagnóstico de un adenocarcinoma diferente a los más frecuentes, asociado a un nivel 16 con el reporte de un adenocarcinoma ductal invasor observándose una alta similaridad. En este mismo grupo se asocia el marcador tumoral CA 15-3 elevado, es decir mayor a 26,4 u/ml, dando a entender que al diagnosticar el tumor en la unión de cuadrantes su marcador tumoral CA 15-3 será elevado.

En este grupo de clases también se relaciona al número de ganglios aislados de nivel II que son de 1 a 3 con las dimensiones del tumor menores a 2 cm en los niveles 1 y 2 y clasificando al tumor como estadio I, dando como resultado una similaridad baja. El grupo de la izquierda asocia en un segundo nivel a las dimensiones 1 y 2 del tumor con valores menores a 2 cm con el diagnóstico de estadio I obteniendo una similaridad alta, asociando estos con el número de ganglios afectados con valores entre 1 y 3 de nivel II. En la clase 31 se asocia de manera significativa al marcador

tumoral en rango normal cuando se diagnóstica un adenocarcinoma ductal infiltrante localizado en la mama izquierda con al número de ganglios afectados entre 1 y 3 de nivel I de un Birads VI de localización de tumor en el cuadrante superior interno.

El grupo de la derecha está conformado por la variable Estadio IV asociado con el número de ganglios afectados de 10 o más de nivel II, y estas a la vez se asocia con el número de ganglios afectados de 10 o más de nivel I. En el mismo grupo se forma un nivel 15 que asocia las variables anteriores descritas con el diagnóstico de un Birads V en la parte de la mama inferior interna, obteniendo una similaridad alta. Al adenocarcinoma de mama que es uno de los diagnósticos con más frecuencia se lo asocia con un estadio IV junto con una afectación de ganglios de 10 o más en los niveles I y II. En el grupo de la derecha se crea la clase 25 que asocia al Birads 0 localizado cuyos ganglios afectados son entre 4 y 9 de nivel I con la dimensión del tumor que es mayor a 5 cm en las dos dimensiones localizado en el cuadrante superior externo de la mama, obteniendo una similaridad elevada. Dentro del mismo grupo de la derecha está el nivel de similaridad 33 que asocia a un estadio II y II con Birads IV con una afectación de ganglios de 4 a 9 de nivel 2 con un ki mayor a 30.

Otra asociación se da entre el diagnóstico de Birads I localizado en el cuadrante inferior externo con un Ki 67 menor o igual a 30 con la localización de la mama, siendo esta en la derecha. A las conexiones descritas anteriormente se asocia el nivel 29 de similaridad a que probablemente si el tumor se localiza en la mama derecha las dimensiones del tumor serán mayores a 2 cm y menor o iguales a 5 cm.

**Tabla 21-3:** Clasificación del árbol jerárquico

No	Clasificación	Interpretación	Coh
1	Estadio I ganglios entre 1 y 3 n2 Birads 4	Significa que el paciente diagnosticado con estadio I junto con el número de ganglios aislados entre 1 y 3 son nivel 2 se obtendrá un Birads 4.	0.299
2	Localización mama izquierda Ductal infiltrante Estadio II y III	Significa que si el tumor se localiza en la mama izquierda siendo un adenocarcinoma ductal infiltrante este se clasificara en un estadio II o III.	0.704

<b>3</b>	reporte otros Ca 15-3 elevado mayor a 2cm menor o igual a 5cm dim2 mayor a 2cm menor o igual a 5cm dim1	Significa que, si se reporta un adenocarcinoma papilar intraquístico, fibroadenoma, papiloma ductal, papilomatosis intraductal o un sarcoma con un marcador tumoral 15-3 elevado, la dimensión del tumor en ambas dimensiones será mayor a 2 cm y menor o igual a 5cm	0.471
<b>4</b>	Superior interno ki mayor a 30 Birads 6 ganglios entre 1 y 3 Ca 15-3 normal	Significa que si el tumor se localiza en el cuadrante superior interno de la mama junto con un KI 67 mayor a 30% probablemente el marcador tumoral CA 15-3 sea normal.	0.273
<b>5</b>	Unión de cuadrantes menor o igual a 2cm dim 1 menor o igual a 2cm dim 2	Significa que, si el tumor se localiza en la unión de cuadrantes, la dimensión del tumor será en las dos dimensiones menor o igual a 2cm.	0.772
<b>6</b>	Ganglios entre 4 y 9 n2 mayor a 5 cm dim2 mayor a 5cm dim 1 Superior externo	Significa que si los ganglios infectados del nivel 2 son entre 4 y 9 el tumor se localizara en el cuadrante superior externo y con dimensiones 1 y 2 mayores a 5 cm.	0.44

**Fuente:** Historias clínicas del Hospital Oncológico SOLCA de la provincia de Chimborazo

**Realizado por:** Quinatoa M., William G.,2020

El gráfico 21-3 muestra la representación gráfica algunas de las reglas de clasificación generadas automáticamente por el software estadístico R-CHIC para el árbol jerárquico.





**Regla R1:** Los pacientes diagnosticados con cáncer de mama con estadio I poseen un tumor con dimensión 1 menor o igual a 2 cm en la mama izquierda cuyos ganglios aislados son entre 1 y 3 de nivel 2, con una intensidad implicativa comprendida entre 80% y 85%.

**Regla R2:** A los tumores localizados en la unión de cuadrantes fueron dimensionados de orden 2 con valores menores o iguales a 2 cm con una intensidad implicativa del 85%, en el cual se detectó de 1 a 3 ganglios aislados de nivel 1 con una intensidad del 80%.

**Regla R3:** A los pacientes que se les ha evaluado con cáncer de mama en el cuadrante superior interno localizado en la mama derecha siendo un adenocarcinoma ductal infiltrante, sus exámenes inmunohistoquímicos serán de un Ki 67 mayor al 30% y un marcador tumoral CA 15-3 normal (0-26.4 u/ml), con una intensidad implicativa comprendida entre 80% y 85%.

## CONCLUSIONES

Los factores pronóstico que se asocian a la mortalidad del cáncer de mama son: localización de la mama afectada, tipo de Birads, reporte de cáncer, clasificación del estadio, cuadrante de la mama, tamaño del tumor en dimensión 1, tamaño del tumor en dimensión 2, número de ganglios aislados nivel I, número de ganglios aislados nivel II, marcador tumoral CA 15-3 y Ki 67, obteniendo un modelo de predicción logístico de 73% según el coeficiente del Cox y Snell, y un análisis implicativo que determino que todas las variables mencionadas anteriormente ayudan a un buen pronóstico.

Al realizar pruebas no paramétricas para los factores de riesgo sociales como ocupación y estado civil, se obtuvo que estas variables no están asociadas a la mortalidad de cáncer de mama con valores p iguales a 0.240 y 0.880 respectivamente, siendo ambas mayores a 0.05, pero estableciendo que la mayor frecuencia de pacientes tiene como ocupación ama de casa y con estado civil casada.

Para los factores materno-culturales se estableció que el nivel de instrucción no es significativo para la mortalidad de cáncer de mama, obteniendo un valor p igual a 0.987, siendo este valor mayor a 0.05, por lo tanto, este factor no es influyente en esta investigación.

Se comparó las tasas brutas de mortalidad según el sexo de los pacientes, siendo en esta investigación del 100% para las mujeres, ya que no se encontraron pacientes hombres en la base de datos que ayuden a la realización de esta investigación.

La prueba no paramétrica Chi-cuadrado comparó los factores pronósticos con la clasificación del tumor TNM (estadio) y se obtuvo que las variables localización de la mama (0.642), reporte de tipo de cáncer (0.078), cuadrante donde se encuentra localizado el tumor (0.986), tamaño del tumor 1 (0.090) y 2 (0.485), número de ganglios aislados de nivel I ( $p=0.727$ ), marcador tumoral CA 15-3 ( $p=0.352$ ) y Ki 67 ( $p=0.304$ ) no presentaron asociación significancia con valores  $p > 0.05$ , en comparación con otros factores pronóstico como la variable Birads (0.019) y número de ganglios aislados de nivel II ( $p=0.001$ ), demostrando que estas últimas están asociadas a la clasificación del tumor TNM (estadio) con valores  $p < 0.05$ .

## **RECOMENDACIONES**

Realizar campañas de concientización a mujeres y hombres sobre los factores de riesgo para superar el cáncer de mama, promoviendo charlas grupales que hablen de los riesgos de dicha enfermedad, ya que un diagnóstico oportuno ayuda a disminuir la tasa de mortalidad.

Digitalizar las historias clínicas de los distintos pacientes que padecen esta enfermedad en los hospitales, ya que al no estarlo se corre el riesgo de tener la existencia de datos faltantes y atípicos, que causen investigaciones con resultados sesgados.

Promover a docentes e investigadores del área médica y estadística al análisis de este tipo de datos, sobre todo la aplicación de las técnicas de regresión logística y análisis implicativo, ya que van de la mano con el mejoramiento de diagnósticos y tratamientos de enfermedades.

Realizar otro tipo de estudios para esta enfermedad, como análisis de discriminante y modelos de supervivencia, ya que estos ayudaran a complementar el estudio realizado sobre el cáncer de mama.

## GLOSARIO

**Coefficientes ( $\beta$ ):** Los coeficientes  $\beta_i$  expresan el logaritmo neperiano del odds ratio (OR) para cada factor de riesgo  $X_i$  (Fiuza y Rodríguez, 2000).

**Estadístico Wald:** El estadístico de Wald es una medida de distancia para la cual el estimador (no restringido) falla en satisfacer las restricciones (Gallón, 2014).

**Odds ratio:** Este cociente, que en inglés se denomina *odds* y para el que no hay una traducción española comúnmente aceptada, indica cuanto más probable es la ocurrencia del evento que su no ocurrencia (Abraira, 2019).

**Razón de las ventajas estimadas ( $\text{Exp}(B)$ ) e intervalo de confianza para  $\text{Exp}(B)$ :** Permitirá, mediante la regresión logística, ajustar por variables de confusión. El estimador del odds ratio para un factor de riesgo cuando este interactúa con otra variable, depende del valor de la variable que está interactuando, por lo tanto, interesa dar los intervalos de confianza para los odds ratio (Silva, 2018, p. 50).

**Prueba de Hosmer-Lemeshow:** Es una prueba estadística para bondad de ajuste para modelos de regresión logística, se utiliza con frecuencia en la predicción del riesgo modelos, además la prueba evalúa si o no las tasas de eventos observados coinciden con las tasas de eventos esperados en los subgrupos de la población modelo (Hosmer, 2013).

**Prueba de Desviante:** Proporcionará una prueba de hipótesis para evaluar el ajuste del modelo. Las hipótesis para evaluar son:

$H_0$ : No existe diferencia entre el modelo estimado y el modelo saturado.

$H_1$ : Existe diferencia entre el modelo estimado y el modelo saturado (Silva, 2018, p. 50).

## BIBLIOGRAFÍA

**AMERICAN SOCIETY OF CLINICAL ONCOLOGY.** Cáncer de mama: Estadísticas. [En línea]. 2012. [Consulta: 25 de marzo 2020]. Disponible en: <https://www.cancer.net/es/tipos-de-c%C3%A1ncer/c%C3%A1ncer-de-pr%C3%B3stata/estad%C3%ADsticas>.

**AMADOU, A.; et al.** Breast cancer in Latin America global burden, patterns, and risk factors. *Salud pública de México*, 2014, vol. 56, pp. 547-554. [Consulta: 8 de febrero de 2019]. Disponible en: <http://www.medigraphic.com/pdfs/salpubmex/sal-2014/sal145q.pdf>

**ALMEIDA, ANDRÉ.; & CASSIA, ADELAIDE.** Socioeconomic inequalities in cancer incidence and mortality: review of ecological studies, 1998-2008. *Saúde Soc. São Paulo*, 2013, vol. 22, pp. 878-891. [Consulta: 8 de febrero de 2019]. Disponible en: [http://www.scielo.br/scielo.php?pid=S0104-12902013000300020&script=sci\\_arttext&tlng=pt](http://www.scielo.br/scielo.php?pid=S0104-12902013000300020&script=sci_arttext&tlng=pt)

**BAENA, A.; et al.** Tendencias e indicadores sociales de la mortalidad por cáncer de mama y cuello uterino. Antioquia, Colombia 2000-2007. *Salud pública de México*, 2011, vol. 53, pp. 420-429. [Consulta: 5 de febrero de 2019]. Disponible en: [https://www.scielosp.org/article/ssm/content/raw/?resource\\_ssm\\_path=/media/assets/spm/v53n6/a04v53n6.pdf](https://www.scielosp.org/article/ssm/content/raw/?resource_ssm_path=/media/assets/spm/v53n6/a04v53n6.pdf)

**BAIDER, LEA.** Cáncer y familia: aspectos teóricos y terapéuticos. *International Journal of Clinical and Health Psychology*, 2003. [Consulta: 8 de febrero de 2019]. Disponible en: <http://www.redalyc.org/articulo.oa?id=33730306>

**BRAVO, LE.; et al.** Burden of breast cancer in Cali, Colombia: 1962-2012, *Salud pública de México*, 2014, vol. 53, pp. 448-456. [Consulta: 8 de febrero de 2019]. Disponible en: <https://www.ncbi.nlm.nih.gov/pubmed/25604291>

**CORONADO, G.; et al.** Breast cancer in the world: Incidence and mortality. *Salud pública de México*, 2011, vol. 53, pp. 372-384. [Consulta: 5 de febrero de 2019]. Disponible en: <https://www.ncbi.nlm.nih.gov/pubmed/22218798>

**CORONADO, G.; BEASLEY, J.; LIVAUDAIS, J.** Alcohol consumption and the risk of breast cancer. *Salud pública de México*, 2011, vol. 53, pp. 440-447. [Consulta: 5 de febrero de 2019]. Disponible en: <https://www.ncbi.nlm.nih.gov/pubmed/22218798>

**DÁVILA, CLAUDIO.; AGUDELO MARCELA.** Average years of life lost due to breast and cervical cancer and the association with the marginalization index in Mexico in 2000 and 2010 , *Cad. Saúde Pública, Rio de Janeiro*, 2013, vol. 22, pp. 1093-1102. [Consulta: 8 de febrero de 2019]. Disponible en: [https://www.researchgate.net/publication/263165729\\_Average\\_years\\_of\\_life\\_lost\\_due\\_to\\_breast\\_and\\_cervical\\_cancer\\_and\\_the\\_association\\_with\\_the\\_marginalization\\_index\\_in\\_Mexico\\_in\\_2000\\_and\\_2010](https://www.researchgate.net/publication/263165729_Average_years_of_life_lost_due_to_breast_and_cervical_cancer_and_the_association_with_the_marginalization_index_in_Mexico_in_2000_and_2010)

**FERNÁNDEZ PÉREZ, M.; et al.** Factores de riesgo modificables en algunos tipos de cáncer. *Revista Electrónica Dr. Zoilo E. Marinello Vidaurreta* [en línea], 2016, [Consulta: 8 de febrero de 2019]. Disponible en: <http://revzoilomarinaldo.sld.cu/index.php/zmv/article/view/940>

**FLORES, L.; et al.** Factores pronósticos relacionados con la supervivencia del cáncer de mama. *Salud pública de México*, 2008, vol. 50, n°. 2, pp. 119-125. [Consulta: 5 de febrero de 2019]. Disponible en: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0036-36342008000200005](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0036-36342008000200005)

**GONZÁLEZ, JOSÉ.; et al.** Factores pronósticos del cáncer de mama. *Rev Cubana Cir*, 2011, vol. 50, pp 130-138. [Consulta: 8 de febrero de 2019]. Disponible en: [http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S0034-74932011000100013&lng=es](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0034-74932011000100013&lng=es).

**LÓPEZ, L.; et al.** Utilización correcta de las técnicas de detección de cáncer de mama en mujeres mexicanas, *Salud pública de México*, 2014, vol. 53, pp. 538-546. [Consulta: 8 de febrero de 2019]. Disponible en: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0036-36342014000500021](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0036-36342014000500021)

**MEDINA, GONZALO.** Características clínicas y pronósticas de los subtipos moleculares de cáncer de mama determinados por inmunohistoquímica. *Rev. Perú. med. exp. salud pública* [en línea]. 2017, vol. 34, n°. 3, pp.472-477. [Consulta 2019-02-08].,Disponible en: [http://www.scielo.org.pe/scielo.php?script=sci\\_arttext&pid=S1726-46342017000300014&lng=es&nrm=iso](http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1726-46342017000300014&lng=es&nrm=iso)

**MOHAR, A.; et al.** Factores pronósticos en pacientes con cáncer de mama y metástasis cerebral como primer sitio de recurrencia. *Salud pública de México*, 2018, vol. 60, (México) pp. 141-150. [Consulta: 5 de febrero de 2019]. Disponible en: <http://webcache.googleusercontent.com/search?q=cache:rsDXCMnn9IgJ:saludpublica.mx/index>

.php/spm/article/download/9082/11453+&cd=1&hl=es-419&ct=clnk&gl=ec

**MORAGA RODRÍGUEZ, ALINA.; et al.** Análisis estadístico implicative para la identificación de factores pronósticos de la mortalidad por cáncer de mama. *MEDISAN* [En línea], 2017, vol. 21, pp. 413-424. [Consulta: 8 de febrero de 2019]. Número ISSN 1029-3019. Disponible en: <http://www.redalyc.org/articulo.oa?id=368450539003>

**NAROD, SA.; & RODRÍGUEZ, A.** Predisposición genética para el cáncer de mama: genes BRCA1 y BRCA2. *Salud pública de México*, 2011, vol. 53, pp. 420-429. [Consulta: 5 de febrero de 2019]. Disponible en: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0036-36342011000500010](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0036-36342011000500010)

**FIUZA, M.; & RODRIGUEZ, J.** *La regresión logística: una herramienta versátil* [En línea]. 2000. pp 447-565. [Consulta: 19 febrero del 2018]. Disponible en: <https://www.revistanefrologia.com/es-la-regresion-logistica-una-herramienta-versatil-articulo-X0211699500035664>

**ORGANIZACIÓN PANAMERICANA DE LA SALUD.** Programa de Cáncer. [En línea]. 2020. [Consulta: 25 marzo 2020]. Disponible en: [https://www.paho.org/hq/index.php?option=com\\_content&view=article&id=292:cancer-program&catid=1872:cancer&Itemid=3904&lang=es](https://www.paho.org/hq/index.php?option=com_content&view=article&id=292:cancer-program&catid=1872:cancer&Itemid=3904&lang=es).

**PAZMIÑO, R.; & MIGUEZ, R.** Software Estadístico CHIC: descubriendo sus potencialidades mediante el análisis de percepción sexual universitaria. *Ciencia Digital*, 2018, vol. 2, pp. 122-139. [Consulta: 6 de febrero de 2019]. Disponible en: <http://cienciadigital.org/revistacienciadigital2/index.php/CienciaDigital/article/view/194/172>

**PERALTA, M.** Cáncer de mama en Chile: datos epidemiológicos: Se describen los análisis epidemiológicos del cáncer de mama en Chile y se relata la experiencia en el Hospital Clínico San Borja-Arriarán, destacando la mayor precocidad en el diagnóstico. *Rev. chil. obstet. ginecol*, 2002, vol.67, pp.439-445. [Consulta: 8 de febrero de 2019]. Disponible en: [https://scielo.conicyt.cl/scielo.php?script=sci\\_arttext&pid=S0717-75262002000600002](https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0717-75262002000600002&lng=es&nrm=iso). ISSN 0048-766X. <http://dx.doi.org/10.4067/S0717-75262002000600002>.

**RUIZ, FRANCISCO.; et al.** Análisis de la supervivencia del cáncer de mama durante el decenio

1999-2008 en un hospital público de Madrid". *Revista Española Salud Pública*, 2011, vol. 86, pp. 589-600. [Consulta: 8 de febrero de 2019]. Disponible en: [http://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S1135-57272012000600005](http://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S1135-57272012000600005)

**SILVA, SAULO.** Modelo predictivo de regresión logística en función a los factores de riesgo para contraer cáncer de mama. Universidad Nacional Santiago Antúnez de Mayolo.

**VENTURELLI, FRANCISCO.** Cáncer vesicular en colecistectomías: Análisis de sobrevida y factores pronósticos. *Rev Chil Cir*, 2008, vol.60, pp.277-281. [Consulta: 8 de febrero de 2019]. Disponible en: [https://scielo.conicyt.cl/scielo.php?script=sci\\_arttext&pid=S0718-40262008000400003&lng=es&nrm=iso](https://scielo.conicyt.cl/scielo.php?script=sci_arttext&pid=S0718-40262008000400003&lng=es&nrm=iso). ISSN 0718-4026. <http://dx.doi.org/10.4067/S0718-40262008000400003>.

## ANEXOS

### ANEXO A: CÓDIGO PARA DIAGRAMAS DE CAJA DE VARIBALES CUANTITATIVAS

#### *Librerías*

```
library(openxlsx)  
library(ggplot2)
```

#### *Lectura de datos*

```
data.original <- read.xlsx("DATOS_CA_DE_MAMA.xlsx")  
  
Base_de_datos_final <- read.xlsx("DATOS_CA_DE_MAMA.xlsx", sheet = "ALAO_PT")  
view(Base_de_datos_final)
```

#### *Funciones*

```
datapar(mfrow=c(2,2))  
boxplot(Base_de_datos_final$Peso,main="Valores atípicos de Peso",xlab="Peso",col="yellow3")  
  
boxplot(Base_de_datos_final$Talla,main="Valores atípicos de Talla",xlab="Talla",col="yellow3")  
  
boxplot(Base_de_datos_final$Edad,main="Valores atípicos de Edad",xlab="Edad",col="green3")  
  
boxplot(regre$`Tamaño tumor dim 1`,main="Valores atípicos",xlab="Tamaño del tumor dimensión  
1",col="yellow3")  
  
boxplot(regre$`Tamaño tumor 2`,main="Valores atípicos",xlab="Tamaño del tumor dimensión  
2",col="green3")  
  
boxplot(regre$`Numero de ganglios aislados Nivel I`,main="Valores atípicos",xlab="Numero de  
ganglios aislados Nivel I",col="yellow3")  
  
boxplot(regre$`Numero de ganglios aislados Nivel II`,main="Valores atípicos",xlab="Numero de  
ganglios aislados Nivel II",col="green3")  
  
boxplot(regre$`CA 15-3 U/ml (0-26.4)`,main="Valores atípicos",xlab="CA 15-3",col="yellow3")  
  
boxplot(regre$`ki 67 (%)`,main="Valores atípicos",xlab="ki 67 (%)",col="yellow3")
```

## ANEXO B: CÓDIGO EN R PARA EL ANÁLISIS IMPLICATIVO

### *Librerías*

```
library(openxlsx)
library(ggplot2)
library(Rgraphvi)
library(graph)
library(stringr)
library(tcltk)
library(tcltk2)
library(Rcpp)
library(rchic)
```

### *Lectura de datos*

```
data.original <- read.xlsx("DATOS_CA_DE_MAMA.xlsx")

Base_de_datos_final$ <- read.xlsx("DATOS_CA_DE_MAMA.xlsx", sheet = "ALAO_PT")
View(Base_de_datos_final$)
```

### *Funciones*

```
rchic()
```



**ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO**

**DIRECCIÓN DE BIBLIOTECAS Y RECURSOS PARA EL  
APRENDIZAJE Y LA INVESTIGACIÓN**



**UNIDAD DE PROCESOS TÉCNICOS  
REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y  
BIBLIOGRAFÍA**

*Fecha de entrega: 20 / 10 /2020*

<b>INFORMACIÓN DEL AUTOR/A (S)</b>
<b>Nombres – Apellidos:</b> William Geovanny Quinatoa Martínez
<b>INFORMACIÓN INSTITUCIONAL</b>
<b>Facultad:</b> Ciencias
<b>Carrera:</b> Ingeniería en Estadística Informática
<b>Título a optar:</b> Ingeniero en Estadística Informática
<b>f. Analista de Biblioteca responsable:</b> Lic. Luis Caminos Vargas Mgs.