

# UNIVERSIDAD DE SONORA DIVISIÓN DE INGENIERÍA



## POSGRADO EN INGENIERÍA INDUSTRIAL MAESTRÍA EN INGENIERÍA EN SISTEMAS Y TECNOLOGÍA

DESARROLLO DE UN MODELO PARA APOYAR EL PROCESO DE  
ATENCIÓN DE PACIENTES CON CÁNCER DE MAMA EN UNA  
INSTITUCIÓN DE SALUD CON MINERÍA DE DATOS

### T E S I S

PRESENTADA POR

**ANGÉLICA ENRÍQUEZ AMAYA**

Desarrollada para cumplir con uno de los  
requerimientos parciales para obtener  
el grado de Maestra en Ingeniería

DIRECTORA DE TESIS  
DRA. RAQUEL TORRES PERALTA

CODIRECTOR  
DR. GUZMÁN GERARDO ALFONSO SÁNCHEZ SCHMITZ

HERMOSILLO, SONORA, MÉXICO.

ENERO 2020

# Universidad de Sonora

Repositorio Institucional UNISON



**"El saber de mis hijos  
hará mi grandeza"**



Excepto si se señala otra cosa, la licencia del ítem se describe como openAccess



"El saber de más hijos  
hace en grande"

Hermosillo, Sonora a 11 de diciembre de 2019

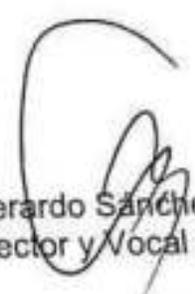
**ANGÉLICA ENRÍQUEZ AMAYA**

Con fundamento en el artículo 66, fracción III, del Reglamento de Estudios de Posgrado vigente, otorgamos a usted nuestra aprobación de la fase escrita del examen de grado, como requisito parcial para la obtención del Grado de Maestra en Ingeniería: Ingeniería en Sistemas y Tecnología.

Por tal motivo este jurado extiende su autorización para que se proceda a la impresión final del documento de tesis: **DESARROLLO DE UN MODELO PARA APOYAR EL PROCESO DE ATENCIÓN DE PACIENTES CON CÁNCER DE MAMA EN UNA INSTITUCIÓN DE SALUD CON MINERÍA DE DATOS** y posteriormente efectuar la fase oral del examen de grado.

ATENTAMENTE

  
Dra. Raquel Torres Peralta  
Directora de tesis y Presidente del jurado

  
Dr. Gerardo Sánchez Schmitz  
Codirector y Vocal del Jurado

  
Dr. Federico Miguel Cirett Galán  
Secretario del Jurado

  
Dr. René Francisco Navarro  
Hernández  
Vocal del Jurado



UNIVERSIDAD DE  
MURCIA

Murcia, España, a 16 de diciembre de 2019

**ANGÉLICA ENRÍQUEZ AMAYA**

Con fundamento en el artículo 66, fracción III, del Reglamento de Estudios de Posgrado de la Universidad de Sonora, otorgo a usted mi aprobación de la fase escrita del examen profesional, como requisito parcial para la obtención del Grado de Maestra en Ingeniería.

Por tal motivo, como sinodal externo y vocal del jurado, extiendo mi autorización para que se proceda a la impresión final del documento de tesis: **DESARROLLO DE UN MODELO PARA APOYAR EL PROCESO DE ATENCIÓN DE PACIENTES CON CÁNCER DE MAMA EN UNA INSTITUCIÓN DE SALUD CON MINERÍA DE DATOS** y posteriormente efectuar la fase oral del examen de grado.

ATENTAMENTE

DR. JOSÉ MANUEL JUÁREZ HERRERO  
UNIVERSIDAD DE MURCIA  
Sinodal Externo y Vocal del Jurado



# RESUMEN

Con la presente investigación se busca desarrollar un modelo, que permita identificar áreas de oportunidad para establecer señalamientos y sugerencias de mejora en el proceso de atención de cáncer de mama. La explotación de estos datos es necesaria para una mejor toma de decisiones a la hora de hacer ajustes en los procedimientos y protocolos que ayude a los médicos y al personal de las instituciones de salud a tener una mayor comprensión del problema.

La propuesta consta de dos ejes principales: 1) utilización de análisis de georreferenciación para la vigilancia epidemiológica y 2) revisión y mejora del proceso de atención a pacientes con cáncer de mama. La metodología utilizada está basada en el uso de la minería de datos y consta de 6 fases: 1) identificación, 2) recopilación de datos, 3) preprocesamiento 4) creación del modelo, 5) interpretación y evaluación, 6) implementación.

Se utilizó el análisis de georreferenciación para la vigilancia epidemiológica donde se observó focos rojos al poniente y sur de la ciudad, la mayor frecuencia de casos de padecimientos de la mama se concentra en las edades de los 35-49 años para el sexo femenino, mientras que para el sexo masculino se da entre 10-18 años. Se identificaron las principales actividades del proceso de atención de cáncer de mama identificando áreas de oportunidad. Finalmente, con esto se logró alcanzar el objetivo de generar conocimiento que pueda ser utilizado en el soporte a la toma de decisiones sobre proceso de atención de cáncer de mama realizando reportes a la institución de los resultados obtenidos para la toma de decisiones y las estrategias de prevención.

# ABSTRACT

The present research seeks to develop a model that allows to identify areas of opportunity to establish indications and suggestions for improvement in the process of breast cancer care. The exploitation of this data is necessary for better decision-making in the adjustment of procedures and protocols to help doctors and staff in health institutions to have a better understanding of the problem.

The proposal has two main axes: (1) use of georeferencing analysis for epidemiological surveillance and (2) review and improvement of the process of care for breast cancer patients. The methodology used is based on the use of data mining and consists of 6 phases: 1) identification, 2) data collection, 3) pre-processing, 4) model creation, 5) interpretation and evaluation and 6) implementation.

Georeferencing analysis was used for epidemiological surveillance where red spots were observed in the west and south of the city, the highest frequency of cases of breast ailments is concentrated in the ages of 35-49 years for the female sex, while for the male sex it is given between 10-18 years. Identified the main activities of the breast cancer care process by identifying areas of opportunity. Finally, this achieved the objective of generating knowledge that can be used to support decision-making on breast cancer care process by reporting to the institution the results obtained to decision-making and prevention strategies.

## **AGRADECIMIENTOS**

A Dios, a la vida, a mi familia especialmente a mis padres Teresita y Jesús María, a mis amigos y compañeros que formaron parte de esta aventura.

Agradezco a la Dra. Raquel Torres Peralta por su liderazgo a lo largo de toda esta investigación, gracias por compartirme sus conocimientos, por sus consejos, ¡INFINITAS GRACIAS! también a mis codirectores, los Doctores Federico Cirett Galán, Gerardo Sánchez, y al Dr. Francisco Navarro, también agradezco al Dr. José Manuel Juárez por su guía y consejos durante mi estancia en Murcia.

Al personal administrativo, al área de quimioterapia y médicos de la Institución que amablemente me dieron un poco de su tiempo que sin ellos no hubiese sido posible realizar parte de este trabajo, pero especialmente a las pacientes por compartir conmigo su proceso ¡INFINITAS GRACIAS!

Al departamento de Ingeniería Industrial y a la misma Universidad de Sonora por darme la oportunidad de pertenecer a sus filas.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) y al Programa de Fortalecimiento de la Calidad Educativa (PFCE) por su apoyo económico brindado en mi estudio de posgrado.

# ÍNDICE GENERAL

RESUMEN .....	i
ABSTRACT .....	ii
AGRADECIMIENTOS .....	iii
ÍNDICE GENERAL .....	iv
ÍNDICE DE FIGURAS .....	vii
ÍNDICE DE TABLAS .....	xi
1. INTRODUCCIÓN.....	1
1.1. Presentación .....	2
1.2. Planteamiento del problema.....	4
1.3. Objetivo general .....	4
1.4. Objetivos específicos .....	5
1.5. Hipótesis.....	5
1.6. Alcances y delimitaciones .....	5
1.7. Justificación.....	5
2. MARCO DE REFERENCIA .....	7
2.1. Salud Pública .....	7
2.2. Cáncer de Mama.....	8
2.3. Aprendizaje Automatizado.....	14
2.4. Minería de Datos .....	15
2.4.1. Clasificación y Predicción .....	17
2.4.2. Adquisición de Datos .....	17
2.4.3. Preparación de Datos .....	17
2.4.4. Minería de Procesos .....	17
2.5. Tecnologías de la Información en Áreas de la Salud .....	19
2.5.1. Minería de Datos como herramienta en el sector Salud .....	19
2.6. Usos más comunes de la minería de datos.....	23
2.7. Algoritmos y técnicas de Minería de Procesos.....	24

2.7.1. Algoritmos de descubrimiento .....	24
2.7.2. Técnicas de descubrimiento.....	25
2.8. Comparativa de técnicas de Minería de Procesos .....	26
2.9. Trabajos previos.....	27
3. MODELO.....	30
3.1. Identificación .....	31
3.2. Recopilación de datos .....	32
3.3. Preprocesamiento de datos.....	33
3.4. Creación del modelo .....	33
3.5. Interpretación y Evaluación .....	36
3.6. Implementación .....	38
4. IMPLEMENTACIÓN .....	39
4.1. Identificación .....	39
4.1.1. Criterios.....	39
4.1.2. Pacientes .....	42
4.1.3. Tiempo entre consultas.....	57
4.2. Recopilación de datos .....	59
4.2.1. Aplicación de encuesta, revisión de documentación y publicaciones oficiales.....	59
4.3. Preparación de datos .....	63
4.4. Creación del modelo .....	63
4.4.1. Diseño del modelo basado en guías clínicas (modelo declarativo).....	64
4.4.2. Diseño del modelo de proceso (práctica clínica real).....	68
4.5. Interpretación y Evaluación .....	78
4.6. Implementación .....	96
5. CONCLUSIONES RECOMENDACIONES Y TRABAJOS FUTUROS .....	98
5.1. Conclusiones.....	98
5.2. Recomendaciones.....	99
5.3. Trabajos futuros .....	101
6. REFERENCIAS.....	102

7. ANEXOS.....	108
7.1. Entrevistas.....	108
7.2. Documento de consentimiento.....	115
7.3. Resultados de la encuesta .....	118

# ÍNDICE DE FIGURAS

Figura 1.1. Proceso de Atención Médica para un paciente con probabilidad de padecer cáncer de mama .....	2
Figura 2.1. Estructura Sector Público (Gómez-Dantés et al., 2011).....	8
Figura 2.2. Algoritmo para la valoración de nódulos detectados por mastografía .....	14
Figura 2.3. Etapas de la minería de datos según López-Ramírez (2017) pp. 40. ....	16
Figura 2.4. Posición de la minería de procesos según Van der Aalst, W. M. (2016).	18
Figura 2.5. Árbol de descomposición en subprocesos (Torres-Sakipova, 2014).....	26
Figura 3.1. Modelo adaptación KDD (1996) y van der Aalst (1999). .....	31
Figura 3.2. Medidas de evaluación.....	37
Figura 4.1. Comportamiento anual al alza de casos de cáncer de mama. ....	43
Figura 4.2. Porcentaje de incidencia con respecto a la población de la institución por cada año, correspondiente a los casos de padecimientos de la mama de acuerdo a la muestra proporcionada por la institución.....	44
Figura 4.3. Frecuencias por tipo de diagnóstico.....	45
Figura 4.4: Grupos de edad en los padecimientos “Fibroadenosis de mama”, “Masa no especificada en la mama” y “Mastalgia-mastodinia-mastitis-dolor mama”. ....	46
Figura 4.5. Casos por Grupos de edad para los casos de “CA mama”, “Carcinoma in situ de la mama” y “Tumor maligno de la mama”. .....	47
Figura 4.6. Incidencia de los padecimientos de la mama por grupo de edad y género. ....	48
Figura 4.7. Diagnósticos por género .....	49
Figura 4.8. Mapa de la ciudad de Hermosillo con zonas calientes de los casos de “CA mama”, “Tumor maligno de la mama” y “Carcinoma in situ de la mama”. ....	50
Figura 4.9. La distribución por grupos de edad y colonia, se puede apreciar que hay un pico en el grupo de edad de los 35-39 años, así como el de los 30-34 años.....	51

Figura 4.10. La distribución de coordenadas geográficas (latitud, longitud) no muestra una segmentación definida. ....	51
Figura 4.11. Código en Python utilizado para realizar la curva del codo en la que aprecia un cambio brusco en la evolución de la inercia donde $k=3$ . ....	53
Figura 4.12. Centroides de la ciudad de Hermosillo.....	53
Figura 4.13. Identificación de 3 clúster diferenciados por colores: clúster 1=rojo; clúster 2=verde; clúster 3=azul. ....	54
Figura 4.14. Georeferenciación de clúster de la muestra de casos de cáncer de mama en Hermosillo. ....	54
Figura 4.15. Casos atípicos entre los grupos de edad de 30-34 años y 34-39 años, por clúster.....	56
Figura 4.16. Promedio en días entre citas por casos únicos en la institución. ....	58
Figura 4.17. Modelo declarativo del proceso de atención de cáncer de mama.....	66
Figura 4.18. Actividades recurrentes en el proceso de cáncer de mama. ....	71
Figura 4.19. Actividades iniciales del proceso de atención de cáncer de mama.....	72
Figura 4.20. Actividades finales del proceso de atención de cáncer de mama. ....	73
Figura 4.21. Descubrimiento del modelo de proceso de atención de cáncer de mama (parte 1).....	75
Figura 4.22. Descubrimiento del modelo de proceso de atención de cáncer de mama (parte 2).....	76
Figura 4.23. Descubrimiento del modelo de proceso de atención de cáncer de mama (parte 3).....	77
Figura 4.24. El modelo de proceso presenta una idoneidad de 79.48%, el cual presenta algunas desviaciones y violaciones con respecto al modelo basado en guías clínicas. ....	79
Figura 4.25. Evaluación del proceso de atención de cáncer de mama mediante 3 indicadores: tiempo de espera promedio, el recíproco y tiempo de permanencia. ....	81

Figura 4.26. Resumen campo G01: tiempos en que tarda un paciente en asistir al médico por primera vez desde el momento que se detectó algo anormal .....	85
Figura 4.27. Resumen campo G13: ¿el tratamiento fue aplicado en tiempo indicado por el médico?.....	86
Figura 4.28. Resumen de campo para I01, ¿Ha recibido información sobre la enfermedad por parte del instituto u otros medios? .....	87
Figura 4.29. Resumen de campo E01, ¿Acudieron a su llamado cuando las necesitó? .....	88
Figura 4.30. Resumen de campo E01, ¿Consideró adecuados los muebles del cuarto/sala?.....	88
Figura 4.31. Resumen de campo E01 en general, ¿diría que el personal de enfermería le ha demostrado interés? .....	89
Figura 4.32. Resumen de campo E01, ¿fueron atentas las personas encargadas de su ingreso?: .....	89
Figura 4.33. Resumen de campo E01, ¿hicieron su trabajo en tiempo adecuado? ..	90
Figura 4.34. Resumen de campo E03, ¿fue cómoda la cama?.....	91
Figura 4.35. Resumen de campo E02, ¿Le explicaron ampliamente la enfermedad? .....	91
Figura 4.36. Resumen de campo E02, antes de realizar cualquier procedimiento o aplicarle algún tratamiento, ¿se lo explicaron? .....	92
Figura 4.37. Resumen de campo E02, ¿cómo fue el trato demostrado por el personal de las comidas? .....	92
Figura 4.38. Resumen de campo E02, ¿le dijeron qué artículos de uso personal necesitaría traer usted y cuáles le proporcionaría el hospital?.....	93
Figura 4.39. Resumen de campo E02, ¿cómo fue el servicio de limpieza durante la hospitalización?.....	93
Figura 4.40. Resumen de campo E03. ¿diría que el personal médico le ha demostrado respeto? .....	94

Figura 4.41. Resumen de campo E03, cuando recibió el diagnóstico, ¿Considera que fue el modo adecuado de hacérselo saber? .....	94
Figura 4.42. Resumen de campo E03, ¿cómo es el trato que recibe del personal de laboratorio? .....	95
Figura 7.1. Bienvenida a la encuesta. ....	108
Figura 7.2. Sección general.....	109
Figura 7.3. Sección estatus general. ....	110
Figura 7.4. Sección Información sobre la enfermedad. ....	111
Figura 7.5. Evaluación personal enfermería.....	112
Figura 7.6. Evaluar personal médico y hospitalización.....	113
Figura 7.7. Evaluar servicio en general. ....	114
Figura 7.8. Sección vivir con la enfermedad.....	114

# ÍNDICE DE TABLAS

Tabla 2.1. Defunciones por causa de cáncer de mama (INEGI, 2015). .....	13
Tabla 2.2. Comparación de modelos sobre minería de proceso (Orellana-García et al., 2016). .....	27
Tabla 3.1. Departamentos involucrados en el proceso de atención. ....	34
Tabla 3.2. Log de Eventos: cada línea representa un evento. ....	35
Tabla 3.3. Variables seleccionadas para la vista de datos minable. ....	35
Tabla 3.4. Indicadores clave de rendimiento (KPIs) .....	38
Tabla 4.1. Diagnósticos más comunes. ....	40
Tabla 4.2. Casos únicos por paciente. ....	40
Tabla 4.3. Casos por grupos de edad. ....	41
Tabla 4.4. Casos por diagnóstico. ....	41
Tabla 4.5. Georeferenciación casos por diagnóstico y grupo de edad. ....	42
Tabla 4.6. Casos especiales en bebés de hasta un año de edad. ....	48
Tabla 4.7. Detalle de clústers. ....	56
Tabla 4.8. Tiempo en días entre citas, cuyos pacientes tienen más de dos citas. ....	57
Tabla 4.9. Tiempo promedio entre citas mayores a 91.5 (3 meses). ....	58
Tabla 4.10. Criterios para referir a una unidad especializada de mama. NOM-041-SSA2-2011. ....	61
Tabla 4.11. Codificación de cáncer de mama según la Clasificación Internacional de Enfermedades. NOM-041-SSA2-2011. ....	61
Tabla 4.12. Métodos terapéuticos. NOM-041-SSA2-2011. ....	61
Tabla 4.13. Variables mínimas para el registro de casos de cáncer de mama. ....	62
Tabla 4.14. Equipo de integración mínimo .....	62

Tabla 4.15. Estándares de referencia. NOM-041-SSA2-2011.....	62
Tabla 4.16. Lista de actividades en el proceso de atención de cáncer de mama.....	65
Tabla 4.17. Semántica TLT; □ significa siempre, ◇ significa eventualmente, ○ significa enseguida de, (x ⊔ y) donde x debe esperar hasta y.....	68
Tabla 4.18. Variables consideradas para el diseño del modelo de proceso.....	68
Tabla 4.19. Ejemplo de log de eventos. ....	70
Tabla 4.20. Resumen logs de eventos. ....	70
Tabla 4.21. Actividades consideradas como cruciales. ....	83
Tabla 4.22. Indicadores para evaluar el desempeño y proceso según la NOM-041- SSA2-2011 .....	84
Tabla 4.23. Necesidades y recomendaciones que realizó la paciente hacia la institución .....	96

# 1. INTRODUCCIÓN

Los recursos informáticos y tecnología están teniendo una contribución muy importante en el proceso para evitar y reducir los errores clínicos, así como también en el proceso de mejora de procesos y flujos de trabajo de las instituciones de salud; la minería de datos ha mejorado varios campos de la vida del ser humano, basándose en patrones y tendencias descubiertas determinando nuevos conocimientos y proporcionando diagnósticos (Silachan y Tantasanawong, 2011).

La implementación de minería de datos ayuda a una reducción del tiempo de espera del paciente en términos de consulta y/o diagnósticos por parte del médico, lo que conlleva a una atención médica de calidad que es una de las claves del éxito de las instituciones de salud. Los métodos de minería de datos incorporan técnicas de estadística, inteligencia artificial, aprendizaje automático (machine learning) y sistemas de bases de datos (Gheorghe y Petre, 2014).

Desde hace varias décadas la tasa de cáncer de mama ha aumentado, siendo el de mayor incidencia entre las mujeres a nivel mundial y ocupa el primer lugar entre las causas de muerte (de la Vara-Salazar et al., 2011; Lorente et al., 2011; Guler, 2017; OMS). Según Ozkan et al. (2017) en su estudio, existen varios factores de riesgo biopsicosociales en el desarrollo del cáncer de mama; los resultados del estudio determinaron que el mayor riesgo de contraer cáncer de mama fue para las mujeres de los 40-59 años.

El incremento de cáncer de mama está vinculado a los estilos de vida, las transiciones demográficas y epidemiológicas; o bien a la falta de programas de detección precoz en países de bajos o medianos ingresos, donde la disponibilidad de servicios es de 35% mientras que en los países de ingresos altos es de 95% (de la Vara-Salazar et al., 2011). La Organización Mundial de la Salud (OMS) promueve programas nacionales para el control de cáncer de mama. Uno de ellos el control integral del

cáncer, que abarca la prevención, la detección precoz, el diagnóstico y tratamiento, la rehabilitación y los cuidados paliativos (OMS).

El objetivo de este capítulo es describir la institución objeto de estudio, así como la problemática a resolver, además se expone todo aquello relacionado con esta investigación como los objetivos, la hipótesis, delimitaciones y justificación del proyecto.

## 1.1. Presentación

La investigación se desarrolló en el Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado de Sonora (ISSSTESON), es un organismo público descentralizado, tiene como función otorgar prestaciones y servicios de seguridad social entre las cuales destacan pensiones y servicios de salud, beneficiando a 180,000 derechohabientes entre sus diferentes organismos afiliados.

El ISSSTESON ofrece el servicio médico bajo un modelo regionalizado y escalonado por tres niveles de atención.



**Figura 1.1.** Proceso de Atención Médica para un paciente con probabilidad de padecer cáncer de mama

La figura 1.1 muestra el proceso de atención Médica que sigue un paciente que tiene probabilidad de padecer cáncer de mama. El paciente llega a la institución por el nivel de atención uno, el cual ofrece servicios generales como atención dental, estudios de laboratorio clínico básico y radiografías simples; a nivel nacional se estima que el promedio de días naturales desde que un paciente con posible cáncer de mama sintió un síntoma hasta que consultó a un médico al respecto es de 30 días.

Cuando un paciente pasa al segundo nivel ya ha transcurrido tiempo desde la primera consulta, y en este nivel se puede recibir servicios de especialidades básicas como la hospitalización general, pruebas especiales de laboratorio clínico, radiología y ultrasonidos.

Por último, en el tercer nivel se encuentran los servicios de alta especialidad, es decir aquellas enfermedades que ponen el riesgo la vida como el cáncer de mama. Una vez que una paciente se encuentra en el nivel tres de atención con sospecha de cáncer de mama, el tiempo para ser diagnosticada es de 37 días naturales, si el diagnóstico es positivo, debe iniciar un tratamiento, y el promedio para iniciar un tratamiento después de la biopsia es de 30 días para las instituciones de salud pública.

Siendo el tiempo de atención un factor clave para la supervivencia del paciente, hace falta identificar posibles cuellos de botella que puedan estar afectando en los tiempos de atención, en cada uno de los niveles con que cuenta el ISSSTESON.

Se estima que el promedio de atención médica a nivel nacional para las instituciones de salud pública es de 139 días naturales. Lo anterior indica que el proceso actual de prestación de servicios sea lento, donde el tiempo transcurrido desde el primer nivel hasta llegar al nivel tres es crucial para un paciente con cáncer de mama, lo que puede agravar la salud del paciente.

Según la muestra proporcionada por ISSSTESON se han registraron 1650 casos desde enero 2013 a enero de 2018 y la cifra sigue una tendencia al alza. Tan solo en

el 2017, en Sonora se presentaron 76 defunciones de cáncer de mama. Las defunciones por esta enfermedad (Secretaría de Salud, 2013) han ido en aumento en los últimos años, pasando de 129 defunciones en el 2005 hasta 176 en 2013.

El ISSSTESON cuenta con bases de datos que contienen registros de los seguimientos a pacientes de cáncer de mama desde la última década, detectando la necesidad de análisis de dichos registros que abarcan el proceso de atención del cáncer desde la detección hasta el término del tratamiento. Al no contar con este análisis no se conoce si existe algún factor en la institución a lo largo del proceso de atención de cáncer de mama que pueda estar generando cuellos de botella.

## **1.2. Planteamiento del problema**

El cáncer de mama es de las principales causas de muerte en mujeres en México y en el mundo. Aunque existe un gran número de estudios sobre la detección y tratamiento del cáncer de mama, en México no se está analizando el proceso de atención de esta enfermedad a profundidad, lo que puede estar generando cuellos de botella en el tiempo de atención, desde la detección hasta que termina un tratamiento; siendo el tiempo un factor crucial para la supervivencia del paciente, existe la imperiosa necesidad de medir la eficacia de los procesos de inicio a fin para hacer posible una evaluación de la atención para su futura mejora.

En el ISSSTESON se desea llevar un mayor control de los procesos de atención del cáncer que permita evaluar los procedimientos administrativos desde la detección hasta la terminación del tratamiento, con el fin de mejorar la atención al paciente.

## **1.3. Objetivo general**

Desarrollar un modelo basado en análisis y minería de datos, que permita identificar áreas de oportunidad para establecer señalamientos y sugerencias de mejora en el proceso de atención de cáncer de mama.

## **1.4. Objetivos específicos**

- Analizar el proceso de atención de cáncer de mama desde la detección hasta terminar el tratamiento, haciendo uso de las tecnologías de la información, análisis y minería de datos.
- Desarrollar y evaluar un modelo que permita extraer conocimiento del proceso de atención de cáncer de mama desde la detección hasta el término del tratamiento.
- Elaborar informes con los señalamientos y sugerencias para mejora de la atención.

## **1.5. Hipótesis**

La implementación de técnicas de minería de datos ayudará a mejorar el proceso de atención de cáncer de mama desde su detección hasta el término de su tratamiento en la institución de salud.

## **1.6. Alcances y delimitaciones**

Limitaciones del proyecto:

- Existe información de derechohabientes que no se podrá revelar con fin de mantener la confidencialidad de los mismos.
- Los resultados de esta investigación son para proporcionar señalamientos y sugerencias de mejora para agilizar el proceso de atención de cáncer de mama.
- La información que se utilizará en esta investigación se centra en las unidades de atención de la ciudad de Hermosillo.
- Se utilizará información correspondiente al período 2013-2018.

## **1.7. Justificación**

El mayor beneficio de esta investigación será la mejora del proceso de atención del cáncer de mama, lo que podría resultar en un aumento en la probabilidad de supervivencia de las pacientes.

Adicionalmente, ayudará a la institución a agilizar procesos para una atención pronta, a obtener una radiografía de la cadena de atención de los pacientes con cáncer para tener una oportunidad de mejora. Por otro lado, una atención pronta puede facilitar la detección de cáncer de mama, lo que da la posibilidad de aplicar tratamientos que suelen ser más eficaces, menos complejos y menos costosos.

Así mismo, una persona que fue diagnosticada con cáncer de mama en fases tempranas y recibe un tratamiento eficaz a tiempo, puede seguir trabajando y seguir apoyando económicamente a su familia.

Es necesario el uso de los recursos tecnológicos con los que se cuenta actualmente para la mejora de procesos en áreas sensibles, como la salud, ayudando a proporcionar servicios médicos de calidad.

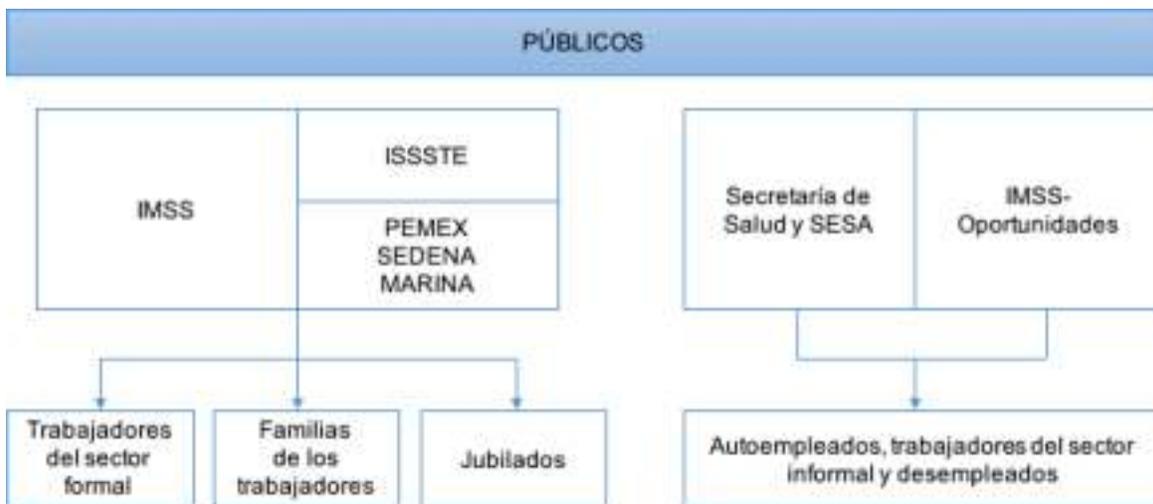
## **2. MARCO DE REFERENCIA**

En este capítulo se definen los temas que comprende este documento, tales como salud pública en México, cáncer de mama y minería de datos; posteriormente se presentan trabajos previos, seguidos por los casos más comunes dentro de la minería de datos. Esto con la finalidad de brindar al lector una visión amplia del contenido de este trabajo.

### **2.1. Salud Pública**

El Instituto Nacional de Salud Pública (INSP) creado en 1987 (Gómez-Dantés et al., 2011) tiene como compromiso ofrecer resultados de investigación a problemas relevantes de la salud pública para prevenir y controlar enfermedades, además de formar a profesionales de la salud (INSP, 2015).

La figura 2.1 muestra la estructura del sector público del sistema de salud de México, entre las se encuentran las instituciones de seguridad social que prestan servicios a los trabajadores del sector formal de la economía, tales como el Instituto Mexicano del Seguro Social (IMSS), Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado (ISSSTE), Petróleos Mexicanos (PEMEX), Secretaría de la Defensa (SEDENA), Secretaría de Marina (SEMAR) entre otros; dentro de las instituciones y programas que atienden a la población sin seguridad social se encuentra la Secretaría de Salud (SSa), Servicios Estatales de Salud (SESA), Programa IMSS-Oportunidades (IMSS-O), Seguro Popular de Salud (SPS) (Gómez-Dantés et al., 2011).



*Figura 2.1. Estructura Sector Público (Gómez-Dantés et al., 2011).*

## 2.2. Cáncer de Mama

El cáncer de mama ha ido en aumento, el principal causante de muerte entre las mujeres a nivel mundial (de la Vara-Salazar et al., 2011; Lorente et al., 2011; Guler, 2017; OMS). La investigación de Ozkan et al. (2017) menciona que, en el desarrollo del cáncer de mama, existen varios riesgos biopsicosociales y, además encontró que las mujeres en el rango de edad de los 40-59 años tienen mayor riesgo de contraer esta enfermedad. Además, Vara-Salazar et al., (2011) comentan que los estilos de vida, las transiciones demográficas y epidemiológicas están vinculadas al incremento de cáncer de mama; o bien según la OMS la falta de programas de detección precoz en países de bajos o medianos ingresos incide en el incremento, siendo la disponibilidad de servicios del 35% y, 95% para los países con altos ingresos.

Según la OMS la incidencia de esta enfermedad varía en todo el mundo, las tasas normalizadas por edad de hasta 99.4 por cada 100,000 personas en América del Norte, mientras que en Europa oriental, América del Sur, África austral y Asia occidental presentan incidencias moderadas, pero en aumento. La incidencia más baja se da en la mayoría de los países africanos, pero también en ellos se observa un incremento de la incidencia de cáncer de mama.

Las tasas de supervivencia del cáncer mamario varían en todo el mundo, desde el 80% o más en América del Norte, Suecia y Japón, pasando por un 60% aproximadamente en los países de ingresos medios, hasta cifras inferiores al 40% en los países de ingresos bajos (Coleman et al., 2008).

La OMS promueve programas nacionales para el control de cáncer de mama. El control integral del cáncer abarca la prevención, la detección precoz, el diagnóstico y tratamiento, la rehabilitación y los cuidados paliativos. Adicionalmente precisa que en los países donde el ingreso es bajo o mediano implementen estrategias que aborden el problema de manera eficaz y eficiente.

Etapas del control integral:

- Según la OMS el primer paso del control integral del cáncer es la prevención, se promueve siguiendo una alimentación saludable, agregando actividad física, control en el consumo de alcohol, manejar niveles equilibrados de peso, siguiendo estas recomendaciones se podría llegar a tener un efecto de reducción de la incidencia de cáncer de mama a largo plazo.
- La detección precoz es otra etapa del control de cáncer de mama, según Anderson et al. (2008), aunque existan medidas de prevención no es suficiente, la atención a tiempo es la piedra angular de la detección del cáncer de mama. Existen dos métodos de detección de la enfermedad: el examen clínico de las mamas y la mamografía (Lorente et al., 2011). Según Yip et al. (2008), la planificación de un programa bien organizado y sostenible que garantice la coordinación, la continuidad y calidad de intervenciones son aspectos esenciales para el éxito de la detección precoz.
- El diagnóstico precoz, es una importante estrategia de diagnóstico, especialmente en países de bajos ingresos, donde la enfermedad es diagnosticada en fases avanzadas y donde los recursos son limitados. Algunos datos sugieren que el diagnóstico precoz puede dar lugar a un "descenso del estadio TNM" (aumento de la proporción de cánceres de mama detectados en

una fase temprana) de la enfermedad, que la haría más vulnerable al tratamiento curativo (Yip et al., 2008).

- La rehabilitación en cuidados paliativos es la posibilidad de brindar ayuda experimentada a una persona y su familia para reajustarse a una situación que es estática y que permanecerá así por semanas (Porta i Sales, 2017).

Según Breast Cancer (2015), para describir los detalles de su cáncer, como el tamaño del tumor y si se ha extendido se conoce como estadios, donde suele expresarse con un número entre 0 y IV. El estadio 0 corresponde al cáncer no invasivo que permanece en su ubicación original y el estadio IV al cáncer invasivo que se propaga fuera de la mama hacia otras partes del cuerpo.

Adicionalmente según American Cancer Society (2018), existen ciertos tipos de cánceres que parecen repetirse en algunas familias, puede deberse a que los miembros están expuestos a los mismos factores o tienen hábitos similares que aumentan el riesgo de padecer la enfermedad. En otros casos, los cánceres que se presentan con mayor frecuencia en algunas familias pueden ser causados por un gen anormal que se transmite de una generación a otra, aunque no es el cáncer lo que se hereda, sino el gen anormal que puede dar lugar al cáncer. Se cree que solo alrededor del 5 al 10 % de todos los casos de cáncer se originan directamente a partir de defectos genéticos (llamados mutaciones) heredados de uno de los padres. Se puede tratar de un síndrome de cáncer familiar si presentan ciertas características tales como la aparición de más de un tipo de cáncer en una misma persona (como una mujer con cáncer de mama y cáncer de ovario).

En las últimas décadas la detección del cáncer de mama es cada vez más temprana gracias a los protocolos introducidos. Por ejemplo, en Estados Unidos en 1995, los casos entre el estadio 0 y 1 se encontraban entre el 56% de los casos, mientras que en 1985 eran de tan solo 45%. Sin embargo, se han detectado casos recidivas (volvió) en un 5 a 40%, que probablemente se debe a una disección axilar inadecuada, una

técnica quirúrgica incompleta o un tratamiento sistémico subóptimo (Stankov et al., 2012). Las demoras de cáncer de mama se han asociado con una menor supervivencia: un tiempo superior a tres meses entre la aparición de síntomas y el inicio del tratamiento disminuye la supervivencia global en 12%. La mayor demora fue para el diagnóstico donde un tercio de las mujeres manifestó que fueron necesarias dos o más consultas por los síntomas iniciales, que esto puede estar relacionado al desconocimiento del médico que tiene el primer contacto con la paciente, como se estableció en Canadá. Mientras que, en Estados Unidos, la causa más frecuente de las demoras son los hallazgos iniciales al examen, no eran significativos para los médicos. En nuestro medio, prácticamente ninguna escuela de medicina tiene cátedra de cáncer, factor que hace más difícil que un médico general, puerta de entrada al sistema, sospeche este diagnóstico y sepa cómo orientarlo adecuadamente (Piñeros et al., 2011).

La demora en iniciar un tratamiento podría reducir la supervivencia en pacientes con cáncer de mama (Ángeles-Llerenas et al. 2016; Allemani, et al., 2015; McLaughlin et al., 2012; Flores-Luna et al. 2008), según Ángeles-Llerenas et al. (2016), el porcentaje de supervivencia para América Latina va de 66 a 78%. Según Flores-Luna et al. (2008), observan que existe un 59.9% de supervivencia a cinco años, los casos fueron detectados en la primera etapa (estadio clínico I) tuvieron una supervivencia mayor de 82%, mientras que los pacientes cuyo caso fue detectado en la última etapa tuvieron un 15% de supervivencia (estadio clínico IV); adicionalmente Richards et al. (1999), mencionan que los pacientes que inician un tratamiento 3 a 6 meses después de los primeros síntomas tuvieron una supervivencia a cinco años significativamente más baja que las que esperaron menos de 3 meses.

Así mismo McLaughlin et al. (2012) comentan que, la demora en iniciar un tratamiento podría reducir la supervivencia en mujeres con cáncer de mama. La pronta atención es crucial, si los retrasos en la atención médica son elevados la salud de la paciente

puede deteriorarse, según Ángeles-Llerenas et al. (2016), existen 4 variables de tiempo a considerar:

- A. tiempo total en días naturales desde el primer síntoma hasta el inicio del tratamiento;
- B. tiempo en días naturales desde el primer síntoma hasta la consulta con un médico acerca de los síntomas;
- C. el tiempo en días naturales desde la recepción de los resultados de la mamografía hasta los resultados de la biopsia diagnóstica, y
- D. el tiempo en días naturales desde la biopsia hasta el inicio del tratamiento.

Adicionalmente Ángeles-Llerenas et al. (2016) comentan que los retrasos en la atención médica pueden ser por 2 motivos:

- 1) retraso del paciente, que surge de un retraso en la búsqueda de atención médica después del autodescubrimiento de un posible síntoma de la enfermedad, y
- 2) retraso del sistema, que incluye, el tiempo para obtener citas, programar pruebas de diagnóstico, recibiendo un diagnóstico definitivo e iniciar el tratamiento.

En el Instituto Nacional de Estadística y Geografía (INEGI) se puede encontrar estadísticas de enfermedades de mortalidad en México como se muestra en la tabla 2.1, además se ve en la parte sombreada las defunciones por la causa de cáncer de mama, las cuales han ido en aumento en los últimos años (INEGI, 2015).

	2011	2012	2013	2014	2015
<b>Total</b>	<b>590,693</b>	<b>602,354</b>	<b>623,599</b>	<b>633,641</b>	<b>655,688</b>
CIE-10/2	590,693	602,354	623,599	633,641	655,688
Ciertas enfermedades infecciosas y parasitarias (A00-B99)	17,734	18,003	18,882	16,780	16,214
Tumores (neoplasias) (C00-D48)	76,215	78,352	80,551	82,830	85,201
Tumores malignos del labio, de la cavidad bucal y de la faringe (C00-C14)	1,039	1,081	1,105	1,197	1,196
Tumores malignos de los órganos digestivos (C15-C26)	22,091	23,132	24,033	24,883	25,870
Tumores malignos de los órganos respiratorios e intratorácicos (C30-C39)	7,752	7,482	7,842	7,796	7,990
Tumores malignos de los huesos y de los cartílagos articulares (C40-C41)	826	850	824	1,029	1,083
Melanoma y otros tumores malignos de la piel (C43-C44)	1,454	1,529	1,566	1,624	1,763
Tumores malignos de los tejidos mesoteliales y de los tejidos blandos (C45-C49)	1,189	1,158	1,244	1,437	1,448
Tumor maligno de la mama (C50)	5,258	5,663	5,597	6,051	6,304
(C500) Tumor maligno del pezón y aréola mamaria	2	3	3	3	5
(C501) Tumor maligno de la porción central de la mama	1	1	3	1	
(C503) Tumor maligno del cuadrante inferior interno de la mama			1		1
(C504) Tumor maligno del cuadrante superior externo de la mama			1	1	
(C505) Tumor maligno del cuadrante inferior externo de la mama					1
(C508) Lesión de sitios contiguos de la mama	5	4		5	8

**Tabla 2.1.** Defunciones por causa de cáncer de mama (INEGI, 2015).

Adicionalmente según la NOM-041-SSA2-2011 (2011), la tendencia es ascendente en todos los grupos de edad, con un incremento mayor en las mujeres mayores de 60 años, en comparación con las mujeres de 40 a 59 años. Se estima que la incidencia y la mortalidad seguirán aumentando de manera importante debido al envejecimiento poblacional, a los cambios en los patrones reproductivos, a una mayor prevalencia de los factores de riesgo y a los problemas para el acceso oportuno a la detección, el diagnóstico temprano y el tratamiento adecuados.

Por lo anterior en México existe la Norma Oficial Mexicana NOM-041-SSA2-2011, que tiene por objetivo la prevención, diagnóstico, tratamiento, control y vigilancia epidemiológica del cáncer de mama que debe ser el instrumento que oriente el desarrollo del programa de cáncer de mama de manera sistemática y sostenida. Los

lineamientos que presenta esta norma es establecer las bases para que las diferentes instituciones que conforman el Sistema Nacional de Salud organicen sus recursos a favorecer estilos de vida saludables en las mujeres mexicanas para disminuir los riesgos de esta enfermedad, además de vigilar de manera constante los factores de riesgo, con el fin de prevenir o detectar tempranamente esta patología y contribuir a la disminución de la morbilidad y mortalidad por esta causa. Existen estudios que confirman que, en los países desarrollados, un programa de tamizaje organizado, es decir son esenciales servicios de alta calidad en todos los procesos desde la detección hasta el tratamiento y rehabilitación, lo anterior tiene el potencial de reducir entre el 20% y 40% de tasa de mortalidad y la carga de la enfermedad en la población en riesgo (NOM-041-SSA2-2011, 2011).

Según la NOM-041-SSA2-2011 (2011), se sigue un algoritmo para la valoración de nódulos detectados por mastografía (figura 2.2):



**Figura 2.2.** Algoritmo para la valoración de nódulos detectados por mastografía

### 2.3. Aprendizaje Automatizado

Anteriormente el aprendizaje automatizado (Machine Learning) está definido como algoritmos que aprenden de forma autónoma partiendo de un histórico, con el fin de hacer predicciones futuras. Es una rama de la inteligencia artificial, ya que su núcleo

se basa en software avanzado de reconocimiento de patrones para adaptarse a las nuevas circunstancias para detectar y extrapolar patrones (López-Ramírez, 2017).

## **2.4. Minería de Datos**

La minería de Datos forma parte del Descubrimiento de Conocimiento en Base de Datos (por sus siglas en inglés KDD), el cual es definido (Oded y Lior (2005); Fayyad, Piatestky-Shapiro, y Smyth, (1996); Morgado-García, Antonio Ponce-de-León-Lima y Rosete-Suárez, (2017)) como un proceso metodológico para encontrar un “modelo” válido, novedoso, potencialmente útil, capaz de describir patrones dentro de los datos contenidos en algún repositorio. El proceso está dividido en 5 fases:

1. selección de datos,
2. preprocesamiento,
3. transformación,
4. data mining,
5. interpretación y evaluación.

La minería de datos es el proceso de extraer conocimiento de grandes cantidades de datos, que consta de varios pasos: extracción y recolección, limpieza e integración, selección y transformación de datos, extracción de conocimiento y finalmente visualización y comunicación (Han, Kamber, y Pei, (2011); Chiotti y Monserrat). Los métodos de minería de datos incorporan técnicas de estadística, inteligencia artificial, aprendizaje automático (machine learning) y sistemas de bases de datos (Gheorghe y Petre, 2014; López-Ramírez, 2014).

Según Zaki y Meira (2014) la minería de datos está dividida en dos tareas: tareas de predicción y tareas descriptivas. Las tareas descriptivas se utilizan para encontrar patrones ocultos en grandes conjuntos de datos así ayudar a la toma de decisiones, para ello utiliza diversas técnicas como asociación, agrupaciones, entre otras. Mientras que la tarea predictiva tiene como objetivo la predicción, lográndolo mediante la

construcción de modelos haciendo uso de conjuntos de reglas, árboles de decisión, redes neuronales y vectores de soporte, por mencionar algunos; siendo esta tarea la más común de la minería de datos y es la que cuenta con la mayoría de las aplicaciones de las empresas.

La minería de datos predictiva tiene tres etapas, como se muestra en la figura 2.3, comenzando con las tareas de colección y almacenamiento de los datos, incluyendo la identificación de las características relevantes de una empresa y el establecimiento de un archivo de almacenamiento para documentarlas (Zaki y Meira, 2014). La primera etapa llamada procesamiento de datos se refiere a la selección, limpieza, el enriquecimiento, la reducción y la transformación de las bases de datos (Camacho-Centellas, 2016). La etapa dos, llamada predicción consiste en la selección y/o reducción de características, construir modelos para la predicción, y finalmente la etapa tres es la interpretación y/o despliegue (Zaki y Meira, 2014).

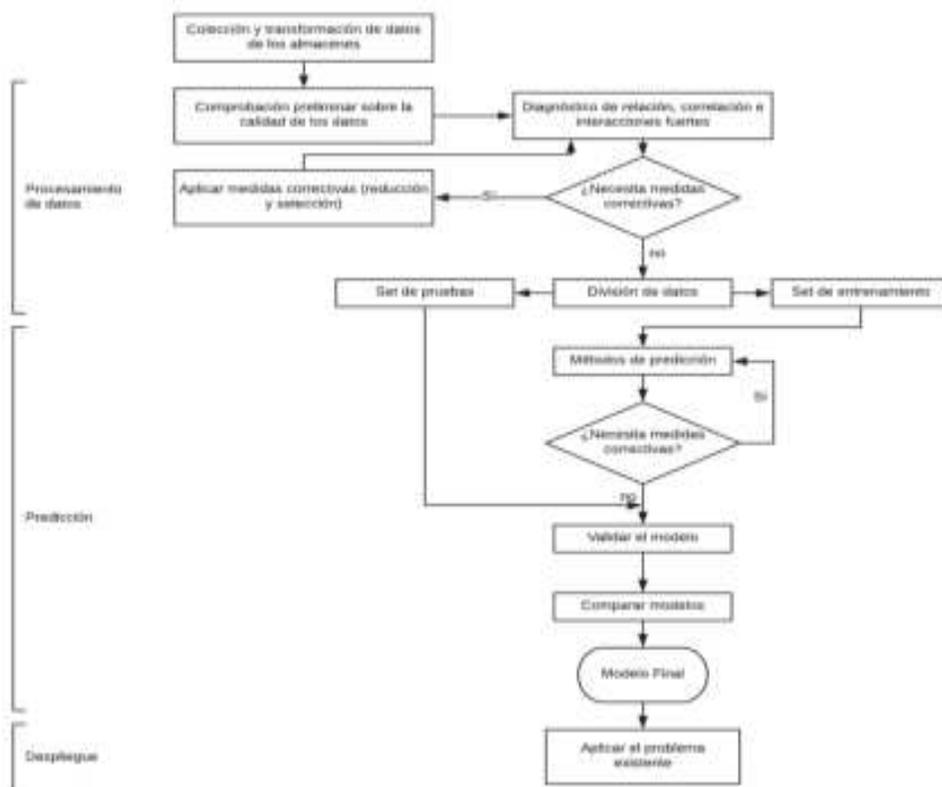


Figura 2.3. Etapas de la minería de datos según López-Ramírez (2017) pp. 40.

### **2.4.1. Clasificación y Predicción**

Según Zaki y Meira (2014), consiste en predecir un resultado determinado en base a una entrada dada, donde el algoritmo procesa un conjunto de entrenamiento que contiene un conjunto de atributos y el resultado, generalmente llamado meta o atributo de predicción. El algoritmo intenta descubrir relaciones entre los atributos que harían posible predecir el resultado, posteriormente al algoritmo se le proporciona un conjunto de datos que no ha visto antes, llamado conjunto de predicción. Para definir la efectividad del algoritmo se compara el número de predicciones realizadas contra el número total de predicciones.

### **2.4.2. Adquisición de Datos**

Existe la necesidad de registrar mediciones y procesar los datos recogidos para su visualización, ya que aportan beneficios para las empresas, por lo que se deben definir qué datos se quieren recopilar y qué método o procedimiento se va a usar para ello (Zaki y Meira, 2014).

### **2.4.3. Preparación de Datos**

Los datos son los más importante para todos los sistemas de aprendizaje supervisado, pero por sí solos no sirven de nada, sino hasta que se extraen conocimientos o inferencias de ellos. Los datos brutos no siempre son los mejores para el análisis, sino que deben ser pre procesados o preparados y transformados para poder explotar de una manera óptima. Para la minería de datos es importante la preparación de los datos y existen diversas técnicas para esta tarea (Zaki y Meira, 2014).

### **2.4.4. Minería de Procesos**

La minería de procesos por sus siglas en inglés PM tiene como objetivo la extracción de conocimiento de los registros de eventos almacenados, donde éstos pueden originarse a partir de diversos tipos de sistemas como lo puede ser los sistemas

genéricos de información empresarial, así como los sistemas de información hospitalaria (HIS) (Rovani et al., 2015). Según (Rudnitckaia, 2015) estas técnicas proporcionan nuevos medios para descubrir, monitorear y mejorar procesos en una variedad de dominios de aplicaciones, donde existen dos factores principales para el creciente interés en la MP:

1. Se registran cada vez más eventos que proporcionan información detallada sobre el historial de procesos.
2. En entornos altamente cambiantes existe la necesidad de mejorar y respaldar los procesos comerciales.

Adicionalmente (Rudnitckaia, 2015) comenta que PM proporciona un puente entre BI (Business Intelligence) y BPM (Back Propagation Network), Minería de datos y workflow. Incluye el descubrimiento de procesos, verificación de conformidad, redes sociales/minería organizacional, construcción automatizada de modelos de simulación, extensión de modelos, reparación de modelos y predicción de casos y recomendaciones basadas en la historia (figura 2.4).



**Figura 2.4.** Posición de la minería de procesos según Van der Aalst, W. M. (2016).

## **2.5. Tecnologías de la Información en Áreas de la Salud**

Los recursos informáticos y tecnología están teniendo una contribución muy importante en el proceso para evitar y reducir los errores clínicos, así como también en la mejora de procesos y flujos de trabajo de las instituciones de salud; la minería de datos ha mejorado varios campos de la vida del ser humano, basándose en patrones y tendencias descubiertas determinando nuevos conocimientos y proporcionando diagnósticos (Silachan y Tantasanawong, 2011). La investigación aplicada que trabaja en el descubrimiento de patrones de comportamiento de datos en diferentes áreas de la medicina se conoce como informática biomédica (Chiotti y Monserrat).

### **2.5.1. Minería de Datos como herramienta en el sector Salud**

La implementación de minería de datos ayuda a una reducción del tiempo de espera del paciente en términos de consulta y/o diagnósticos por parte del médico, lo que conlleva a una atención médica de calidad que es una de las claves del éxito de las instituciones de salud (Gheorghe y Petre, 2014). Otro beneficio de la minería de datos en estas instituciones es la identificación de tratamientos efectivos y mejores prácticas (Koh y Tan, 2005).

Según en el estudio de (Koh y Tan, 2005) las aplicaciones de minería de datos en la salud pueden tener un gran potencial y utilidad. Sin embargo, el éxito depende de la disponibilidad de datos de salud limpios, es fundamental que las instituciones de atención médica consideren la mejor forma de capturar, almacenar, preparar y extraer los datos. Se debe construir una estandarización del vocabulario clínico y del intercambio de datos entre organizaciones para mejorar los beneficios de las aplicaciones de minería de datos. Además, dado que los datos contenidos en estas instituciones no están limitados a tan solo datos legales, como notas de médicos o registros clínicos, es necesario explorar también el uso de minería de textos para ampliar el alcance de la minería de datos de la salud puede hacer actualmente.

Por lo anterior las organizaciones que realizan sus propios depósitos de datos, la minería de datos y análisis están mejor posicionadas para cumplir con las necesidades a largo plazo de los planes de salud. El análisis se puede refinar y personalizar rápidamente ayudando a facilitar la transparencia de los datos y una toma de decisiones más racional (Koh, H.C. y Leong, 2001; Benko y Wilson, 2003).

Existen algunos casos de estudio en la medicina general como los que se presentan a continuación:

Según Singh (2015) el paso más importante desde el punto de vista del tratamiento clínico, y posible recuperación, es el diagnóstico precoz de cáncer de mama, ya sea realizado por autoexamen, mamografía o resonancia magnética, por lo que propone un modelo de predicción de cáncer de mama usando algoritmos de clasificación de minería de datos basado en varios parámetros, incluidos los síntomas de cáncer de mama, mutación genética y otros factores de riesgo de la enfermedad. El resultado obtenido es que predice con precisión la posibilidad de enfermedad de cáncer de mama ya que verifica la mutación en el gen y simula la relación entre varios factores de riesgo y síntomas, ayudando a investigadores a cumplir las normas que se centran en la identificación de la mutación genética que causa esta enfermedad.

Otro estudio fue realizado con el fin de detectar tempranamente el cáncer de mama mediante el uso de un método de clasificación de minería de datos, el método utilizado llamado Adaboost-SVM que puede ser mejorado añadiendo gradualmente nuevas características al clasificador. Como resultado mostró que el enfoque propuesto es mejor que los enfoques convencionales ayudando a los médicos a mejorar la precisión del diagnóstico y reducir la tasa de detección falsa (Peng et al., 2014).

En Alabama las infecciones adquiridas en un hospital matan a unas 90,000 personas cada año, por lo que surge la necesidad de encontrar estos brotes con el fin de reducir el número de infecciones. El sistema implementado llamado Sistema de Vigilancia de Minería de Datos (DMSS) que encuentra patrones en datos médicos que son

demasiado sutiles o complejos para que los humanos los detecten. El período de análisis fue de 11 meses en el cual se encontraron 41 brotes y tiene planeado convertirse en un método de vigilancia nacional para detectar brotes de enfermedades infecciosas transmitidas por los alimentos (Kreuze, 2001).

Otra investigación fue que el gobierno de Singapur en 1999 introdujo la financiación basada en el uso de CaseMix a los hospitales públicos esperando que el financiamiento de la atención médica produzca beneficios significativos. El estudio menciona que la minería de datos es implementada en el contexto de CaseMix y que, además, complementa los métodos estadísticos, permitiendo a los hospitales comprender su comportamiento para conducir a servicios más rentables adicionalmente se ayuda a comprender los datos de readmisión y detectar patrones sistemáticos, si los hay (Koh y Leong, 2001).

Adicionalmente la minería de datos en el sector sanitario ha ofrecido una clara ventaja competitiva al utilizar Business Intelligence. A través de la aplicación, ha sido capaz de identificar tendencias y correlaciones previamente desconocida, se está convirtiendo rápidamente popular en las industrias de uso intensivo de datos, como financiero, minorista y manufactura (Veletsos, 2003).

Las predicciones hechas en el sector salud mediante software SAS Enterprise Miner forma parte de otro estudio, y que proporciona una suite de herramientas para descubrir cómo los genes interactúan en su promoción o la prevención de enfermedades, incluyendo redes neuronales, árboles de decisión, ensambles, logística y de regresión de mínimos cuadrados. SAS se refiere a la metodología utilizada en la minería de datos como el SEMMA facilitando la aplicación estadística exploratoria y las técnicas de visualización, seleccionar y transformar las variables predictivas (Thompson y Roberson, 2000).

Otra implementación de la minería está en el caso de Chiotti y Monserrat, tiene como propósito descubrir patrones de consumo de medicamentos que generen conocimiento útil para la toma de decisiones. Se utilizó el proceso de minería de datos para descubrir conocimientos utilizando varios algoritmos de segmentación (“clustering”), donde los resultados más significativos se obtuvieron con el algoritmo k-means, además de utilizar un algoritmo de árbol de clasificación. Los resultados obtenidos explican el 80% del consumo total de medicamentos por sexo y por estación del año, divididos en varias clases como: medicamentos para el sistema cardiovascular, para el tracto alimentario y metabólico, sistema nervioso, anti-infecciosos y medicamentos para el sistema respiratorio.

Otro caso práctico de minería de datos aplicado en el sector salud se encuentra en la investigación de Uma y Savita (2017), cuyo objetivo fue evaluar con qué precisión predecir la recurrencia del cáncer de mama utilizando algoritmos de minería de datos; donde la detección temprana de esta enfermedad puede proporcionar una ventaja potencial en el tratamiento de ésta. Se utilizaron 4 algoritmos de clasificación y 4 algoritmos de agrupamiento donde el resultado indicó que el árbol de decisión C5.0 y SVM son mejores predictores de cáncer de mama con un 81% de precisión, mientras que c-means tienen la menor precisión con un 37% entre los otros algoritmos utilizados en la investigación.

Otra investigación es la de Chandra, Patel, y Chandra (2017), en el cual comentan que con el uso de diferentes técnicas como clasificación Naïve Bayes y Weka como herramienta de aprendizaje automático, árboles de decisión y random forest (bosques aleatorios) se puede predecir el cáncer oral; los resultados obtenidos reflejan que Naïve Bayes tuvo un alto grado de precisión para predecir esta enfermedad.

Por último, las compañías farmacéuticas también pueden beneficiarse de los sistemas de Gestión de las Relaciones con los Clientes (por sus siglas en inglés CRM) de atención médica y minería de datos. Al rastrear qué médicos prescriben qué

medicamentos y para qué fines, las compañías farmacéuticas pueden decidir a quién orientar, mostrar cuál es el plan de tratamiento más económico o menos costoso para una dolencia, ayudar a identificar a los médicos cuyas prácticas son adecuadas para ensayos clínicos específicos (para ejemplo, médicos que tratan a un gran número de un grupo específico de pacientes), y mapee el curso de una epidemia para apoyar a vendedores farmacéuticos, médicos y pacientes. Las compañías farmacéuticas también pueden aplicar la extracción de datos a grandes cantidades de datos genómicos para predecir cómo la composición genética de un paciente determina su respuesta a una terapia farmacológica (Koh y Tan, 2005).

## **2.6. Usos más comunes de la minería de datos**

Van der Aalst (2016), menciona que hay diferentes fuentes y están generando grandes volúmenes de datos. Por lo que el análisis de grandes cantidades de datos es una necesidad, ya que las empresas son beneficiadas al recopilar y analizar sus datos: los hospitales pueden detectar tendencias anomalías en sus registros de pacientes, las empresas de ciberseguridad y detección de intrusos en la red informática, supervisión del consumo de energía de los electrodomésticos; análisis de patrones en bioinformática y datos farmacéuticos; datos financieros y de inteligencia empresarial; detectar tendencias en blogs, Twitter, sólo por mencionar algunos (Han et al., 2011). La minería ha incursionado en diversas áreas de la vida del ser humano como se muestra en la figura 2.5.

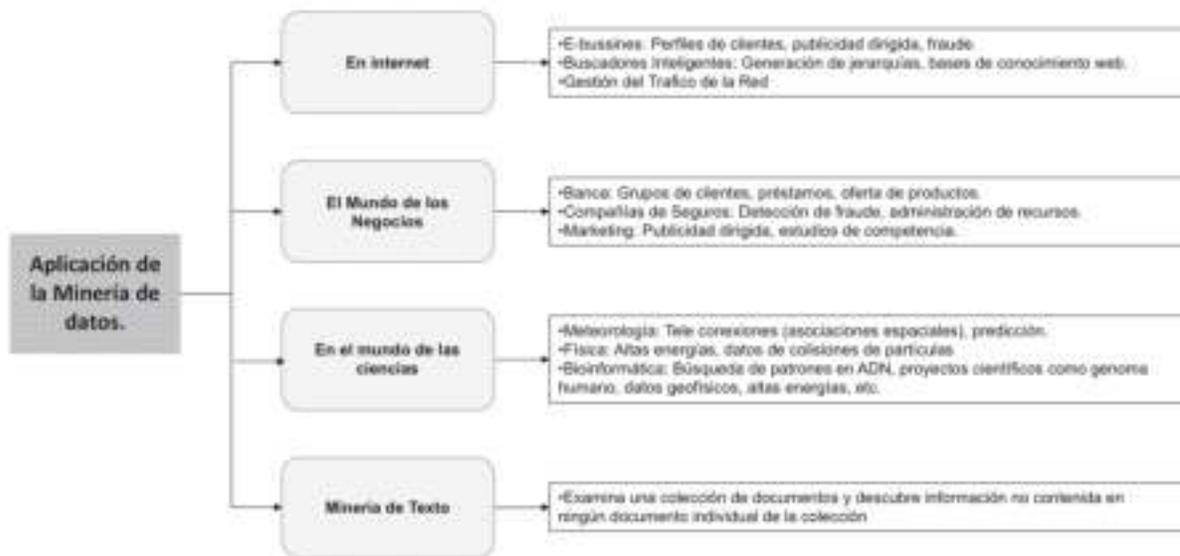


Figura 2.5. Aplicación de la minería de datos (Camacho-Centellas, 2016).

## 2.7. Algoritmos y técnicas de Minería de Procesos

La finalidad de los algoritmos y técnicas de procesos es obtener modelos que apoyen a los analistas, auditores y directivos de las empresas en sus actividades de chequeo, análisis, mejora y toma de decisiones de los procesos de su entidad (Torres-Sakipova, 2014).

### 2.7.1. Algoritmos de descubrimiento

Según Torres-Sakipova (2014) existen varios algoritmos los cuales se describen a continuación:

- $\alpha$ : creado por van der Aalst en 2004, fue de los primeros algoritmos de descubrimiento, está basado en la captura, desde un registro de eventos, de las relaciones siguientes usando el sistema binario. Una de las características es que asume que el registro de eventos está completo, por lo que presenta problemas para lidiar con el ruido, además, no tiene en cuenta las frecuencias de las actividades. A partir de este algoritmo surgieron el  $\alpha^+$  y el  $\alpha^{++}$ , que incorpora el tratamiento a ciclos cortos. Posteriormente en el 2009, surge el algoritmo  $\beta$ , que puede distinguir concurrencia y ciclos cortos.

- Minería heurística: Creado en 2006 por Weijters y otros, toma en cuenta las frecuencias de los eventos, es usado para derivar las relaciones causales y de paralelismo, pero no incorpora los eventos poco frecuentes al modelo, provocando que no se presenten caminos que se ejecutan con poca frecuencia, pero pudieran ser importantes.
- Minería genética: Presentado en el 2006, simula la evolución en la naturaleza. Parte de una población inicial de individuos a la que se le realizan varias iteraciones del algoritmo genético acercándose cada vez más a la solución apropiada.
- Minería difusa: creado por Günther y van der Aalst, está basado en medidas de significancia y correlación para visualizar el comportamiento en los registros de eventos
- Otro algoritmo es el inductivo según Bottrighi et al. (2018), divide el registro de eventos recursivamente en sublogs. Los sublogs se descomponen hasta que se refieren a una sola actividad.

### **2.7.2. Técnicas de descubrimiento**

Torres-Sakipova (2014) comenta en su investigación que se debe sacrificar la precisión para lograr niveles de abstracción y obtener modelos más entendibles, es decir, reduciendo el alcance del modelo se simplifica la conducta del modelo, contribuyendo significativamente a la reducción de la complejidad del mismo.

Adicionalmente Torres-Sakipova (2014), comenta que la descomposición de procesos en árboles de bloques de construcción permite agrupar las tareas automáticamente y resalta los aspectos más significativos del proceso en cada momento. El descubrimiento de los subprocesos que componen el proceso analizado, sus dependencias y correlaciones, permiten una mayor precisión en el diagnóstico realizado, en la figura 2.6 se observa el árbol de descomposición en subprocesos.

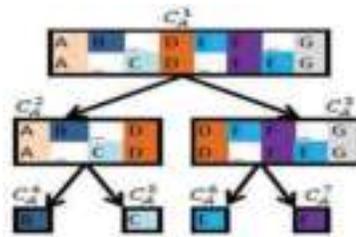


Figura 2.5. Árbol de descomposición en subprocesos (Torres-Sakipova, 2014).

## 2.8. Comparativa de técnicas de Minería de Procesos

Las técnicas de minería de procesos se utilizan en diversos sectores tales como: la industria, los servicios web, la inteligencia de negocios y la salud. La investigación de Orellana-García et al. (2016), recopiló información sobre los principales modelos en el tema de minería de procesos para aplicar técnicas en el descubrimiento, chequeo de conformidad y mejoramiento de los procesos y muestra una comparativa entre las distintos modelos. En función de seleccionar los elementos y características útiles para el entorno hospitalario se realizó un análisis de los modelos, donde se resume en la tabla 2.2:

Características	Modelos				
	Jans	Giessel	Rozinat	Bozkaya	van der Aalst
Diagnóstico preliminar de los datos		SI		SI	SI
Extraer registro de eventos desde la fuente de datos	SI	SI	SI	SI	SI
Generar el flujo de actividades	SI	SI	SI	SI	SI
Caracterización del proceso de negocio			SI	SI	SI
Análisis a partir de tipo de proceso				SI	SI
Obtener modelo integrado	SI				SI
Utilización de técnicas	SI	SI	SI	SI	SI
Análisis de rendimiento		SI	SI	SI	SI
Análisis de roles	SI	SI	SI	SI	
Análisis con el cliente				SI	SI
Análisis desde la perspectiva organizacional	SI	SI	SI	SI	SI
Análisis desde la perspectiva de control de flujo	SI	SI	SI	SI	SI
Análisis desde la perspectiva de tiempo	SI	SI	SI	SI	SI
Análisis desde la perspectiva de caso				SI	SI

Otras perspectivas	SI	SI	SI
Procedimiento para analizar el proceso			SI
Elaborado para un dominio de aplicación	SI		
Retroalimentación de los resultados			SI
Herramientas para el análisis	SI		
Considera profesionales no expertos			
Especificación de las técnicas a utilizar		SI	SI

**Tabla 2.2.** Comparación de modelos sobre minería de proceso (Orellana-García et al., 2016).

## 2.9. Trabajos previos

La investigación de Bottrighi et al. (2018) hace una aportación al estado del arte cuya evaluación experimental fue dirigida al manejo del accidente cerebrovascular produciendo resultados satisfactorios. Los registros de eventos guardan las secuencias de las acciones que se han completado en la organización, esto constituye una fuente de conocimiento de experiencias que son capaces de soportar tres tareas y cuyo objetivo final es mejorar la calidad o la eficiencia de los procesos:

- 1) Descubrimiento del modelo de proceso
- 2) Recuperación de ruta desde el modelo de proceso
- 3) Recuperación del seguimiento

Aportando al estado del arte en cada una de estas tareas esta investigación tiene como enfoque el descubrimiento del modelo de proceso, partiendo del registro de eventos que permite comprender y analizar los procedimientos realmente implementados en una organización, identificando: cuellos de botella, las diferencias con respecto a un estándar existente y el cumplimiento de las normas comerciales con el objetivo de mejorar su corrección y eficacia en la práctica. El enfoque está centrado en el descubrimiento del modelo de proceso, en el cual se propone ayudar al analista a identificar un buen equilibrio entre el sobreajuste y el ajuste insuficiente. Para lograr

ese objetivo, se diseñó una herramienta llamada SIM (Semantic Interactive Miner) que analiza el modelo actual, y con abstracción semántica para construir modelos cada vez más generalizados. Esta herramienta se comparó con dos conocidos algoritmos de minería de procesos: el minero inductor y el minero heurístico, donde el resultado fue innovador, dando la capacidad de facilitar que el analista utiliza directamente su conocimiento de dominio para conducir el descubrimiento del modelo de proceso, una característica que puede ser extremadamente ventajosa en aplicaciones ricas en conocimiento. Como investigaciones futuras se planea llevar a cabo una evaluación más extensa, centrándose en contextos de cuidado de la salud, donde los expertos de dominio (médicos) tienen un "profundo" conocimiento, por lo que se espera que este estudio, proporcione a los expertos las herramientas adecuadas para analizar el modelo y que ayude a proporcionar avances críticos con respecto a los tradicionales enfoques.

En otra investigación (Orellana-García et al., 2016) comentan que el sector salud enfrenta retos a diario por lo que hace necesario incorporar nuevas técnicas y herramientas que permitan mejorar el resultado de la actividad hospitalaria, y una forma de lograrlo es cuando las actividades y los recursos relacionados se gestionan como un proceso, esta forma de gestión logra identificar, estabilizar y controlar mejor las áreas de responsabilidad médica. Los procesos se consideran el núcleo operativo de gran parte de las organizaciones, donde la mejora continua de los mismos puede contribuir a disminuir las debilidades y afianzar las fortalezas de la organización, así mismo lograr el aumento de la productividad de la productividad y un mejor servicio sanitario en las instituciones de salud. La minería de procesos es una disciplina para analizar los procesos reales, partiendo de la extracción de conocimiento de los registros de eventos ampliamente disponibles en los actuales sistemas de información. El descubrimiento de procesos permite obtener modelos, en los que se pueden observar eventualidades visuales en los procesos modelados, y que en los registros de tienen los HIS no se está explotando para detectar dichas eventualidades. El objetivo de esta investigación es el desarrollo de una personalización de Visual Miner

Inductivo (complemento de ProM) que permite detectar sucesos que puedan estar pasando en los procesos hospitalarios, se desarrolló un componente para extraer información de un registro de eventos del HIS. El resultado obtenido con Visual Miner Inductivo fue que se presenta en forma de nodos las actividades o transiciones que forman parte de los procesos; obteniendo un modelo de proceso animado, en el que se pueden visualizar con mayor facilidad las actividades frecuentes o no, además de cuellos de botellas y desviaciones, que influyen en un mal funcionamiento de la institución y puede influir en una deficiente atención a pacientes. El complemento desarrollado forma parte del HIS añadiendo valor agregado tanto para analistas de procesos como para los desarrolladores del sistema.

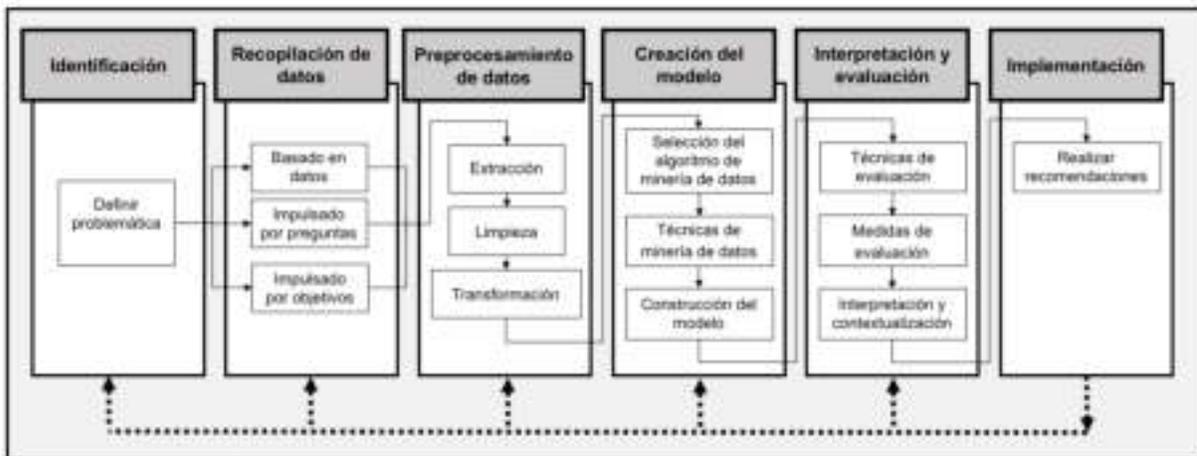
### 3. MODELO

A continuación, se realiza la propuesta de modelo con la finalidad de detectar áreas de oportunidad para establecer señalamientos y sugerencias de mejora en el proceso de atención de cáncer de mama. Como parte de la evaluación de técnicas y estrategias para la solución del problema, se identificaron modelos que se examinaron en el capítulo dos para la minería de datos como Jans, Jans, Rozinat, Bozkaya y van der Aalst, así como el proceso KDD.

El estudio es una combinación del enfoque cuantitativo y el cualitativo llamado enfoque mixto, utiliza diversas fuentes de información y tipos de datos como: pruebas de laboratorio, historial clínico y entrevistas que pueden incluir preguntas cerradas y/o abiertas, los cuales son integrados, analizados e interpretados en conjunto generando meta inferencias.

La base del modelo propuesto es el proceso KDD (1989), el cual incluye no sólo la obtención de los modelos o patrones, sino también la evaluación e interpretación. Incluye las tareas de selección, limpieza, transformación de los datos, extracción de patrones y modelos adecuados, evaluarlos e interpretarlos y convertirlos en conocimiento; y hacer el conocimiento disponible para su uso. A este proceso se le incorporan algunas técnicas de minería de datos que provienen del van der Aalst 1999, cuya finalidad es guiar el desarrollo para la detección de variabilidad en el proceso de atención hospitalaria como el del cáncer de mama, desde su detección hasta el término de su tratamiento, evaluando cada paso en la atención al paciente y detectando cuellos de botella o causas de retrasos en los procedimientos.

Las revisiones de literatura revisadas sugieren una serie de pasos partiendo de registros de eventos, de los que se desprende el modelo de la figura 3.1.



*Figura 3.1. Modelo adaptación KDD (1996) y van der Aalst (1999).*

En la figura 3.1 se presenta el modelo propuesto, el cual es una adecuación de los modelos de Van der Aalst y KDD, el cual se compone de 6 fases, donde cada una contiene tareas independientes, de igual forma se observa la secuencia de cada una de ellas; dependiendo de los resultados que se obtengan en cada fase es posible que se tenga que retroceder:

- A. Identificación
- B. Recolección de datos
- C. Preprocesamiento de datos
- D. Creación del modelo
- E. Interpretación y evaluación
- F. Implementación

A continuación, se describen cada una de las fases y tareas que la conforman.

### **3.1. Identificación**

Esta fase consiste en obtener un análisis exploratorio y descriptivo de la situación actual sobre el proceso de atención hospitalaria del cáncer como puede ser el de mama en la población derechohabiente de ISSSTESON.

Para ello se recomienda el uso de la estadística descriptiva que ayude en el análisis e interpretación de los datos como histogramas y tablas. El diagnóstico inicial consiste en conocer:

Pacientes:

- A. Análisis de casos únicos
- B. Diagnósticos
- C. Grupos de edad
- D. Casos especiales
- E. Colonias mayormente afectadas

Tiempos:

- A. Tiempos de los períodos entre citas
- B. ¿Quiénes son las que tienen citas espaciadas?, ¿Qué características tienen?

Obtener las características anteriores permite conocer cómo actualmente se está llevando a cabo el proceso de atención hospitalaria en la institución. Las consideraciones tomadas para este análisis se muestran en el anexo 7.1.

## 3.2. Recopilación de datos

En esta fase el objetivo es reconocer y reunir los datos con los que se va a trabajar.

Existen varias formas de obtener los datos:

- A. Basado en datos: en la institución existen sistemas de información la cual contienen datos de un evento, que corresponde al seguimiento de un paciente; y se espera que surjan ideas valiosas al analizar los datos de eventos.
- B. Impulsado por preguntas: se sugiere utilizar técnicas como encuestas y entrevistas. Se diseñaron tres tipos, las cuáles serán aplicadas al personal administrativo, médico y paciente; los diseños se pueden observar en el anexo 7.1.
- C. Impulsado por objetivos: el objetivo es mejorar los procesos en relación a los indicadores claves de rendimiento (por sus siglas en inglés KPIs), se identifican tres dimensiones: tiempo, costo y calidad.

Una vez que se determinan las fuentes de información que pueden ser útiles y dónde conseguirlas, los datos obtenidos deben ser integrados, se sugiere que esta tarea sea realizada mediante un almacén de datos, archivos de texto, hojas de cálculo, etc.

### **3.3. Preprocesamiento de datos**

Los datos al venir de distintas fuentes, pueden contener valores erróneos o faltantes. La etapa de preparación de datos está compuesta por tres tareas:

1. Extracción: en este proceso se deben extraer del sistema los datos que se serán utilizados en el análisis.
2. Limpieza: los datos extraídos deben ser preparados para un fácil manejo, deben tener una estructura adecuada para su posterior transformación, y existen varias estrategias para el manejo de datos faltantes, en blanco, datos inconsistentes o fuera de rango.
3. Transformación: en esta tarea se generan nuevas variables a partir de las ya existentes, consolidando los datos de una forma necesaria para la siguiente fase.

Estas tareas son importantes porque se asegura que los datos a minar sean de calidad, es decir relevantes y necesarios. Al terminar esta fase se obtiene una vista minable.

### **3.4. Creación del modelo**

El problema actual de la mejora en el proceso de atención hospitalaria como sería el cáncer de mama puede clasificarse como un problema de minería de procesos. Teniendo en cuenta las necesidades el modelo puede ser descubierto usando las técnicas de descubrimiento de procesos, utilizando el algoritmo Inductive Visual Miner con la herramienta ProM.

En el proceso de KDD será necesario explorar varios modelos hasta encontrar el mejor obtener el modelo de flujo de actividades que resuelva el problema, y es posible que se tenga que retroceder a las fases anteriores y hacer cambios en los datos, o incluso, modificar la definición del problema.

El modelo de flujo de actividades puede ser extendido a un proceso integrado teniendo en cuenta las diferentes perspectivas que a continuación se describen:

- A. agregar la perspectiva organizacional. Es importante identificar los departamentos o entidades organizativas que conectan las actividades dentro del proceso de atención del cáncer, en la tabla 3.1, se muestran los departamentos involucrados:

Departamentos	Descripción
<b>Área</b>	[Consulta Externa General, Coordinación CE General]
<b>Centro de Atención</b>	[CIAS Sur, CIAS Centro, CIAS Norte]
<b>Consultorio</b>	[5001, etc.]
<b>Inter consultorio</b>	[Ginecología y obstetricia, Clínica de Mama, Cirugía oncológica, Quimioterapia]
<b>Especialidades</b>	[Medicina General, Medicina Familiar, Medicina Integrada, Obesidad]
<b>Tipo de Médico</b>	[General, Especialista]

**Tabla 3.1.** Departamentos involucrados en el proceso de atención.

- B. agregar la perspectiva del tiempo. Para conocer las distribuciones de probabilidad que describen adecuadamente los tiempos de atención hospitalaria, se pueden utilizar las marcas de tiempo y las frecuencias. Se tomarán en cuenta 8 tiempos de atención para la operación del paciente:

1. fecha de mamografía (con sospecha),
2. fecha de salud al centro oncológico,
3. fecha en que el paciente que se ve por primera vez en el oncológico,
4. fecha biopsia,
5. fecha de resultado biopsia
6. fecha de estudios preoperatorios,
7. fecha enviada al hospital para operación,
8. fecha en que el paciente es operado.

- C. agregar la perspectiva del caso. Los atributos incorporados en el registro son utilizados para la vista minable. Esto demuestra que los datos son relevantes y deben ser incluidos en el modelo. En la tabla 3.2 se muestra un ejemplo de un log o bitácora de eventos.

Event id	Descripción	Evento/Actividad	Usuario	dd/mm/ yyyy hh:mm
1		Inicio	Usuario 1	02/01/2013 12:00:00 a.m.
2	Registrarse	Procesando	Usuario 1	02/01/2013 12:00:00 a.m.
3	Registrarse	Terminado	Usuario 1	02/01/2013 12:00:00 a.m.
4	Enviar cuestionario	Procesando	Usuario 1	02/01/2013 12:00:00 a.m.
5	Evaluar	Procesando	Usuario 1	02/01/2013 12:00:00 a.m.
6	Enviar cuestionario	Terminando	Usuario 1	02/01/2013 12:00:00 a.m.
7	Recibir cuestionario	Procesando	Usuario 1	02/01/2013 12:00:00 a.m.
8	Recibir cuestionario	Terminado	Usuario 1	02/01/2013 12:00:00 a.m.
9	Evaluar	Terminado	Usuario 1	02/01/2013 12:00:00 a.m.
10	Archivar	Procesando	Usuario 1	02/01/2013 12:00:00 a.m.
11	Archivar	Terminado	Usuario 1	02/01/2013 12:00:00 a.m.
12		Final	Usuario 1	02/01/2013 12:00:00 a.m.

**Tabla 3.2.** Log de Eventos: cada línea representa un evento.

- D. agregar otras perspectivas. Si se tienen otras variables pueden ser añadidas en el modelo para complementar la vista minable, la tabla 3.3 muestra una lista de variables que pueden considerarse para la atención hospitalaria.

#	Nombre de la variable	Descripción	Tipo
1	Tamaño del tumor	Tamaño del tumor	Discreto
2	Año del diagnóstico	Año del diagnóstico	Discreto
3	Etapa de la enfermedad	Etapa de la enfermedad	Discreto
4	Confirmación de diagnóstico	Método de confirmación del cáncer	Discreto
5	Estado de vida del paciente	Estado de vida del paciente	Binario

**Tabla 3.3.** Variables seleccionadas para la vista de datos minable.

### 3.5. Interpretación y Evaluación

El modelo descubierto debe ser verificado y evaluado antes de reemplazar el modelo antiguo. Como parte de la etapa es necesario realizar las siguientes tareas:

#### A. Técnicas de evaluación

Para conocer si el modelo descubierto se puede considerar como “buen modelo”, se puede verificar a través del análisis de verificación de conformidad, es la forma de visualizar las diferencias entre el proceso propuesto presentado en el modelo y el proceso real expresado en los registros de eventos, señalando posibles problemas para abordar. Además, la verificación de conformidad facilita el análisis de los procesos para verificar su cumplimiento con las mejores prácticas, como se especifican en guías clínicas.

#### B. Medidas de evaluación de modelos

En la figura 3.2 se muestran las cuatro medidas de evaluación para analizar los resultados de las técnicas de descubrimiento de procesos. El objetivo es medir la fidelidad con que los datos son descritos por el modelo:

1. Replay fitness: cuantifica hasta qué punto el modelo descubierto puede ser replicado sin errores.
2. Sencillez: la complejidad de un modelo de proceso es capturada por la dimensión de simplicidad y, existen varias técnicas que lo hacen simple por ejemplo el algoritmo  $\alpha$ .
3. Precisión: el modelo descubierto no debe permitir un comportamiento no relacionado a lo que se vio en el registro de eventos.

4. Generalización: esta medida evalúa hasta qué punto el modelo resultante podrá reproducir el comportamiento futuro del proceso, a lo que puede verse como una medida de confianza en la precisión.



**Figura 3.2.** Medidas de evaluación

Las cuatro mediciones de calidad son necesarias para realizar el descubrimiento de procesos y ninguna debe quedar fuera. Sin embargo, se considera la dimensión de Fitness la más importante de ellas. Se aconseja asignar pesos a cada medición, por ejemplo: fitness: peso 10; precisión: peso 5; generalización: peso 1; simplicidad: peso 1.

- C. Interpretación y contextualización: también se debe evaluar el contexto donde el modelo se va a utilizar, para ello existen varios indicadores, típicamente se encuentran tres dimensiones de desempeño: tiempo, costo y calidad, cada dimensión puede definir diferentes indicadores clave de rendimiento, donde cada una tiene acciones de mejora para el cumplimiento de sus objetivos, por ejemplo, para reducir el tiempo de flujo al proporcionar información que conduzca al rediseño o ajuste del proceso hospitalario, como se aprecia en la tabla 3.4.

#	Indicadores	Descripción	Objetivo	Acción
1	El tiempo de flujo	Tiempo desde el inicio hasta el final del proceso	Mejorar los KPIs	<ul style="list-style-type: none"> <li>• Rediseño (mejora del proceso),</li> </ul>
2	El tiempo de servicio	Tiempo por actividad	relacionados con el tiempo	<ul style="list-style-type: none"> <li>• Ajuste (mejorar control),</li> </ul>
3	El tiempo de espera	Tiempo de espera para un paciente		<ul style="list-style-type: none"> <li>• Intervenir (manejar el problema de manera ad-hoc),</li> </ul>
5	Reducción de costes	Costos en tratamientos	Mejorar los KPIs	<ul style="list-style-type: none"> <li>• Apoyo (detectar, predecir y recomendar)</li> </ul>

			relacionados con el costo
6	Mejora de la calidad	Mejora en el servicio de atención hospitalaria	Mejorar los KPIs relacionados con la calidad

**Tabla 3.4.** Indicadores clave de rendimiento (KPIs)

El criterio de evaluación del desempeño de la propuesta será calculado sobre la dimensión tiempo, referente a la atención en el proceso hospitalario como lo es el cáncer de mama, con el fin dar sugerencias de mejora al proceso.

### 3.6. Implementación

Una vez construido y validado el modelo lo siguiente es generar informes que incluyan sugerencias de mejora como apoyo al proceso de atención de pacientes de cáncer de mama. Es necesario realizar la difusión de los hallazgos encontrados al personal involucrado de la institución, una de las formas puede ser mediante reuniones y publicidad en redes sociales, para que a su vez se pueda generar material de comunicación personalizada a pacientes, y sean informados de las medidas a tomar para la detección precoz de la enfermedad.

Es importante medir cómo evoluciona el modelo. Es posible que los patrones cambien a lo largo de tiempo, por lo que el modelo puede ser re-evaluado, re-entrenado y/o reconstruido.

## 4. IMPLEMENTACIÓN

En este capítulo se presenta el desarrollo y la implementación del modelo propuesto en el capítulo previo, la cual fue aplicada en la institución de salud ISSSTESON.

A continuación, se detallan las actividades que se realizaron en cada una de las fases que componen el modelo.

### 4.1. Identificación

Como parte de la primera fase es obtener un análisis exploratorio y descriptivo de la situación de la institución.

Con base en la muestra proporcionada por la institución extraída de los HIS o sistemas de registro de consultas médicas generales se obtuvo el siguiente análisis preliminar siempre considerado la anonimidad de las personas:

Pacientes:

- A. Análisis de casos únicos
- B. Diagnósticos
- C. Grupos de edad
- D. Casos especiales
- E. Colonias mayormente afectadas

Tiempos:

- A. Tiempos de los períodos entre citas
- B. ¿Quiénes son las que tienen citas espaciadas?, ¿Qué características tienen?

#### 4.1.1. Criterios

Para realizar el análisis inicial fue necesario un procesamiento de los datos, por lo que se tomaron en cuenta los siguientes criterios:

1. Total de diagnósticos: Se encontraron 20 diagnósticos diferentes, de padecimientos de la mama.

2. Diagnósticos seleccionados: los cálculos ejecutados fueron realizados con los 6 diagnósticos más comunes (tabla 4.1):

No.	Diagnósticos más comunes
1	CA mama
2	Carcinoma in situ de la mama
3	Fibroadenosis de mama
4	Masa no especificada en la mama
5	Mastalgia-mastodinia-mastitis-dolor mama
6	Tumor maligno de la mama

**Tabla 4.1.** Diagnósticos más comunes.

Los diagnósticos “CA mama”, “Carcinoma in situ de la mama” y “Tumor maligno de la mama”, si bien se consideran como cáncer de mama el diagnóstico dependerá de qué células de la mama se vuelven cancerosas, también depende del criterio del médico.

3. Casos únicos por paciente: el cálculo fue realizado tomando en cuenta al paciente como registro único, excluyendo duplicados donde el paciente registra más de un diagnóstico.
- Considerando 20 diagnósticos, la muestra por casos únicos es de 1,650 (100%) y considerando 6 diagnósticos la muestra por casos únicos es de 1,217 (73.76%), como se muestra en la tabla 4.2.

Consideración	Casos únicos	% de Representación
20 diagnósticos	1,650	100.00%
6 diagnósticos	1,217	73.76%

**Tabla 4 2.** Casos únicos por paciente.

4. Casos por grupos de edad: el cálculo fue realizado considerando al paciente como registro único en cada grupo de edad, de manera que si el paciente pasó por varios grupos fue contabilizado en cada uno.

- La muestra que representa los 20 diagnósticos divididos por grupos de edad es de 1,701 (100%) casos, mientras que la muestra que representa los 6

diagnósticos divididos por grupos de edad es de 1,245 (73.19%) casos, como se aprecia en la tabla 4.3.

<b>Consideración</b>	<b>Casos</b>	<b>% de Representación</b>
20 diagnósticos	1,701	100.00%
6 diagnósticos	1,245	73.25%

**Tabla 4.3. Casos por grupos de edad.**

5. Casos por diagnóstico: se realizó considerando al paciente como registro único en cada diagnóstico, por lo que si el paciente tuvo uno o más padecimientos se contabilizó en cada uno de ellos.

- La muestra de casos por diagnóstico es de 1,835 (100%) casos considerando los 20 diagnósticos, y la muestra de casos por diagnóstico es de 1,310 (71.39%) casos considerando los 6 diagnósticos, como se muestra en la tabla 4.4.

<b>Consideración</b>	<b>Casos</b>	<b>% de Representación</b>
20 diagnósticos	1,835	100.00%
6 diagnósticos	1,310	71.39%

**Tabla 4.4. Casos por diagnóstico.**

6. Colonias mayormente afectadas vía Georeferenciación: el cálculo fue considerando los diagnósticos de “CA de mama”, “Tumor maligno de la mama” y el diagnóstico que puede evolucionar a cáncer de mama “Carcinoma in situ de la mama”, se tomó al paciente como registro único en cada diagnóstico y por grupo de edad, por lo que si el paciente tuvo uno o más padecimientos se contabilizó en cada uno de ellos y, si el paciente pasó por varios grupos fue contabilizado en cada uno (tabla 4.5).

- La muestra de casos por diagnóstico es de 1,310 (100%) casos considerando los 20 diagnósticos; y la muestra de casos por diagnóstico y grupo de edad es de 169 (12.90%) casos considerando los 3 diagnósticos. Se encontraron registros con diagnósticos duplicados los cuales fueron eliminados y algunos

registros con colonias no especificadas también fueron excluidos haciendo un total de 9 registros eliminados (tabla 4.5).

• Consideración	Casos	% de Representación
20 diagnósticos	1,835	100.00%
3 diagnósticos	169	12.90%

**Tabla 4.5.** Georeferenciación casos por diagnóstico y grupo de edad.

La identidad de los pacientes no es revelada en ningún momento ni existe un enlace a sus datos personales. La información georeferenciada no se presenta con alta precisión para evitar la identificación de derechohabientes. Hay que tener en cuenta que el instituto cuenta con una rotación constante de afiliados, aunque hay una base que permanece fija.

## 4.1.2. Pacientes

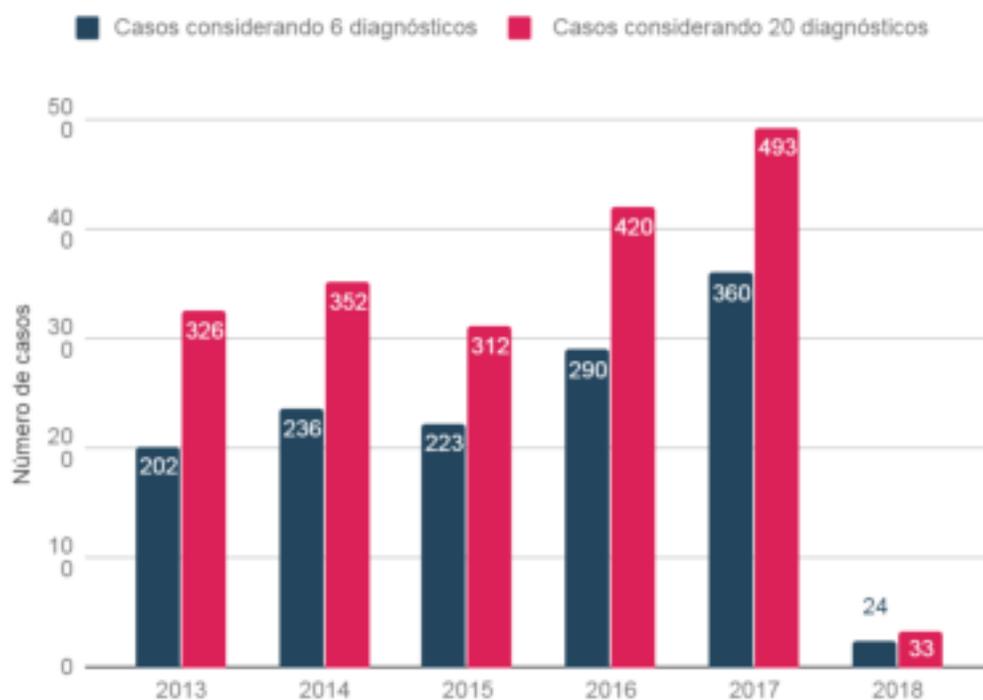
### A. Análisis de casos únicos

Tomando en cuenta todos diagnósticos registrados se obtuvo un diagnóstico inicial de la institución, mostrando un aumento en los casos únicos de padecimientos de la mama desde el 2013 al 2017. En la muestra proporcionada por la institución se tienen registrados 1,650 (100%) casos desde 2013 hasta inicios del 2018; el historial de diagnósticos encontrado en la institución es variado, contando 20 tipos de padecimientos de la mama. El análisis descriptivo tomará en cuenta los 6 padecimientos con mayor incidencia, contabilizando 1,217 (73.76%) casos. El cálculo fue realizado con un registro único por paciente, excluyendo los duplicados donde el paciente tiene o tuvo dos diagnósticos.

Se seleccionaron casos atípicos y típicos de los diagnósticos de la enfermedad que se tienen en la institución, entre los casos con la menor incidencia se encuentran 7 casos del sexo femenino igual o menores a un año y, un caso del sexo masculino entre los

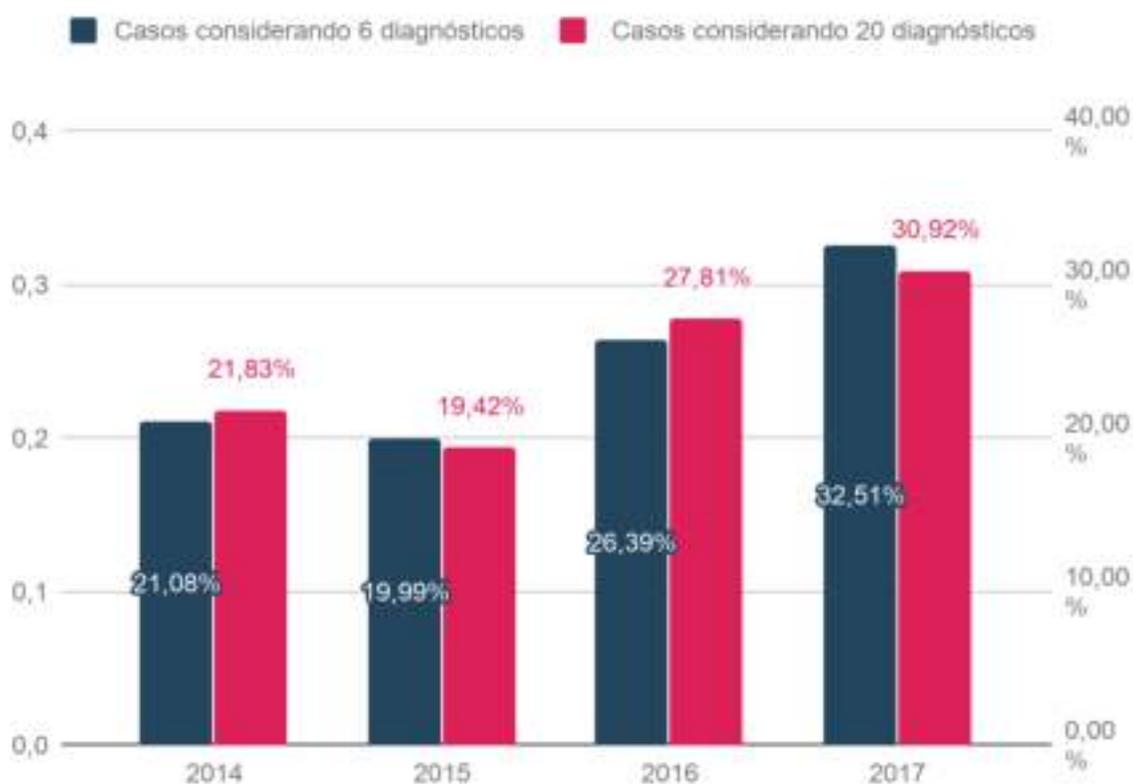
5-9 años; mientras que los casos con la mayor incidencia se da en las edades de 40-44 años.

El análisis de la muestra de datos del sistema de consultas y seguimiento a pacientes de ISSSTESON arrojó que en 2016 hubo un repunte en el número de casos y sigue en aumento en 2017 (los datos de análisis contemplan sólo enero de 2018), como se muestra en la figura 4.1.



*Figura 4.1. Comportamiento anual al alza de casos de cáncer de mama.*

En la figura 4.2 se aprecian los porcentajes de casos de los padecimientos de la mama de acuerdo a la muestra, siendo el 2017 con la mayor incidencia con un 32.51% considerando los 6 diagnósticos y un 30.92% para los 20 diagnósticos. Para el cálculo del índice de incidencia se tomó en cuenta la población de derechohabientes de la institución en cada año.

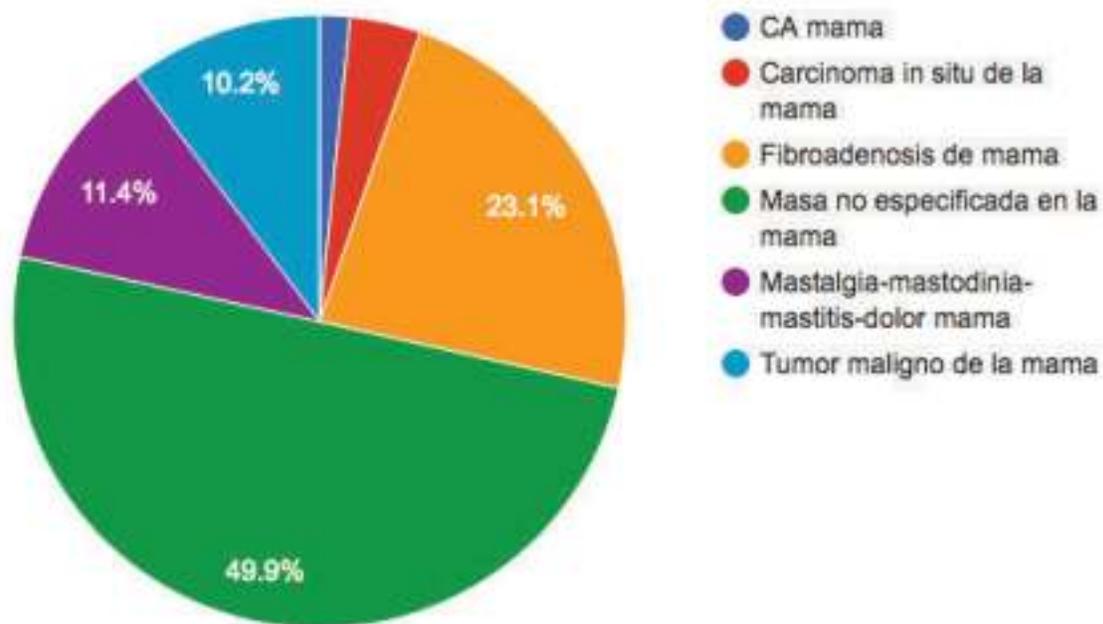


**Figura 4.2.** Porcentaje de incidencia con respecto a la población de la institución por cada año, correspondiente a los casos de padecimientos de la mama de acuerdo a la muestra proporcionada por la institución.

## B. Diagnósticos

El análisis de casos por diagnóstico arrojó que, de los 1,310 (cálculo realizado donde el paciente puede estar incluido en uno o varios diagnósticos) casos encontrados en el período a evaluar, la incidencia mayor se da en el diagnóstico de “Masa no especificada en la mama” con 654 (49.92%) casos, seguido de la “Fibroadenosis de

mama” con 303 (23.13%) casos, el tercer lugar lo ocupa el padecimiento “Mastalgia-mastodinia-mastitis-dolor mama” con 149 (11.37%) casos, “Tumor maligno de la mama” se encuentra en cuarto lugar con 134 casos (10.23%), mientras que el “CA Mama” con 21 (1.6%) casos es el diagnóstico con menos incidencia, como se muestra en la figura 4.3.



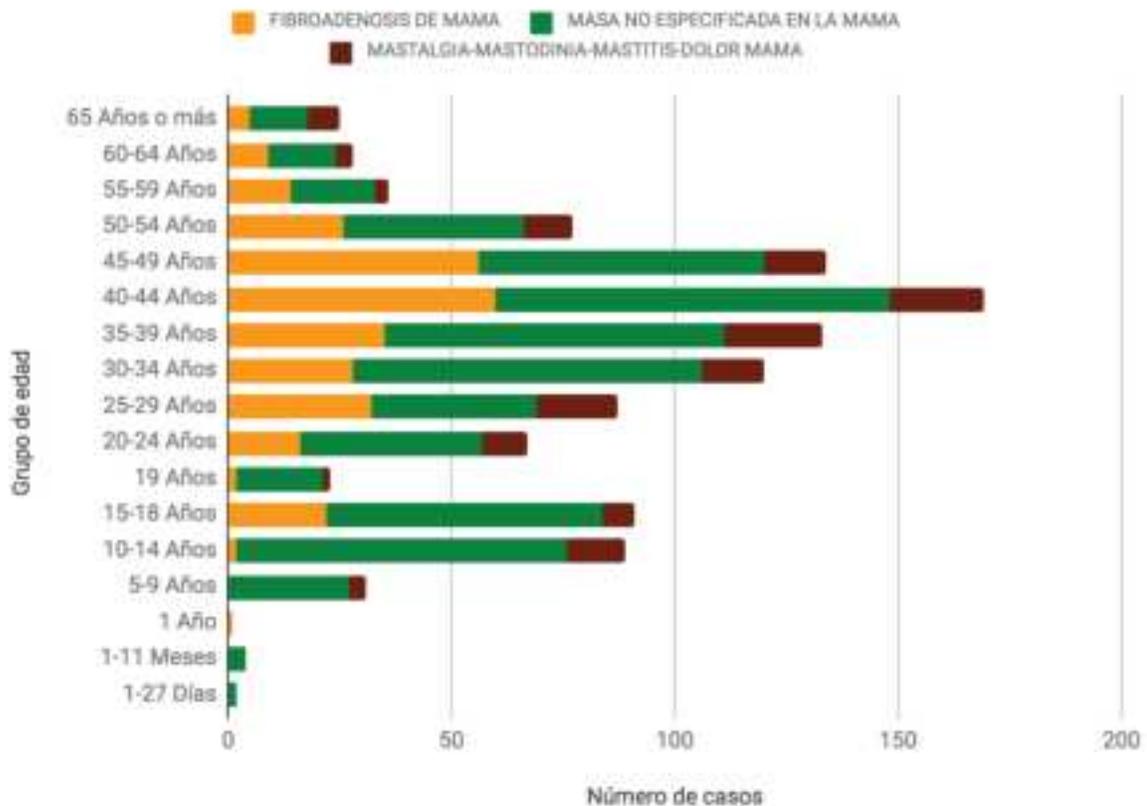
*Figura 4.3. Frecuencias por tipo de diagnóstico.*

### C. Grupos de edad

El análisis por edad muestra que la mayor incidencia corresponde a las edades de 40-44 años con 14.30% casos, seguido por los pacientes de 45-49 años con 12.45%. Las edades de 35-39 años y 30-34 años ocupan el tercer y cuarto lugar con 11.00% y 9.48% respectivamente, siendo el diagnóstico con mayores casos el “Masa no especificada en mama”.

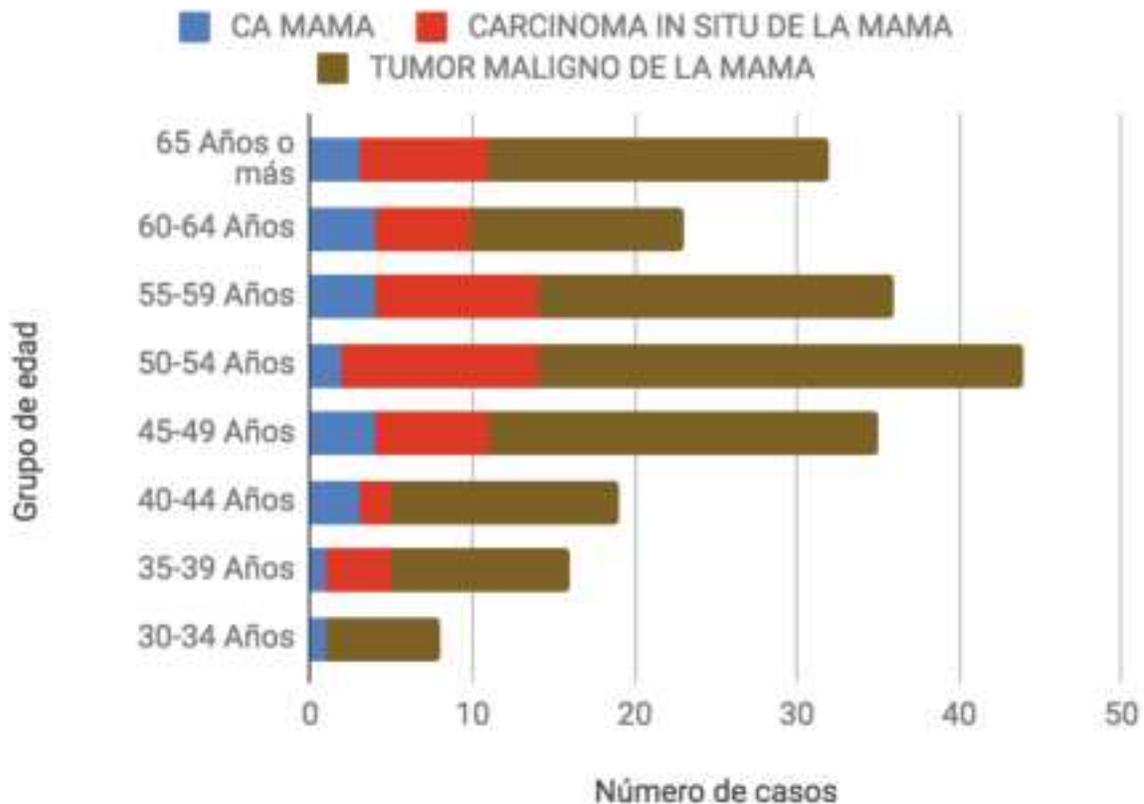
Se encontraron 7 casos atípicos en bebés entre los grupos de edad de 1 año, 1-27 días y 1-11 meses siendo los grupos con menor incidencia con 0.08%, 0.16%, 0.32% respectivamente; los principales diagnósticos encontrados para los casos atípicos son

“Masa no especificada en mama” y “Fibroadenosis de mama” como se muestra en la figura 4.4.



**Figura 4.4:** Grupos de edad en los padecimientos “Fibroadenosis de mama”, “Masa no especificada en la mama” y “Mastalgia-mastodinia-mastitis-dolor mama”.

De igual manera se observó que los casos de diagnóstico “CA mama” tiene presencia en cada grupo de datos a partir de los 30-34 años hasta los 65 años o más con al menos un caso y, para el diagnóstico de “Carcinoma in situ de la mama” siendo un diagnóstico que puede detonar en cáncer de mama se encontró que los registros para éste son a partir de los 35-39 años hasta los 65 años o más, donde la mayor incidencia se da entre los 50-54 años. Otro padecimiento es el “Tumor maligno de la mama” se ha presentado desde los 30 hasta los 65 años o más, la principal incidencia se da en los grupos de edad de los 50-54 años como se aprecia en la figura 4.5.

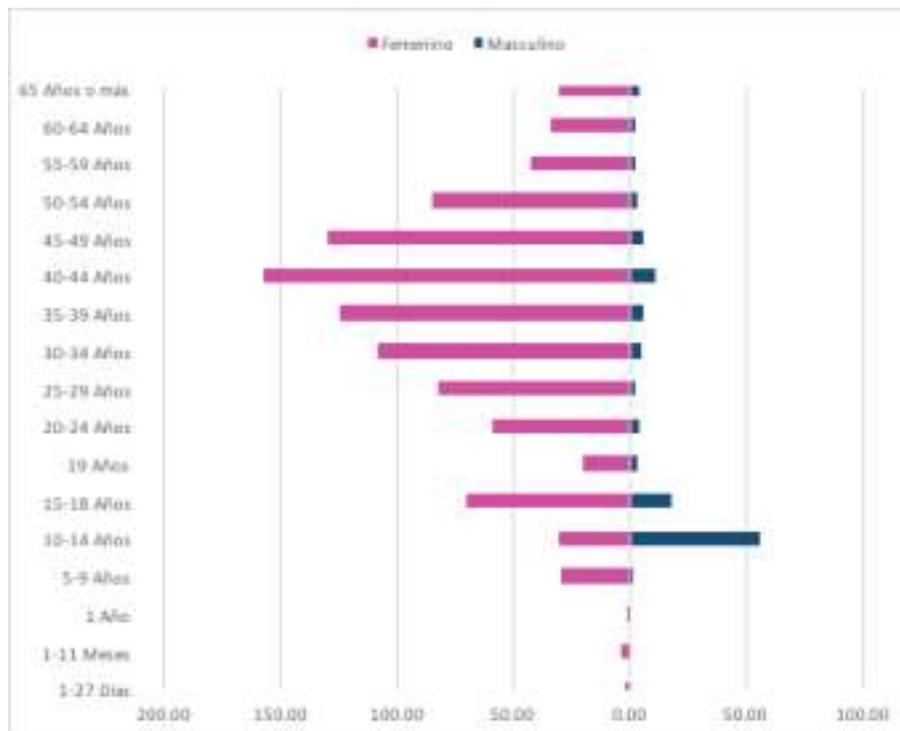


**Figura 4.5.** Casos por Grupos de edad para los casos de “CA mama”, “Carcinoma in situ de la mama” y “Tumor maligno de la mama”.

#### D. Casos especiales

Se analizó el comportamiento de la enfermedad por género (el cálculo excluye registros que no tienen género definido) para observar aquellas situaciones especiales que pueden estar presentes como se muestra en la figura 4.6; se observó que la mayor incidencia en estos diagnósticos de la mama se da en el sexo femenino en las edades de 40-44 años, seguido de 45-49, 35-39 y 30-34 (segundo, tercero y cuarto lugar respectivamente). Sin embargo, para el género masculino la mayor incidencia se da en edades tempranas, donde la edad predominante con este padecimiento es de 10-14 años y en segundo lugar de 15-18, después vienen las edades 40-44, 45-49, 35-39 y 30-34 consideradas como edades adultas (contrarias al sexo femenino que son las que tienen mayor incidencia). Sin embargo, se observan algunos compartimentos

atípicos en ambos géneros, en el caso del género femenino se detectaron casos en bebés menores de un año y, para el género masculino la edad se da entre 5-9 años.



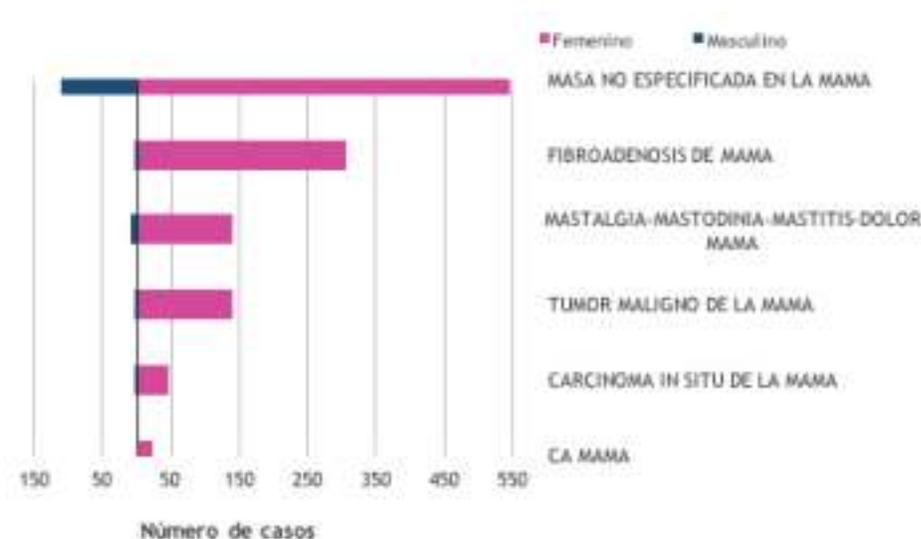
**Figura 4.6.** Incidencia de los padecimientos de la mama por grupo de edad y género.

Los casos atípicos encontrados en el análisis se muestran en la tabla 4.6, se destacan 2 diagnósticos: “Masa no especificada en la mama” con 6 casos en edades de 1-11 meses (4 casos) y 1-27 días (2 casos) todos del sexo femenino, “Fibroadenosis de mama” un caso de una bebé de un año.

Diagnóstico	Sexo	Grupo de edad	Casos
Masa no especificada en la mama	Femenino	1-11 Meses	4
Masa no especificada en la mama	Femenino	1-27 Días	2
Fibroadenosis de mama	Femenino	1 Año	1
<b>Total</b>			<b>7</b>

**Tabla 4.6.** Casos especiales en bebés de hasta un año de edad.

En el caso de los diagnósticos con más frecuencia tanto para el género femenino (547 casos) y masculino (112 casos) es “Masa no especificada en la mama”, sin embargo, en el caso de las mujeres el segundo diagnóstico es el de “Fibroadenosis de mama” con 307 casos y para hombres sólo un caso registrado. La segunda incidencia en el caso de los hombres es “Mastalgia-mastodinia-mastitis-dolor mama” con 10 casos; para ambos géneros el diagnóstico menos frecuente es “CA mama” con 22 casos en mujeres y en el caso de hombres no se tienen registrados casos (figura 4.7).



*Figura 4.7. Diagnósticos por género*

### E. Colonias mayormente afectadas

Se aplicó el algoritmo de agrupamiento K-Medias (o K-means) para clasificar los casos de acuerdo a su cercanía geográfica con el fin de encontrar clústers con comportamiento anormal, por ejemplo, una concentración atípica en un área reducida que pudiera indicar algún factor externo (exposición a químicos, etc). Para el cálculo de K se utilizó el Método del codo (Elbow Method).

Para el análisis de georreferenciación se obtuvo un mapa de calor del municipio de Hermosillo (figura 4.8) considerando los diagnósticos “CA mama”, “Carcinoma in situ de la mama” y “Tumor maligno de la mama” 169 casos de muestra proporcionada. No

se observan clústers que indiquen un comportamiento anormal, sin embargo, podemos identificar las áreas donde existe una mayor incidencia (colonia Altares y Colonia Villa Bonita al poniente y sur de Hermosillo). Para la georreferenciación se utilizó “Google Fusion Tables” (figura 4.8) es un servicio gratuito de Google y es compatible con archivos en formato “.csv”, siendo este formato el de origen que contiene los registros analizados.

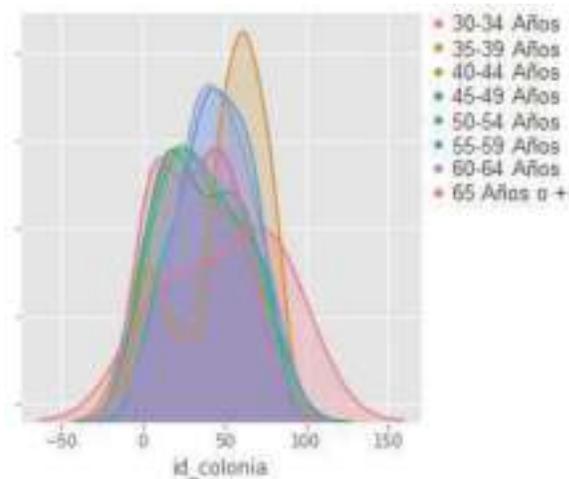


**Figura 4.8.** Mapa de la ciudad de Hermosillo con zonas calientes de los casos de “CA mama”, “Tumor maligno de la mama” y “Carcinoma in situ de la mama”.

Aunque los datos proporcionados son una fracción de los casos totales, esta metodología pudiera ser aplicada al total de casos de distintas instituciones de salud para lograr un diagnóstico más preciso sobre la situación de cáncer en distintas ciudades del país.

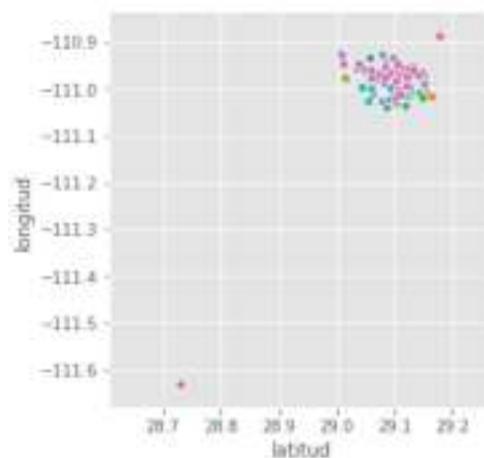
### Visualización de Datos

Antes de realizar el agrupamiento de los datos se graficaron para tener una idea de la dispersión de los mismos: la figura 4.9 no muestra una segmentación clara, sin embargo, en las distribuciones por grupo de edad y por colonia, se observan casos atípicos en los grupos de edad de 35-39 años y 30-34 años.



**Figura 4.9.** La distribución por grupos de edad y colonia, se puede apreciar que hay un pico en el grupo de edad de los 35-39 años, así como el de los 30-34 años.

La figura 4.10 muestra la distribución mediante coordenadas geográficas (latitud, longitud), no se observa una segmentación definida, sin embargo, se aprecian unos puntos dispersos: en la colonia Cuauhtemoc (28.7297221, -111.6313889) se tiene registro de 3 casos, 2 por “Tumor maligno de la mama” (en las edades de 50-54 años y 65 años o más), el tercer caso corresponde a “CA de mama” en las edades de 60-64 años. De nuevo, estos resultados son significativos para esta muestra, sin embargo, necesitamos más datos para determinar con mayor precisión la incidencia de casos de cáncer de mama en el área.



**Figura 4.10.** La distribución de coordenadas geográficas (latitud, longitud) no muestra una segmentación definida.

### **Generación de grupos por medio de algoritmos de agrupamiento**

El instituto requiere tener conocimiento de grupos de pacientes por área geográfica, diagnóstico y edad para dar seguimiento y evaluar la posibilidad de impartir pláticas de cuidados y concientización sobre el cáncer de mama en los distintos centros de atención de la ciudad.

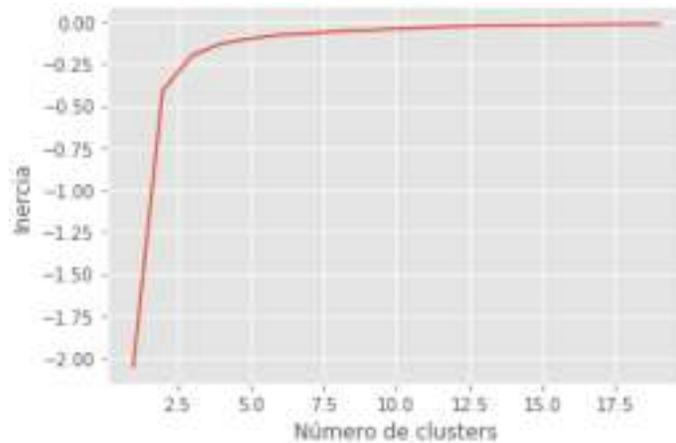
Para este propósito generamos un modelo de clasificación que agrupa a los pacientes de acuerdo a sus características. El resultado de esta clasificación servirá para tomar decisiones sobre campañas de medicina preventiva, pero, además, las autoridades podrán ver si hay algún comportamiento inusual en distintas áreas de la ciudad e investigar las posibles causas.

#### **Definir la entrada**

Se utilizó el algoritmo K-means utilizando lenguaje de programación Python para que agrupe los casos de pacientes por su cercanía geográfica y por el grupo de edad, algunos datos de la muestra se consideraban outliers o valores atípicos, los cuales fueron eliminados.

#### **Obtener el valor K**

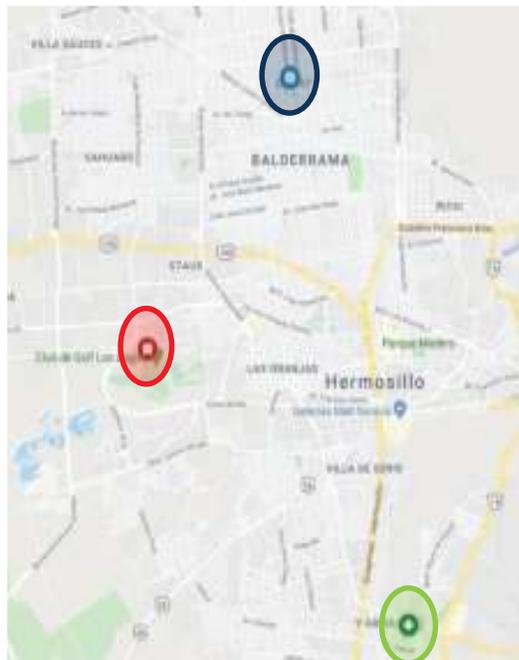
Se utilizó el Método del codo (Elbow Method) para seleccionar el número óptimo de clústers, conocido como valor de K. El código utilizado para conocer k y, la gráfica donde se aprecia el brusco cambio en la evolución de la inercia se pueden apreciar en la figura 4.11, el punto de “codo” indica que k=3 es la cantidad adecuada de clústeres.



**Figura 4.11.** Código en Python utilizado para realizar la curva del codo en la que aprecia un cambio brusco en la evolución de la inercia donde  $k=3$ .

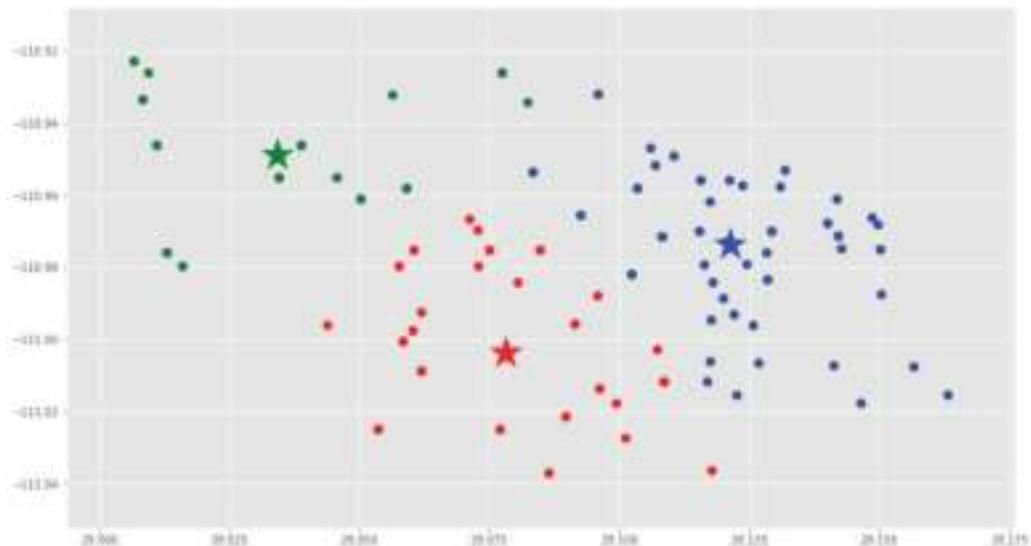
### Algoritmo de agrupamiento K-Means

Se ejecutó el algoritmo de K-Means para 3 clústers y se obtuvo las etiquetas y centroides arrojados en latitud y longitud. El clúster 1 se ubica por el poniente de la ciudad (29.07824814, -111.0038168), el clúster 2 se ubica al sur (29.03431695, -110.94894747) y el clúster 3 al norte de la ciudad (29.1214188, -110.97376487), como se muestra en la figura 4.12.



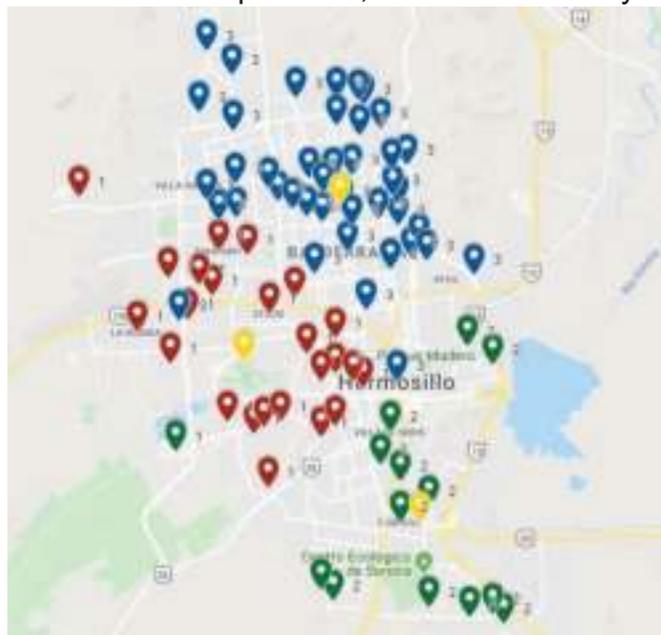
**Figura 4.12.** Centroides de la ciudad de Hermosillo

En la figura 4.13 se aprecia la figura generada que muestra los 3 clústers diferenciados por colores, el clúster 1 se representa de color rojo, el clúster 2 por el verde y el clúster 3 de color azul.



**Figura 4.13.** Identificación de 3 clúster diferenciados por colores: clúster 1=rojo; clúster 2=verde; clúster 3=azul.

En la figura 4.14 se presentan los casos georeferenciados a lo largo de la ciudad de Hermosillo, ubicando el clúster uno al poniente, el clúster 2 al sur y el clúster 3 al norte.



**Figura 4.14.** Georeferenciación de clúster de la muestra de casos de cáncer de mama en Hermosillo.

### Análisis de clúster

La tabla 4.7 muestra el análisis de cada clúster, se encontró que de los 165 (97.63%, cifra que quedó al eliminar los outliers) casos de la muestra, el clúster 1 tiene 52 casos (30.76%), el clúster 2 tiene 41 (24.26%) casos, mientras que el clúster 3 cuenta con 72 (42.60%) casos. Los Casos grupos de edades están representados del 1 al 8 (1="30-34 años"; 2="34-39 años"; 3="40-44 Años"; 4="45-49 Años"; 5="50-54 Años"; 6="55-59 Años"; 7="60-64 Años"; 8="65 Años o +"), mientras que los casos por diagnóstico se representan del 1 al 3 (1="CA de mama"; 2="Carcinoma in situ de la mama"; 3="Tumor maligno de la mama").

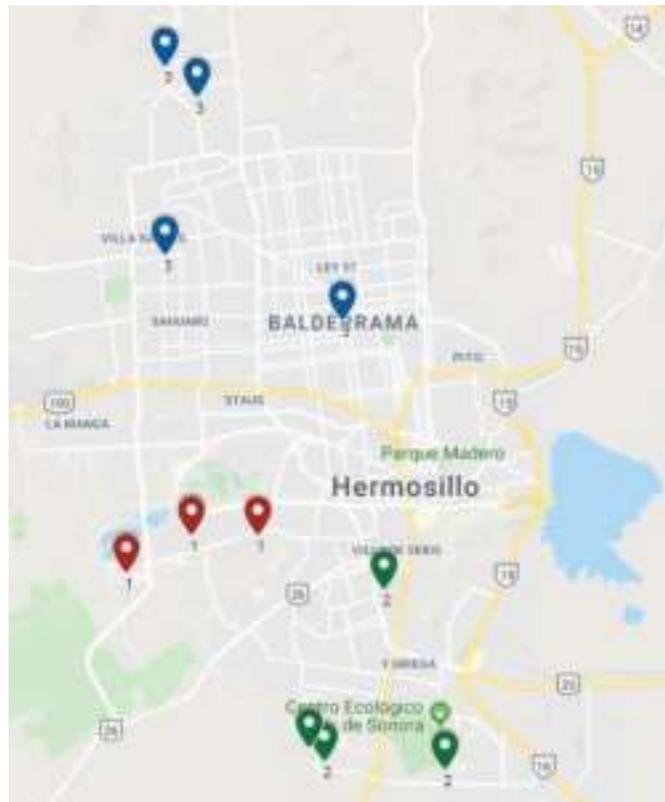
En el clúster 1 la mayor incidencia se da en los grupos de edad de 45-49 años y los casos atípicos se encuentran en las edades de 30-34 años. Para el clúster 2 la mayor incidencia se da en el grupo de edades 50-54 años y, los casos atípicos están en el grupo de 30-34 años al igual que el clúster 1. En el caso del clúster 3 la incidencia se da a partir de los 65 Años o más, al igual que en los demás clústers los casos atípicos está en el rango de edades de 30-34 años.

Clúster	Total de Casos	Casos por Grupo de edades	Casos por diagnósticos	Colonia	Casos únicos	Casos por grupos de edad
1	52	1=1,	1=7,	Colonia el Sahuaro	4	7
		2=5,	2=13,	Colonia Villa Bonita	5	7
		3=8,	3=32.	Fraccionamiento Urbi	2	4
		4=10,		Villa Campestre		
		5=7,				
		6=11,		*se consideraron las colonias con mayor incidencia		
		7=3,				
		8=7.				
2	41	1=3,	1=11,	Colonia Altares	6	7
		2=5,	2=7,	Colonia las lomas	4	6
		3=3,	3=23.	Colonia los Naranjos	2	4
		4=3,		Colonia quintas del Sol	1	4
		5=9,				
		6=8,		*se consideraron las colonias con mayor incidencia		
		7=6,				
		8=4.				

3	72	1=1, 2=3, 3=6, 4=13, 5=1, 6=8, 7=11, 8=17.	1=3, 2=21, 3=48.	Colonia Balderrama Colonia Olivares Colonia Bugambilias	5 5 3	6 5 4
				*se consideraron las colonias con mayor incidencia		

**Tabla 4.7. Detalle de clústers.**

La figura 4.15 se muestra un análisis de los casos atípicos entre los grupos de edad de 30-34 años y 34-39 años. De estos grupos de edades se encontraron 5 casos en el clúster 1 (marcas color rojo) 3 casos en Villa Bonita, Urbi Villa y Portal del Pitic con un caso cada uno. En el clúster 2 (marcas color verde) se encontraron 4 casos ubicados al sur de la ciudad, las colonias que figuran son los Altares, Emiliano Zapata, Quintas del Sol y Villas del Sur, todas con un caso. En el clúster 3 (marcas color azul) se encontraron en la muestra 4 casos en las colonias la Balderrama, las Dunas, Privada San German y Pueblitos.



**Figura 4.15. Casos atípicos entre los grupos de edad de 30-34 años y 34-39 años, por clúster.**

## DetECCIÓN TEMPRANA DE CÁNCER DE MAMA

La prevención temprana de cáncer puede aumentar la probabilidad de supervivencia del paciente, por lo que el instituto requiere conocer los tiempos de atención del proceso de cáncer de mama, esto para mejorar o reenfocar campañas de prevención, cambiar o eliminar actividades que se realizan durante el proceso.

### 4.1.3. Tiempo entre consultas

#### A. Tiempos de los períodos entre citas

Según los registros de medicina general, el análisis arroja un promedio de tiempo en días entre cada consulta de 121.01 (3.97 meses, el cálculo fue tomando 30 días por mes), esto incluyendo 194 (100%) derechohabientes que han asistido una o más veces al centro de atención. Mientras que, si se toma otra muestra con derechohabientes que han asistido más de 5 [20 casos únicos (10.00%)] ocasiones al médico y, que pudiera decirse que no se fueron a otro hospital, el promedio es de 99.89 días (3.28 meses), encontrando el menor tiempo promedio encontrado fue de 21.62 días (0.71 meses) entre cada cita, mientras que el tiempo promedio máximo alcanza los 290 días (9.53 meses) como se aprecia en la tabla 4.8.

Consideración	Casos únicos	% de Representación	Promedio entre citas (días)	Min. Promedio entre citas (días)	Max. Promedio entre citas (días)
2 citas o +	194	100.00 %	121.01	0.5	736
5 citas o +	20	10.00 %	99.89	21.625	290

*Tabla 4.8. Tiempo en días entre citas, cuyos pacientes tienen más de dos citas.*

#### B. ¿Quiénes son las que tienen citas espaciadas?, ¿Qué características tienen?

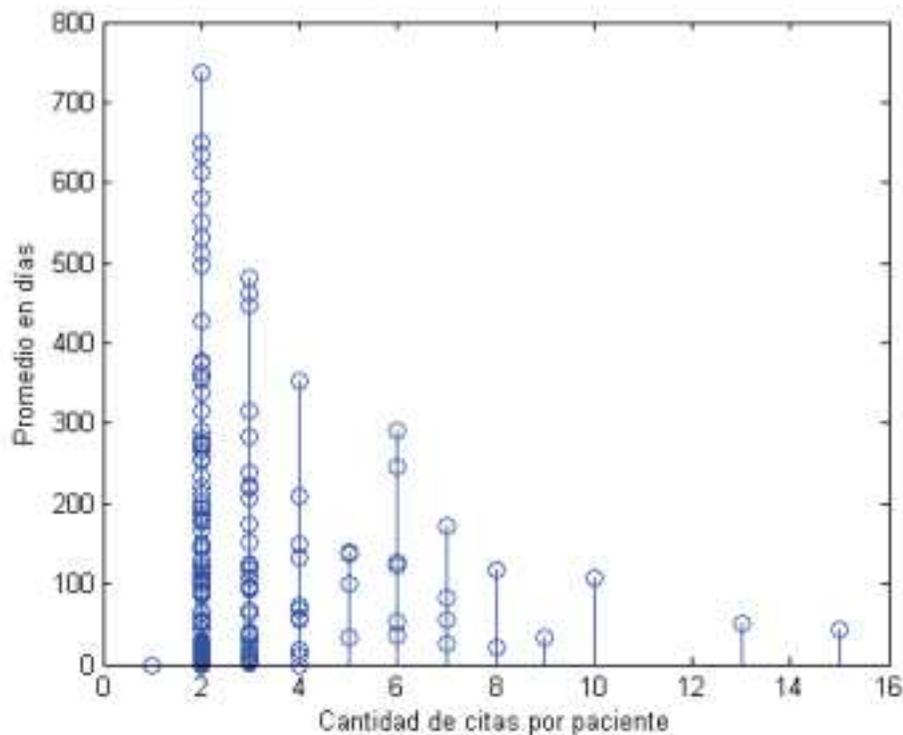
Tomando en cuenta lo citado en la literatura, donde se menciona que el iniciar un tratamiento de 3 a 6 meses después de los primeros síntomas, los pacientes tuvieron una supervivencia a cinco años significativamente más baja que las que esperaron menos de 3 meses, se obtiene como resultado 37 casos donde el promedio de citas oscila entre los 91.5 días (3 meses) y 178.5 días (5.8 meses), además se encontraron 47 derechohabientes cuyo promedio de citas va de los 187.5 días (6.16 meses) hasta

los 736 días (24.19 meses) como se observa en la tabla 4.9. Los cálculos fueron realizados a partir de los pacientes que tienen registrados 2 o más citas.

Consideración	Casos únicos	% de Representación	Promedio entre citas (días)	Min. Promedio entre citas (días)	Max. Promedio entre citas (días)
3 a 6 meses	37	19.00 %	126.71	91.5	178.5
6 meses o +	47	24.00 %	352.52	187.5	736

**Tabla 4.9.** Tiempo promedio entre citas mayores a 91.5 (3 meses).

De la muestra proporcional, en la figura 4.16 se puede observar el promedio en días por derechohabiente registrado a lo largo del periodo de estudio, de los 194 (100.00%) pacientes 174 (90.00%) tienen registradas entre 2 y 4, y el tiempo entre citas oscila entre 0 días y 736 días (24.19 meses).



**Figura 4.16.** Promedio en días entre citas por casos únicos en la institución.

## **4.2. Recopilación de datos**

En esta fase se presentan las herramientas, documentos y publicaciones utilizados para la recopilación de datos.

### **4.2.1. Aplicación de encuesta, revisión de documentación y publicaciones oficiales**

#### **Impulsado por preguntas**

Para evaluar los procesos de atención a pacientes con cáncer de mama en ISSSTESON, se aplicó una encuesta que extrae información de los pacientes atendidos en la institución. Se aplicó a 19 pacientes que han sido diagnosticados con cáncer de mama, las cuales estaban en tratamiento activo.

La importancia de esta encuesta es conocer paso a paso el proceso de atención recibido por los pacientes, permitiendo además conocer las percepciones del cliente acerca de los estilos de prácticas, de los aspectos administrativos y de las modalidades de tratamiento, ya que la satisfacción está asociada a mejores resultados de la salud y al cumplimiento del tratamiento.

El instrumento utilizado consta de cinco secciones:

- a) General (13 reactivos),
- b) estatus actual (6 reactivos),
- c) información sobre la enfermedad (5 reactivos),
- d) evaluación personal de enfermería (11 reactivos con calificación del 1 al 10); evaluación de personal de hospitalización y médico (28 reactivos con calificación del 1 al 10); evaluación servicio general (4 reactivos con calificación del 1 al 10 y un reactivo con respuesta abierta);
- e) vivir con la enfermedad (2 reactivos con respuesta abierta).

### **Criterios institucionales**

En una entrevista con el personal administrativo, se informó que en el instituto prevalece el criterio del médico al tiempo de sospecha de cáncer de mama. El seguimiento de la norma Oficial Mexicana NOM-041-SSA2-2011 para la prevención, diagnóstico, tratamiento, control y vigilancia epidemiológica del cáncer de mama se da, sin embargo, el médico indica el tipo de tratamiento o el paso a otro departamento según las características de los pacientes con cáncer de mama.

Las consideraciones de la institución dependen en gran medida del médico:

- Criterios de referencia de mujeres con sintomatología mamaria sospechosa de cáncer: todos pasan a clínica de mama
- Las experiencias de los médicos influyen en qué tipo de tratamiento o el paso a otro departamento según las características de los pacientes
- La cirugía, incluso si está prescrita por las guías clínicas, puede ser a veces crítica para determinados pacientes.
- Campañas preventivas

### **Criterios NOM-041-SSA2-2011**

La NOM tiene como objetivo establecer los lineamientos para la promoción de la salud, prevención, diagnóstico, tratamiento, control y vigilancia epidemiológica del cáncer de mama. Es de observancia obligatoria para todo el personal de salud, profesional y auxiliar de los sectores público, social y privado pertenecientes al Sistema Nacional de Salud que brinden atención médica en materia de cáncer de mama.

La tabla 4.10 muestra los criterios para la referencia. Se deben referir a una unidad especializada de mama a las mujeres con:

<b>Criterios para referencia</b>
Tumoración mamaria de características malignas a cualquier edad
Alteraciones de la piel como ulceración, retracción de la piel o pezón, engrosamiento de la piel
Nueva tumorización en mujeres con nodularidad preexistente

Nodularidad asimétrica que persiste después de la menstruación en mujeres menores de 35 años con antecedentes familiares de cáncer de mama o en mujeres de 35 o más años de edad

Descarga sanguinolenta, abundante o persistente por el pezón

**Tabla 4.10.** Criterios para referir a una unidad especializada de mama. NOM-041-SSA2-2011.

El cáncer de mama se debe codificar, de acuerdo con la Clasificación Internacional de Enfermedades de la siguiente manera como se muestra en la tabla 4.11:

Clasificación	Codificación
Tumor maligno de la mama	C50
Tumor maligno del pezón y areola mamaria	C50.0
Tumor maligno de la porción central de la mama	C50.1
Tumor maligno del cuadrante superior interno de la mama	C50.2
Tumor maligno del cuadrante inferior interno de la mama	C50.3
Tumor maligno del cuadrante superior externo de la mama	C50.4
Tumor maligno del cuadrante inferior externo de la mama	C50.5
Tumor maligno de la propagación axilar de la mama	C50.6
Lesión de sitios contiguos de la mama	C50.8
Tumor maligno de la mama, parte no especificada	C50.9

**Tabla 4.11.** Codificación de cáncer de mama según la Clasificación Internacional de Enfermedades. NOM-041-SSA2-2011.

Según la NOM los métodos terapéuticos que se emplean en la actualidad para tratar el cáncer mamario son (tabla 4.12):

Métodos terapéuticos
Cirugía
Radioterapia
Quimioterapia
Hormonoterapia
Biológicos

**Tabla 4.12.** Métodos terapéuticos. NOM-041-SSA2-2011.

Las variables mínimas para el registro de casos de cáncer de mama se detallan en la tabla 4.13.

Variables
Nombre
Sexo
Fecha de nacimiento
Dirección de residencia habitual
Fecha de diagnóstico

Sitio anatómico del tumor primario
Tipo histológico
Medida del tumor primario
Etapa clínica
Clasificación TNM

**Tabla 4.13.** Variables mínimas para el registro de casos de cáncer de mama.

El equipo mínimo de integración de las unidades de detección y diagnóstico lo constituyen: recepcionista, técnico, físico médico, radiólogo (tabla 4.14).

<b>Equipo</b>
Recepcionista
Técnico
Físico médico
Radiólogo

**Tabla 4.14.** Equipo de integración mínimo

La NOM menciona algunos indicadores para evaluar el desempeño y proceso de la Norma (tabla 4.15):

Indicador	Fórmula	Estándar
Oportunidad para el diagnóstico	$((\text{Total de mujeres referidas con diagnóstico definitivo en menos de tres citas}) / \text{Total de mujeres referidas para evaluación diagnóstica}) * 100$	> 95%
Oportunidad para la notificación del resultado	$((\text{Total de mujeres con resultado notificado en menos de 21 días}) / \text{Total de mastografías de tamizaje realizadas}) * 100$	> 95%
Oportunidad de la referencia para tratamiento	$((\text{Total de mujeres con diagnóstico confirmatorio de cáncer de mama referidas a un centro oncológico en menos de 10 días hábiles}) / \text{Total de mujeres con diagnóstico confirmatorio de cáncer de mama}) * 100$	> 90%
Oportunidad de tratamiento	$((\text{Total de mujeres con diagnóstico de cáncer que iniciaron tratamiento en menos de 15 días}) / \text{Total de mujeres con diagnóstico de cáncer que iniciaron tratamiento}) * 100$	> 90%
Sobrevida	(No. De mujeres vivas a 1,3 y 5 años del diagnóstico, según etapa) / No. De mujeres con cáncer en seguimiento a 1, 3 y 5 años del diagnóstico según etapa	0 > 99% I > 93% II > 72 % III > 41% IV > 18%

**Tabla 4.15.** Estándares de referencia. NOM-041-SSA2-2011.

### **4.3. Preparación de datos**

La extracción de datos se realizó de las entrevistas hechas al personal administrativo, médico como a pacientes. Las entrevistas se realizaron de manera personal, los pacientes que participaron son pacientes en tratamiento activo.

Las entrevistas realizadas al personal administrativo y médico se utilizaron para crear el modelo de guías clínicas, fueron entrevistas orales, los datos generados fueron limpiados y transformados en nuevas variables que indican las tareas o actividades realizadas por el instituto durante el proceso de atención de cáncer de mama.

Las entrevistas a pacientes fueron realizadas oralmente y algunas están almacenadas en formato “.mp3” otras en formato “.csv”, se realizó la limpieza de los datos generando una base de datos única en formato “.csv”, formando nuevas variables para la creación del modelo de proceso; de las 19 entrevistas hechas, se descartó una entrevista por no contener información necesaria para el diseño del modelo de procesos, por lo que se consideraron 18.

Se consideró las estructuras de datos que aceptan cada software a utilizar, teniendo especial cuidado con los acentos, ya que pueden interferir en las etapas posteriores (creación del modelo e interpretación).

### **4.4. Creación del modelo**

Una vez que ya se tienen los criterios institucionales y de la NOM-041-SSA2-2011 junto con el procesamiento de datos realizados de las entrevistas, se procede a alimentar el algoritmo Declare Designer para la creación del modelo basado en guías clínicas.

#### 4.4.1. Diseño del modelo basado en guías clínicas (modelo declarativo)

Los modelos declarativos se utilizan para analizar procesos que funcionan en entornos inestables e impredecibles como los procesos de tratamiento médico, ya que dependen de las experiencias de los médicos y de las características clínicas que tiene cada paciente. Los modelos declarativos proporcionan a los participantes del proceso un conjunto de restricciones a seguir durante la ejecución del mismo. Es un modelo de flujo de trabajo que permite observar la representación formal del proceso de atención de cáncer de mama mediante el diseño de las actividades y las restricciones que caracterizan el comportamiento del proceso (Rovani et. al, 2015).

- Elección del software

Se utilizó Plug-in Declare Designer para la creación del modelo basado en guías clínicas. Es un lenguaje basado en lógica temporal lineal (por sus siglas en inglés LTL) con semántica basada en trases y cada uno con su representación gráfica.

Las entidades que forman el modelo se pueden clasificar en cuatro grupos:

- existencia: implican solo una actividad y definen la cardinalidad permitida o la posición de una actividad en un caso.
- relación: definen una dependencia entre dos actividades.
- relación negativa: definen una dependencia negativa entre dos actividades.
- elección: definen relaciones alternativas entre actividades.

La figura 4.17 muestra el modelo declarativo que describe el proceso de atención para pacientes de cáncer de mama. El modelo declarativo para este proceso involucra 29 actividades (tabla 4.16), con 35 restricciones.

No.	Actividad	Nomenclatura
1	Poner cita con médico general	PCMG
2	Acude con médico general	AMG
3	Estudios de laboratorios ordinarios	ELO
4	Estudios de laboratorios urgentes	ELU
5	Estudios de rayos x ordinarios	ERO
6	Estudios de rayos x urgentes	ERU
7	Poner cita clínica de mama	PCM
8	Acude a clínica de mama	ACM
9	Estudios preoperatorios	EP
10	Operación	OP
11	Poner cita revisión periódica	PRP
12	Asistir revisión periódica	ARP
13	Alta del paciente	AP
14	Poner cita con ginecólogo	PG
15	Acude con el ginecólogo	AG
16	Realizar Ultrasonido	RU
17	Poner cita con oncólogo	PO
18	Acude al oncólogo	AO
19	Estudios de laboratorio	EL
20	Emite receta	EL
21	Poner cita en quimioterapia	PQ
22	Registro	R
23	Entrega receta en compras	ERC
24	Verificar receta	VR
25	No confirma paciente	NP
26	Confirma paciente	SP
27	No se aplica tratamiento	NT
28	Compra medicamento	CM
29	Aplica tratamiento	AT

**Tabla 4.16.** Lista de actividades en el proceso de atención de cáncer de mama.

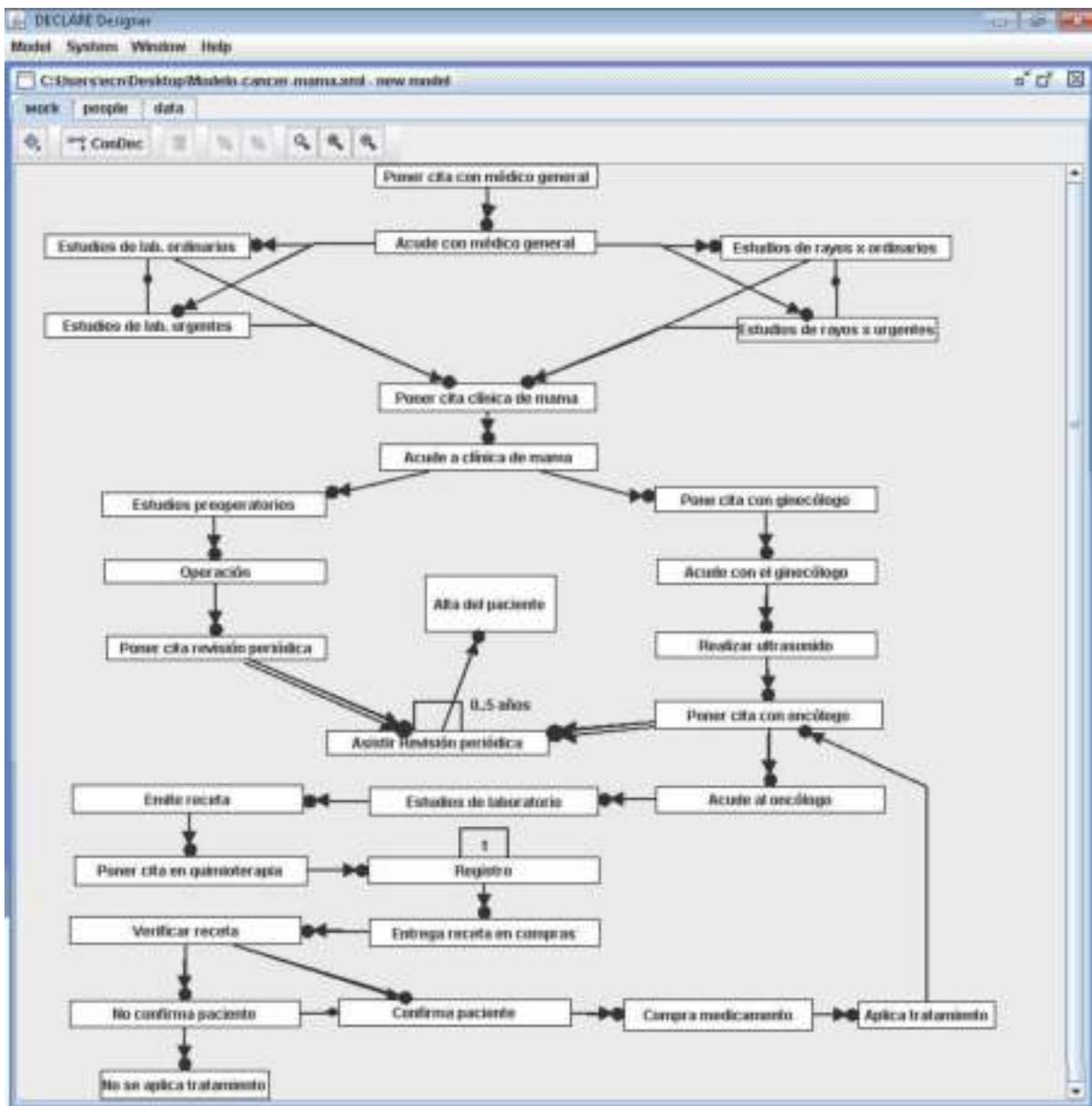


Figura 4.17. Modelo declarativo del proceso de atención de cáncer de mama.

En la tabla 4.18 muestra la semántica del modelo declare cuenta con: 35 restricciones entre actividades, semántica LTL y, la activación de la siguiente actividad del proceso de atención de cáncer mama de la institución, contiene 35 restricciones con 3 tipos de cardinalidades:

- A) relación precedence  $(A, B) = (\neg B \sqcup A) \vee \square(\neg B)$  significa que B sólo puede suceder después de que haya pasado A,

B) exclusive choice (A, B) =  $(\diamond A \vee \diamond B) \wedge \neg(\diamond A \wedge \diamond B)$  donde A o B tienen que pasar pero no ambas,

C) alternate precedence (A, B) =  $((\neg B \sqcup A) \vee \square(\neg B)) \wedge \square(B \rightarrow \bigcirc((\neg B \sqcup A) \vee \square(\neg B)))$  donde las actividades A y B deben alternarse sin repeticiones.

Restricciones	Semántica LTL	Activación
Precedence (PMG, AMG)	$(\neg \text{AMG} \sqcup \text{PMG}) \vee \square(\neg \text{AMG})$	AMG
Precedence (AMG, ELO)	$(\neg \text{ELO} \sqcup \text{AMG}) \vee \square(\neg \text{ELO})$	ELO
Precedence (AMG, ELU)	$(\neg \text{ELU} \sqcup \text{AMG}) \vee \square(\neg \text{ELU})$	ELU
Exclusive choice (ELO,ELU)	$(\diamond \text{ELO} \vee \diamond \text{ELU}) \wedge \neg(\diamond \text{ELO} \wedge \diamond \text{ELU})$	ELO, ELU
Precedence (AMG, ERO)	$(\neg \text{AMG} \sqcup \text{ERO}) \vee \square(\neg \text{ERO})$	ERO
Precedence (AMG, ERU)	$(\neg \text{AMG} \sqcup \text{ERU}) \vee \square(\neg \text{ERU})$	ERU
Exclusive choice (ERO,ERU)	$(\diamond \text{ERO} \vee \diamond \text{ERU}) \wedge \neg(\diamond \text{ERO} \wedge \diamond \text{ERU})$	ERO, ERU
Precedence (ELO, PCM)	$(\neg \text{PCM} \sqcup \text{ELO}) \vee \square(\neg \text{PCM})$	PCM
Precedence (ELU, PCM)	$(\neg \text{PCM} \sqcup \text{ELU}) \vee \square(\neg \text{PCM})$	PCM
Precedence (ERO, PCM)	$(\neg \text{PCM} \sqcup \text{ERO}) \vee \square(\neg \text{PCM})$	PCM
Precedence (ERU, PCM)	$(\neg \text{PCM} \sqcup \text{ERU}) \vee \square(\neg \text{PCM})$	PCM
Precedence (PCM, ACM)	$(\neg \text{ACM} \sqcup \text{PCM}) \vee \square(\neg \text{ACM})$	ACM
Precedence (ACM, EP)	$(\neg \text{EP} \sqcup \text{ACM}) \vee \square(\neg \text{EP})$	EP
Precedence (EP, OP)	$(\neg \text{OP} \sqcup \text{EP}) \vee \square(\neg \text{OP})$	OP
Precedence (OP, PRP)	$(\neg \text{PRP} \sqcup \text{OP}) \vee \square(\neg \text{PRP})$	PRP
Alternate precedence (PRP, ARP)	$((\neg \text{ARP} \sqcup \text{PRP}) \vee \square(\neg \text{ARP})) \wedge \square(\text{ARP} \rightarrow \bigcirc((\neg \text{ARP} \sqcup \text{PRP}) \vee \square(\neg \text{ARP})))$	PRP, ARP
Precedence (ARP, AP)	$(\neg \text{AP} \sqcup \text{ARP}) \vee \square(\neg \text{AP})$	AP
Precedence (ACM, PG)	$(\neg \text{PG} \sqcup \text{ACM}) \vee \square(\neg \text{PG})$	PG
Precedence (PG, AG)	$(\neg \text{AG} \sqcup \text{PG}) \vee \square(\neg \text{AG})$	AG
Precedence (AG, RU)	$(\neg \text{RU} \sqcup \text{AG}) \vee \square(\neg \text{RU})$	RU
Precedence (RU, PO)	$(\neg \text{PO} \sqcup \text{RU}) \vee \square(\neg \text{PO})$	PO
Alternate precedence (PO, ARP)	$((\neg \text{ARP} \sqcup \text{PO}) \vee \square(\neg \text{ARP})) \wedge \square(\text{ARP} \rightarrow \bigcirc((\neg \text{ARP} \sqcup \text{PO}) \vee \square(\neg \text{ARP})))$	PO, ARP
Precedence (PO, AO)	$(\neg \text{AO} \sqcup \text{PO}) \vee \square(\neg \text{AO})$	AO
Precedence (AO, EL)	$(\neg \text{EL} \sqcup \text{AO}) \vee \square(\neg \text{EL})$	EL
Precedence (EL, ER)	$(\neg \text{ER} \sqcup \text{EL}) \vee \square(\neg \text{ER})$	ER
Precedence (ER, PQ)	$(\neg \text{PQ} \sqcup \text{ER}) \vee \square(\neg \text{PQ})$	PQ
Precedence (PQ, R)	$(\neg \text{R} \sqcup \text{PQ}) \vee \square(\neg \text{R})$	R
Precedence (R, ERC)	$(\neg \text{ERC} \sqcup \text{R}) \vee \square(\neg \text{ERC})$	ERC
Precedence (ERC, VR)	$(\neg \text{VR} \sqcup \text{ERC}) \vee \square(\neg \text{VR})$	VR
Precedence (VR, NP)	$(\neg \text{NP} \sqcup \text{VR}) \vee \square(\neg \text{NP})$	NP
Precedence (VR, CP)	$(\neg \text{CP} \sqcup \text{VR}) \vee \square(\neg \text{CP})$	CP
Exclusive choice (NP,CP)	$(\diamond \text{NP} \vee \diamond \text{CP}) \wedge \neg(\diamond \text{NP} \wedge \diamond \text{CP})$	NP, CP

Precedence (CP, CM)	$(\neg CM \sqcup CP) \vee \square(\neg CM)$	CM
Precedence (CM, AT)	$(\neg AT \sqcup CM) \vee \square(\neg AT)$	AT
Precedence (AT, PO)	$(\neg PO \sqcup AT) \vee \square(\neg PO)$	PO

**Tabla 4.17.** Semántica TLT;  $\square$  significa siempre,  $\diamond$  significa eventualmente,  $\bigcirc$  significa enseguida de,  $(x \sqcup y)$  donde  $x$  debe esperar hasta  $y$ .

#### 4.4.2. Diseño del modelo de proceso (práctica clínica real)

Las modelos de ejecuciones reales parten de la extracción de conocimiento de los registros de eventos disponibles en los sistemas actuales de información. El descubrimiento del proceso permite descubrir eventualidades visuales en los procesos modelados (García et. al, 2016; Rovani et. al, 2015).

- Elección del software

Para aplicar técnicas de minería de procesos, fue esencial extraer registros de eventos, los datos extraídos por medio de entrevistas a pacientes de cáncer de mama de la institución necesitaron ser transformados, por lo que se utilizó XES (eXtensible Event Stream) como estándar de IEEE SA para la minería de procesos, siendo el formato compatible con herramientas como ProM, el algoritmo utilizado fue Visual Miner Inductivo que permitió detectar sucesos que puedan estar pasando en los procesos hospitalarios (Rovani et. al, 2015).

Se consideraron 5 variables para la creación del modelo (usuario, event\_id, actividad, inicio y fin) como se muestra en la tabla 4.18.

Variables	Tipo de dato
Usuario	String
Event_Id	Discreto
Actividad	String
Inicio	Timestamp
Fin	Timestamp

**Tabla 4.18.** Variables consideradas para el diseño del modelo de proceso.

Para la realización de cada log de eventos se tomó en cuenta las variables (tabla 4.19):

- Usuario: variable considera como string el cual indica los pacientes entrevistados (18 usuarios),
- event\_id: variable de tipo discreto que indica que cada registro es único,
- actividad: se consideró toda actividad o clase registrada durante las entrevistas, que representan cada paso que dio el usuario durante el proceso de atención de cáncer de mama, algunas actividades donde las pacientes no recordaban la fecha exacta se omitieron, por ejemplo “poner cita con médico general”, “poner cita con cirujano”, “poner cita con oncólogo”.
- Inicio y fin: son variables de tipo fecha, que representan la perspectiva del tiempo donde describen los tiempos de atención hospitalaria de los pacientes entrevistados.

Por lo tanto, se realizaron 18 logs de eventos, referente a cada paciente, en la tabla 4.19 se muestra un ejemplo de uno; las actividades que realizaron las pacientes en hospitales o laboratorios particulares fueron eliminados.

Usuario	Event_Id	Actividad	Inicio	Fin
Usuario_1	0	Poner cita con médico general	01/08/2016 00:00:00	01/08/2016 01:00:00
Usuario_1	1	Acude con médico general	15/08/2016 00:00:00	15/08/2016 01:00:00
Usuario_1	2	Diagnostico	15/08/2015 00:00:00	20/12/2016 00:00:00
Usuario_1	3	Mamografía	15/09/2016 00:00:00	15/09/2016 01:00:00
Usuario_1	4	Ultrasonido	15/09/2016 02:00:00	15/09/2016 03:00:00
Usuario_1	5	Acude con ginecólogo	15/10/2016 00:00:00	15/10/2016 01:00:00
Usuario_1	6	Mamografía	15/11/2016 00:00:00	15/11/2016 01:00:00
Usuario_1	7	Ultrasonido	15/11/2016 02:00:00	15/11/2016 03:00:00
Usuario_1	8	Acude a clínica de mama	03/12/2016 00:00:00	03/12/2016 01:00:00
Usuario_1	9	Internada	03/12/2016 02:00:00	17/12/2016 00:00:00
Usuario_1	10	Biopsia	14/12/2016 00:00:00	20/12/2016 00:00:00
Usuario_1	11	Acude al oncólogo	02/01/2017 00:00:00	02/01/2017 01:00:00
Usuario_1	12	Acude al oncólogo	20/01/2017 00:00:00	20/01/2017 01:00:00
Usuario_1	13	Recibe quimio roja	23/01/2017 00:00:00	13/02/2017 01:00:00
Usuario_1	14	Acude al oncólogo	13/02/2017 01:00:00	13/02/2017 02:00:00

Usuario_1	15	Recibe quimio roja	13/02/2017 02:00:00	06/03/2017 02:00:00
Usuario_1	16	Acude al oncólogo	06/03/2017 02:00:00	06/03/2017 03:00:00
Usuario_1	17	Recibe quimio roja	06/03/2017 03:00:00	27/03/2017 03:00:00
Usuario_1	18	Acude al oncólogo	27/03/2017 03:00:00	27/03/2017 04:00:00
Usuario_1	19	Recibe quimio roja	27/03/2017 04:00:00	17/04/2017 04:00:00
Usuario_1	20	Acude al oncólogo	17/04/2017 04:00:00	17/04/2017 05:00:00
Usuario_1	21	Recibe quimio blanca	17/04/2017 05:00:00	08/05/2017 05:00:00
Usuario_1	22	Acude al oncólogo	08/05/2017 05:00:00	08/05/2017 06:00:00
Usuario_1	23	Operación	09/06/2017 00:00:00	09/06/2017 07:12:00
Usuario_1	24	Acude al oncólogo	09/06/2017 07:12:00	09/06/2017 08:12:00
Usuario_1	25	Recibe quimio blanca	10/07/2017 00:00:00	31/07/2017 00:00:00
Usuario_1	26	Acude al oncólogo	31/07/2017 00:00:00	31/07/2017 01:00:00
Usuario_1	27	Recibe quimio blanca	31/07/2017 02:00:00	21/08/2017 02:00:00
Usuario_1	28	Acude al oncólogo	21/08/2017 02:00:00	21/08/2017 03:00:00
Usuario_1	29	Recibe quimio blanca	21/08/2017 03:00:00	11/09/2017 03:00:00
Usuario_1	30	Acude al oncólogo	11/09/2017 03:00:00	11/09/2017 04:00:00

**Tabla 4.19.** Ejemplo de log de eventos.

Se cargaron los 18 logs al software ProM donde se ejecutó el algoritmo “XES Event Log”, se obtuvo el siguiente análisis:

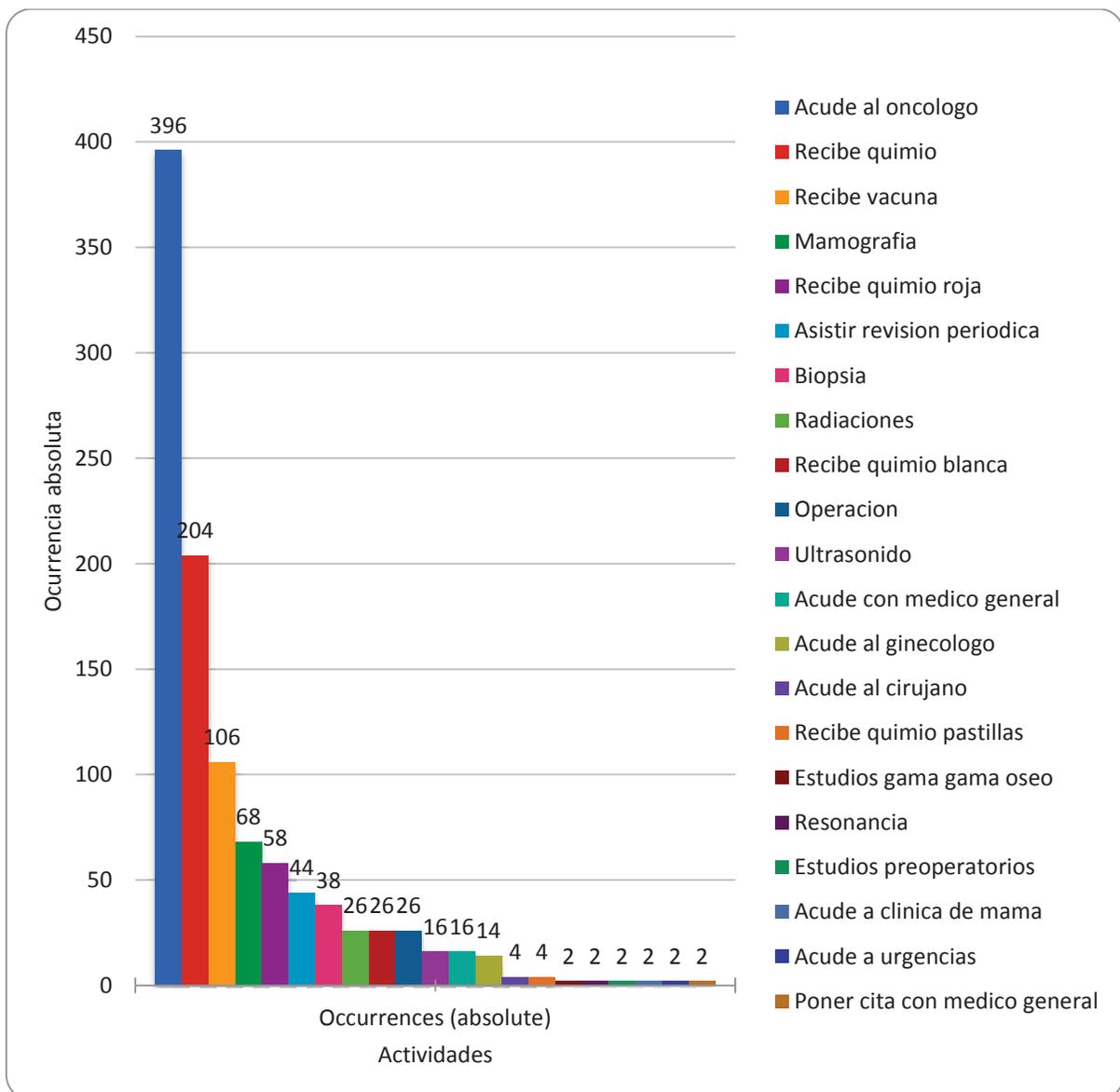
Se consideraron 18 instancias que representan a cada uno de los pacientes entrevistados con un total de 1058 eventos, las fechas de los eventos se encuentran entre las fechas del 21 de julio 2000 al 16 de agosto 2019 (tabla 4.20), con una media de 59 eventos por paciente.

<b>Eventos</b>	<b>Cantidad</b>
Total de instancias de procesos	<b>18</b>
Total de eventos	<b>1058</b>
Fecha de Inicio	<b>21 Julio 2000</b>
Fecha de fin	<b>16 Agosto 2019</b>
Media	<b>59 eventos</b>

**Tabla 4.20.** Resumen logs de eventos.

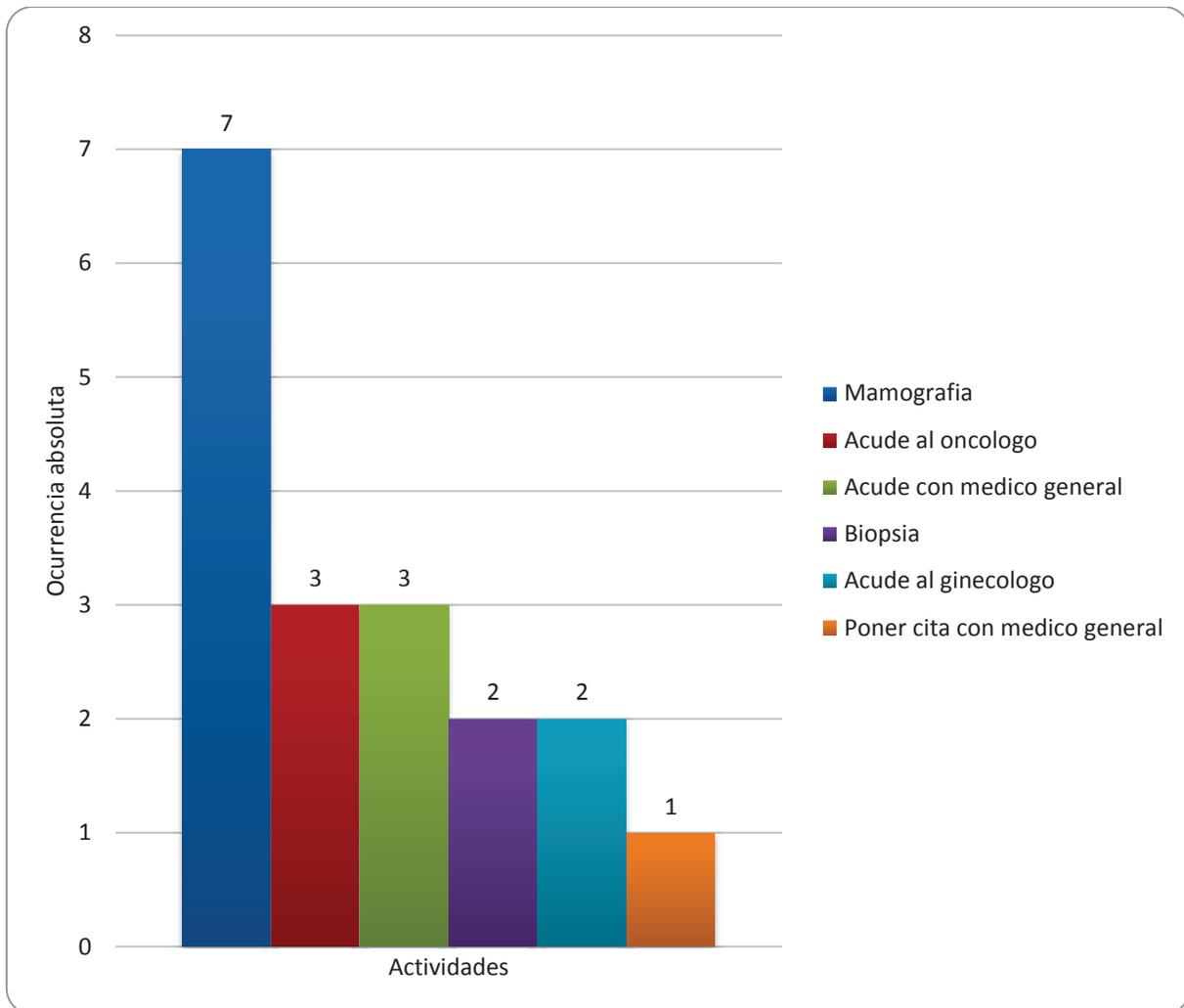
Se encontraron 21 clases o actividades únicas entre los 18 logs de eventos cargados a ProM; la actividad “Acude al oncólogo” es la que tiene mayor frecuencia con 396

(37.47%) ocurrencias, seguida de “recibe quimio” con 204 (19.28%) ocurrencias, “recibe vacuna” con 106 (10.02%) ocurrencias, “mamografía” con 68 (6.43%) ocurrencias; en las actividades con menos ocurrencia se encuentran “estudios gama óseo”, “resonancia”, “estudios preoperatorios”, “acude a clínica de mama”, “acude a urgencias”, “poner cita con médico general” con dos ocurrencias (0.19%), como se muestra en la figura 4.18.



**Figura 4.18.** Actividades recurrentes en el proceso de cáncer de mama.

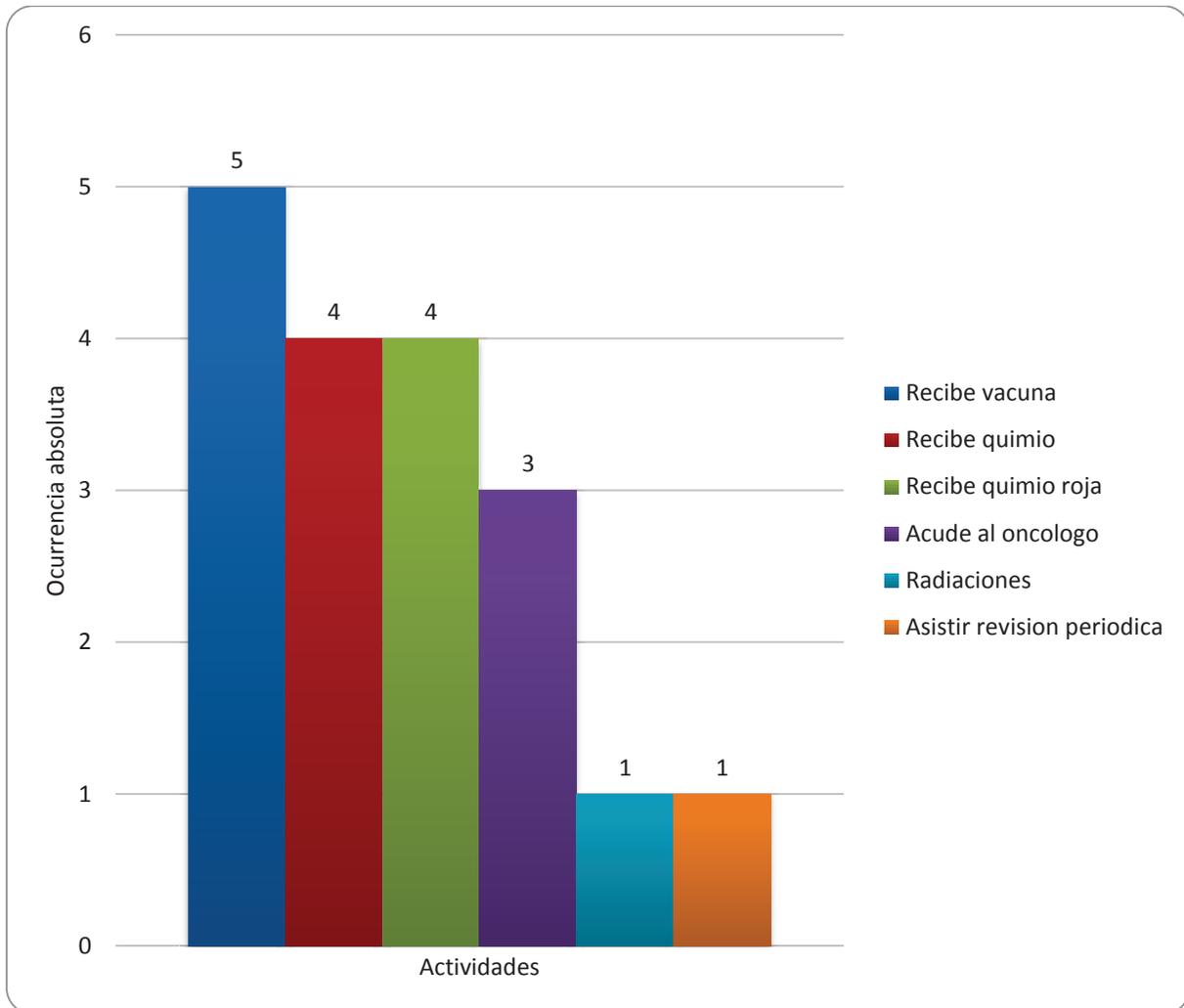
Se identificaron 6 actividades como iniciales del proceso: “mamografía” con presencia en 7 (38.89%) pacientes, “acude con médico general” y “acude al oncólogo” con 3 (16.67%) ocurrencias, “biopsia” y “acude al ginecólogo” con 2 (11.11%) ocurrencias y, “poner cita con médico general” con una ocurrencia (5.56%) como se observa en la figura 4.19.



**Figura 4.19.** Actividades iniciales del proceso de atención de cáncer de mama.

Para las actividades finales del proceso de atención de cáncer de mama se identificaron: con mayor frecuencia se detectó “recibe vacuna” en 5 (27.78%) pacientes, “recibe quimio” y “recibe quimio roja” en 4 (22.22%) pacientes, “acude al

oncólogo” con 3 (16.67%) pacientes y, tanto “radiaciones” como “asistir a revisión periódica” se presentaron en un (5.56%) paciente cada una (figura 4.20); estas actividades que representan la aplicación de tratamientos y chequeo de rutina.



**Figura 4.20.** Actividades finales del proceso de atención de cáncer de mama.

Se ejecutó el algoritmo Plugin Visual Miner para el descubrimiento del proceso de atención de cáncer de mama, donde se aprecia el proceso de manera visual con 846 (80%) paths o rutas del total (1058). Se obtuvo en el modelo 18 actividades; esto quiere decir que el 80% de los logs se ejecutaran de manera satisfactoria en este modelo descubierto.

El proceso inicia con la actividad poner cita con médico general, luego puede acudir al médico general, urgencias o a realizar ultrasonido; del ultrasonido el paciente puede acudir al ginecólogo o a la clínica de mama, el cual puede enviar al paciente a realizarse una resonancia o acudir al cirujano (figura 4.21) o, pasa a realizarse una biopsia, en algunas ocasiones los pacientes pasan a recibir tratamiento como quimioterapia (quimio roja, quimio en pastillas), las candidatas a operación se canalizan para estudios preoperatorios, aunque éstos estudios no se realizan a personas menores de 40 años, por lo que pueden saltarse este paso antes de la cirugía; después de la cirugía se pueden canalizar para recibir tratamiento si es necesario (quimio roja, quimio blanca, vacunas u otro tipo de tratamiento) (figura 4.22) a través del oncólogo; luego del tratamiento o de la cirugía las pacientes pasan a la actividad de revisión periódica cada 6 meses por 5 años donde se les realiza mamografías como parte de la prevención por reincidencia, pasados los 5 años se les da de alta y las revisiones periódicas se llevan a cabo cada año (figura 4.23).

Sin embargo, algunos pacientes pueden comenzar el proceso de atención acudiendo al oncólogo directamente, y son canalizadas con el tratamiento de quimio o vacuna (figura 4.23); algunos pacientes que acuden directamente al oncólogo son pacientes que fueron diagnosticados en instituciones particulares.

En la primera parte del modelo se obtuvieron 8 actividades (poner cita con médico general, acude con médico general, acude a urgencias, ultrasonido, acude al ginecólogo, acude a clínica de mama, resonancia, acude al cirujano) como se muestra en la figura 4.21.

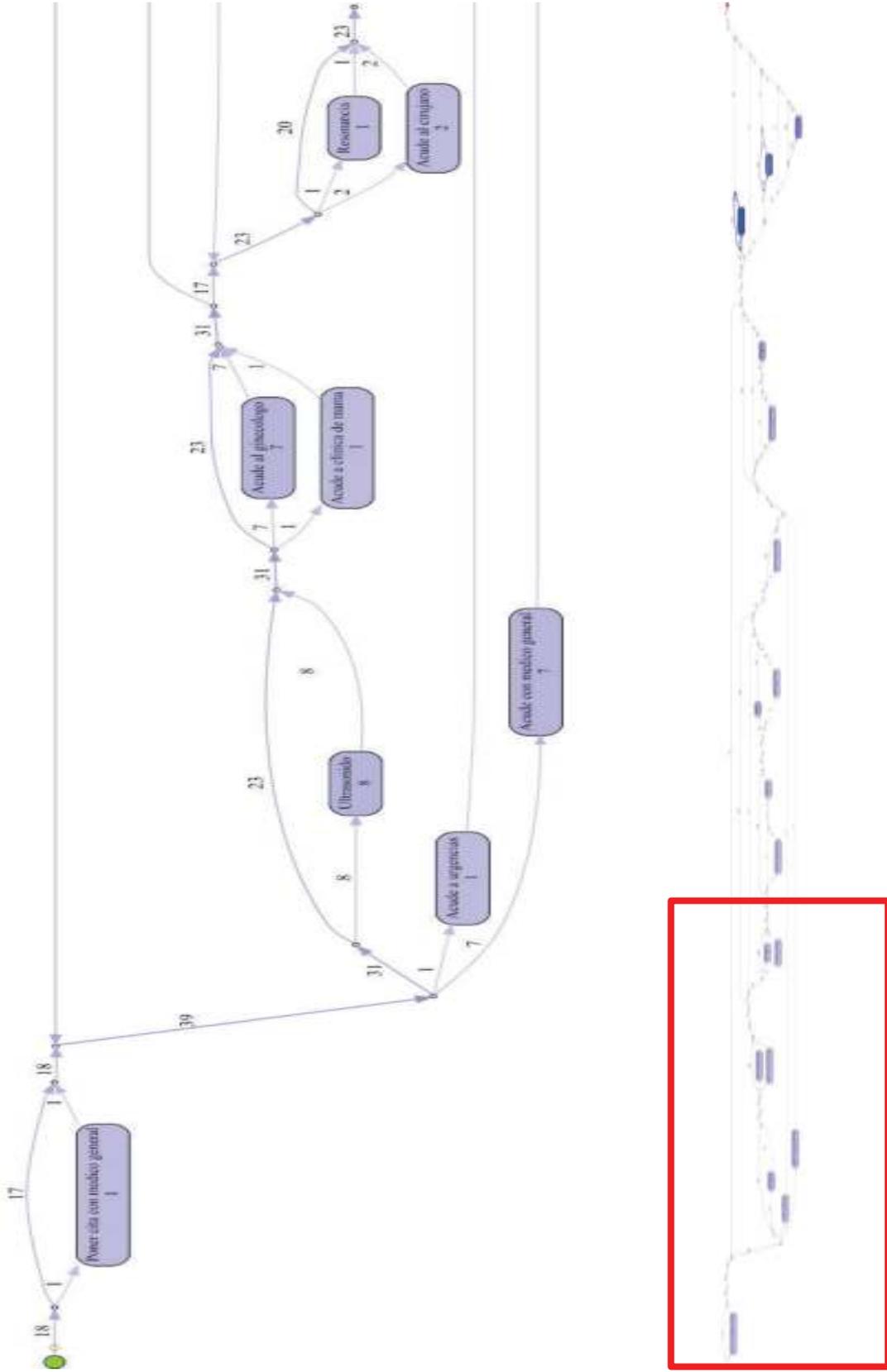


Figura 4.21. Descubrimiento del modelo de proceso de atención de cáncer de mama (parte 1).

En la segunda parte del modelo contempla 5 actividades preoperatorias (estudios preoperatorios, operación, recibe quimio roja, biopsia, recibe quimio pastillas) como se muestra en la figura 4.22.

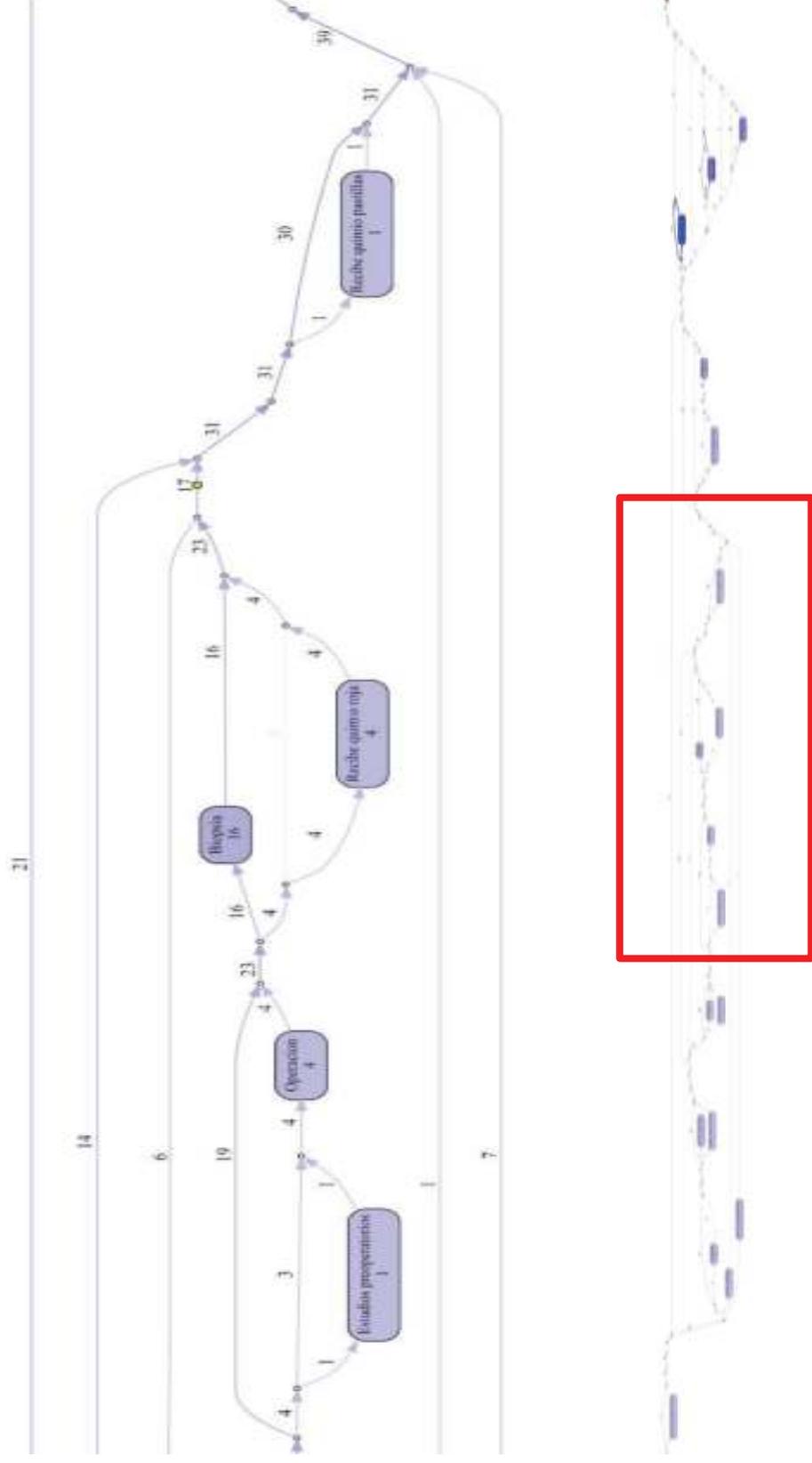


Figura 4.22. Descubrimiento del modelo de proceso de atención de cáncer de mama (parte 2).

En la tercera parte del modelo se tienen 5 actividades (asiste a revisión periódica, mamografía acude al oncólogo, recibe quimio, recibe vacuna) las cuales son las más frecuentes del proceso, como se muestra en la figura 4.23.

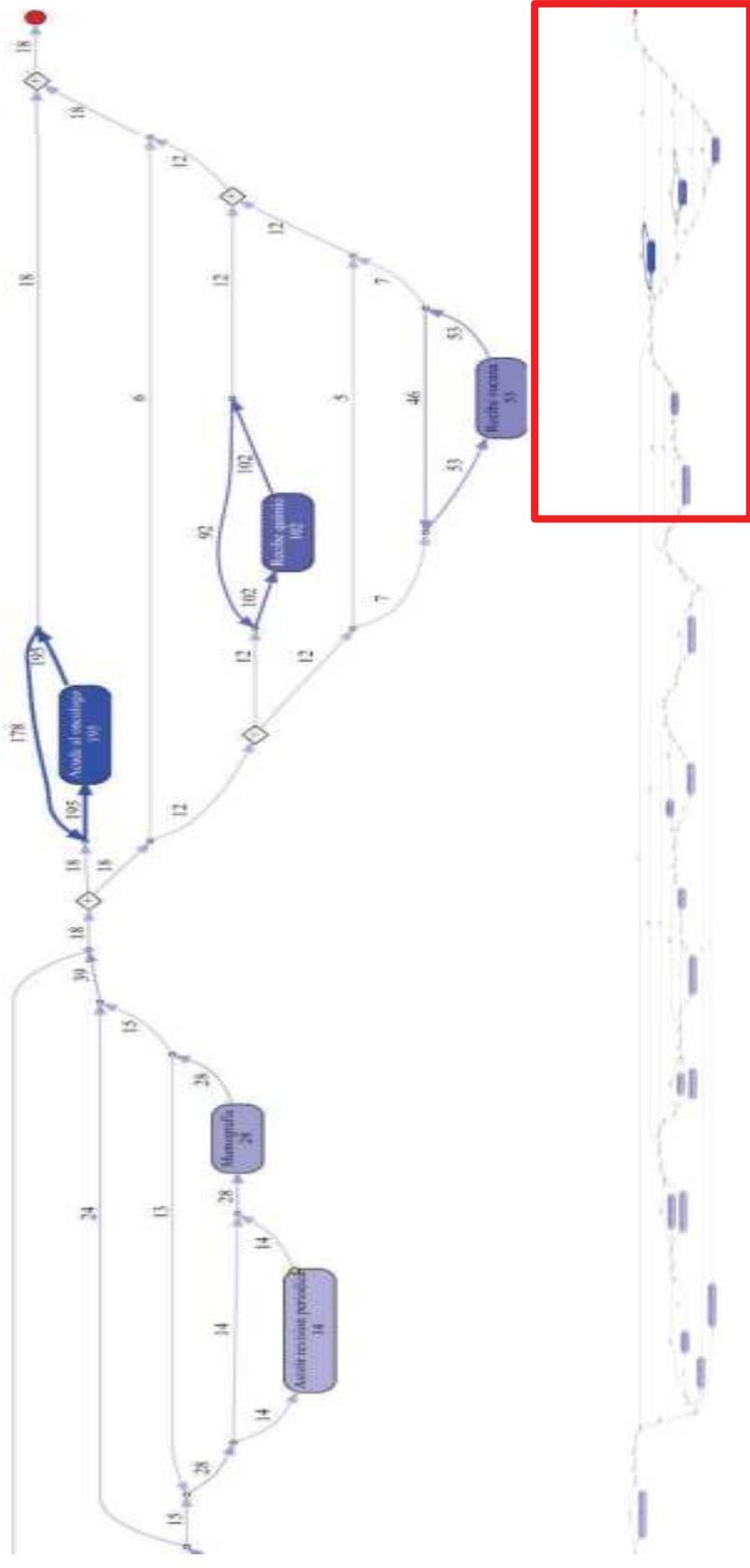


Figura 4.23. Descubrimiento del modelo de proceso de atención de cáncer de mama (parte 3).

## 4.5. Interpretación y Evaluación

El resultado obtenido con Visual Miner Inductivo arrojó en forma de nodos las actividades o transiciones que forman parte del proceso de atención de cáncer de mama; en el cual se pueden apreciar con mayor facilidad las actividades frecuentes o no, cuellos de botella y desviaciones que puedan estar afectando al funcionamiento de la institución e influir en la atención a pacientes de cáncer de mama (Rovani et. al, 2015).

Se tomó el modelo declarativo (basado en guías clínicas) como referencia y así poder cotejarlo con cada uno de los procesos de las 18 pacientes (modelo de proceso descubierto en Visual Miner Inductivo). Para realizar esta evaluación se utilizó el algoritmo Declare Replay para verificar la idoneidad del modelo descubierto con el modelo basado en guías clínicas, se realizó la evaluación con las métricas de modelos fitness, obteniendo un 0.7948 de idoneidad, es decir que el proceso seguido por las 18 pacientes entrevistadas tiene 79.48% de similitud con la guía clínica de la institución; este algoritmo se ejecuta dentro de la herramienta ProM.

Se encontraron algunas actividades con ciertas desviaciones o violaciones en el proceso que se muestran en la figura 4.24:

1. existen eventos o actividades que estaban contempladas dentro del modelo de guías clínicas de la institución y fueron omitidas durante el proceso, por ejemplo “Poner cita con médico general”, “Poner cita clínica de mama”, “Poner cita con oncólogo”, en algunos casos por ser menores de 40 años no se realizan “Estudios preoperatorios”
2. hay actividades que se realizaron y debieron ser precedidas de otras de manera obligatoria y éstas omitieron, el por ejemplo si acude al oncólogo la actividad que la precede es “Poner cita oncólogo”, esto se debió a que algunas pacientes no recordaban la fecha exacta en la que pusieron la cita con el oncólogo por lo

que esta actividad se omitió en la limpieza de datos; otros casos fueron que algunas pacientes no se realizaron exámenes de laboratorio, sino que se los hicieron de manera particular, por lo que acudían a la institución con los resultados y eran canalizadas de manera pronta al oncólogo y por con siguiente de manera rápida a cirugía y/o a quimioterapia.



**Figura 4.24.** El modelo de proceso presenta una idoneidad de 79.48%, el cual presenta algunas desviaciones y violaciones con respecto al modelo basado en guías clínicas.

Otra de las evaluaciones para evaluar el desempeño del proceso se calculó en base a la dimensión tiempo, identificando 3 indicadores:

- el tiempo de espera promedio (Average wait time): el tiempo promedio que un paciente espera (en días) para una prueba de diagnóstico, cirugía o tratamiento (Gutierrez, 2009),
- el recíproco (average service time): tiempo promedio requerido (en días) para atender a un paciente,
- tiempo de permanencia (average sojourn time): tiempo promedio (en días) entre la tarea anterior y una tarea determinada (Mans et. al, 2008).

La actividad que ha sido más duradera es la de “recibe quimio pastillas”, se estima que fue un tratamiento que recibió un paciente durante casi 5 años, el tiempo de espera promedio de espera es de 30 días, mientras que el recíproco tiene un promedio de 1675 días y, el tiempo de permanencia 1705 días. Los indicadores más bajos los tienen las actividades “mamografía” y “acude al cirujano”; la actividad “operación” (tiempo de espera = 10 y 8 días; recíproco=1 y 0 días; tiempo de permanencia =3 y 8 días respectivamente).

Las actividades relacionadas con el oncólogo y tratamientos los resultados son: “acude al oncólogo” (tiempo de espera = 35 días; recíproco = 0 días; tiempo de permanencia = 35 días); “recibe quimio” (tiempo de espera = 30 días; recíproco= 22 días; tiempo de permanencia = 51 días); “recibe vacuna” (tiempo de espera = 65 días; recíproco=18 días; tiempo de permanencia = 84 días); “recibe quimio roja” (tiempo de espera = 17 días; recíproco=22 días; tiempo de permanencia = 40 días); “recibe quimio blanca” (tiempo de espera = 0 días; recíproco=21 días; tiempo de permanencia = 21 días), como se aprecia en la figura 4.25.

Los datos anteriormente mencionados son resultado del análisis del proceso de acuerdo a las entrevistas. Esta información requiere validación de un especialista y contrastarla con las recomendaciones de la NOM.

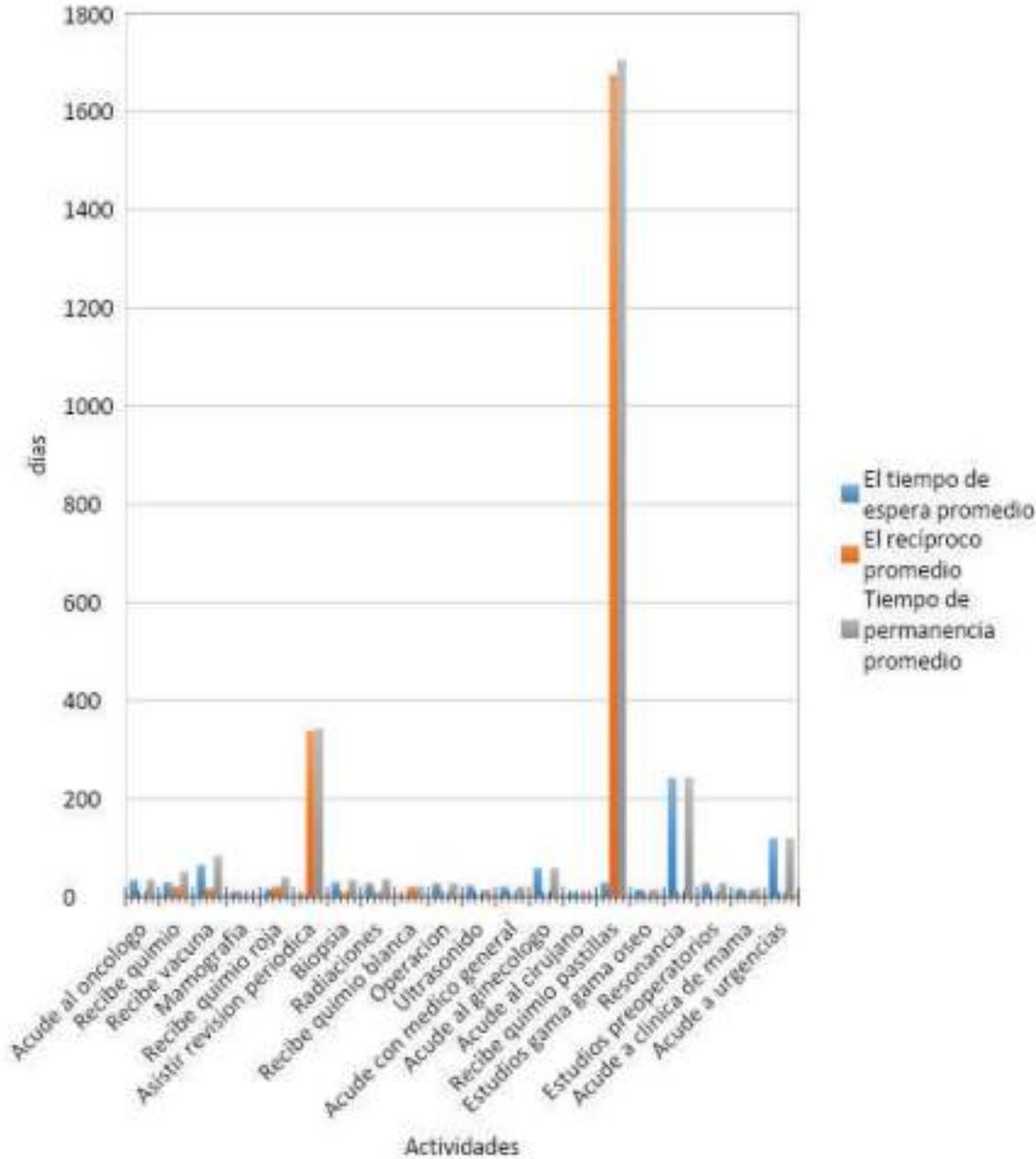


Figura 4.25. Evaluación del proceso de atención de cáncer de mama mediante 3 indicadores: tiempo de espera promedio, el recíproco y tiempo de permanencia.

En la tabla 4.21. se muestran las actividades que se consideran cruciales para el proceso de atención de cáncer de mama:

1. días para primera consulta
2. días desde la primera consulta hasta la mamografía
3. días para la cita con el oncólogo después de la mamografía
4. días desde que se realizó la biopsia hasta acudir al oncólogo
5. días desde que se realizó la biopsia hasta que le entregaron el resultado
6. desde que se dio cuenta hasta que le dieron el tratamiento
7. desde que se dio cuenta hasta la operación, o desde la primera cita o mamografía (se consideraron fechas de actividades realizadas en hospitales particulares)
8. días desde que el entregaron el resultado biopsia hasta que se le aplicó el tratamiento
9. días desde que se le entrega el resultado biopsia hasta que se realizaron la operación

Se consideran que las dos primeras son las que arrojan resultados más importantes.

1. días para primera consulta algunas pacientes tardaron más de 4 años en acudir al médico, tiene una desviación estándar de 1289.5 días (3.53 años) y una media de 666.5 días (1.82 años);
2. la otra actividad que puede considerarse un tiempo razonable es la de días desde la primera consulta hasta la mamografía, cuya desviación estándar fue de 10.6 días y una media de 8 días, como se muestra en la tabla 4.21. El promedio de días entre cada cita es de 20.64 según datos de la encuesta.

<b>concepto</b>	<b>Total de respuestas</b>	<b>media</b>	<b>min</b>	<b>max</b>	<b>desviación estándar</b>
<i>días para primera consulta</i>	12.0	666.5	4.0	4380.0	1289.5
<i>días desde la primera consulta hasta la mamografía</i>	8.0	8.0	0.0	31.0	10.6
<i>cita con el oncólogo después de la mamografía</i>	14.0	72.1	0.0	249.0	93.7
<i>biopsia- oncólogo</i>	12.0	18.6	0.0	114.0	32.1
<i>biopsia -resultado</i>	17.0	9.1	0.0	15.0	5.3
<i>desde que se dio cuenta hasta que le dieron el tratamiento</i>	16.0	678.8	45.0	4394.0	1128.3
<i>desde que se dio cuenta hasta la operación, o desde la primera cita o mamografía</i>	13.0	378.3	15.0	1915.0	507.3
<i>resultado biopsia - tratamiento</i>	17.0	49.2	2.0	167.0	47.1
<i>resultado biopsia- operación</i>	12.0	123.4	6.0	320.0	107.8

**Tabla 4.21.** Actividades consideradas como cruciales.

Con el fin de tener una referencia del proceso que se sigue en la institución los resultados obtenidos de la encuesta también fueron contrastados con lo especificado por la NOM-041-SSA2-2011 y, se encontraron dos indicadores relevantes “Oportunidad de tratamiento” con 26.3%, donde el porcentaje de referencia es debe ser mayor de 90% según la norma, el otro indicador fue “sobrevida” con 25% como se aprecia en la tabla 4.22.

<i>Indicador</i>	Fórmula	Estándar	resultado
<b>Oportunidad para el diagnóstico</b>	((Total de mujeres referidas con diagnóstico definitivo en menos de tres citas)/Total de mujeres referidas para evaluación diagnóstica)*100	> 95%	60 %
<b>Oportunidad para la notificación del resultado</b>	((Total de mujeres con resultado notificado en menos de 21 días)/Total de mastografías de tamizaje realizadas)*100	> 95%	100 %
<b>Oportunidad de la referencia para tratamiento</b>	((Total de mujeres con diagnóstico confirmatorio de cáncer de mama referidas a un centro oncológico en menos de 10 días hábiles)/Total de mujeres con diagnóstico confirmatorio de cáncer de mama)*100	> 90%	77.8 %
<b>Oportunidad de tratamiento</b>	((Total de mujeres con diagnóstico de cáncer que iniciaron tratamiento en menos de 15 días)/Total de mujeres con diagnóstico de cáncer que iniciaron tratamiento)*100	> 90%	26.3%
<b>Sobrevida</b>	(No. De mujeres vivas a 1,3 y 5 años del diagnóstico, según etapa) /No. De mujeres con cáncer en seguimiento a 1, 3 y 5 años del diagnóstico según etapa *Sobrevida: se tomó todas las pacientes como un único cáncer (algunas pacientes no sabían qué tipo de cáncer padecían o padecen) y se tomó de 1 a 3 años	0 > 99%; I > 93%; II > 72%; III>41%; IV>18%	25%

**Tabla 4.22.** Indicadores para evaluar el desempeño y proceso según la NOM-041-SSA2-2011

Con la encuesta realizada fue posible conocer paso a paso el proceso de atención recibido por los pacientes, pero además se realizó un análisis cualitativo permitiendo conocer las percepciones del cliente acerca de los estilos de prácticas, de los aspectos administrativos, atención brindada por parte del personal médico, hospitalización y enfermería.

## Resumen de resultados de la encuesta

Para el análisis cualitativo que se realizó se tomaron en cuenta los 19 pacientes y, se obtuvieron los siguientes resultados:

1. En el área de tiempos en general se obtuvieron: la pregunta ¿cuánto tiempo pasó desde que se descubrió o palpó algo anormal hasta que fue a consulta?, los datos que arrojó es que las pacientes están demorando hasta 4 años en asistir a consulta, por lo que pudiera indicar que las campañas preventivas debieran no sólo estar enfocadas en exploración sino también en fomentar que las pacientes asistan regularmente al médico (figura 4.26).



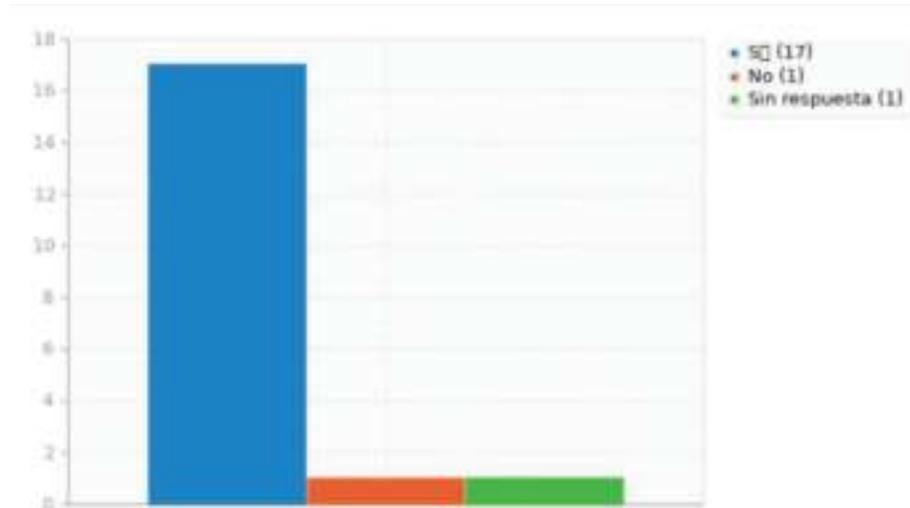
**Figura 4.26.** Resumen campo G01: tiempos en que tarda un paciente en asistir al médico por primera vez desde el momento que se detectó algo anormal

Otra de las preguntas que arrojaron datos relevantes fue: ¿todas las sesiones de tratamiento se aplicaron en tiempo de acuerdo a lo indicado por el médico?, con un 87.47% de las pacientes entrevistadas afirman que el tratamiento fue aplicado en tiempo, contra un 5.29% que no se aplicó en tiempo y un 5.29% restante no tuvo respuesta como se muestra en la figura 4.27.

## Resumen de campo para G13

12. ¿Todas sesiones de tratamiento se aplicaron en tiempo de acuerdo a lo indicado por el médico?

Opción	Cuenta	Porcentaje
Sí (Y)	17	89.47%
No (N)	1	5.26%
Sin respuesta	1	5.26%

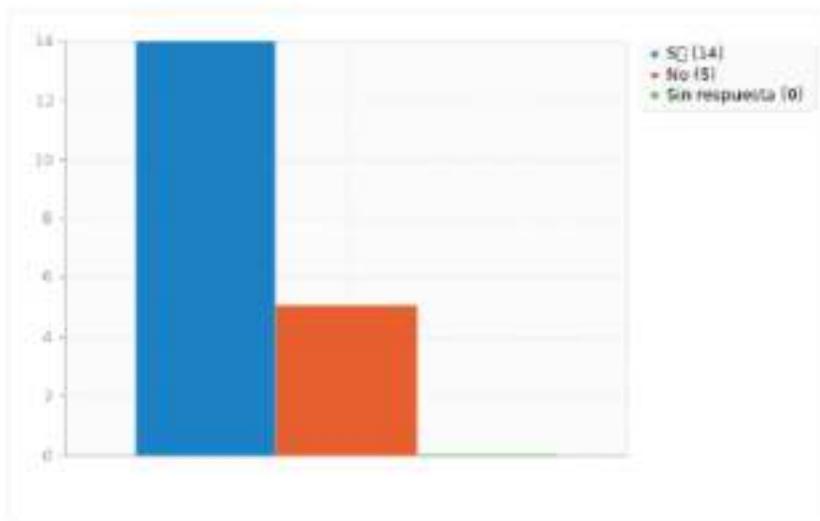


**Figura 4.27.** Resumen campo G13: ¿el tratamiento fue aplicado en tiempo indicado por el médico?

Un punto de mejora para la institución lo arroja la pregunta que indica que las pacientes con un 26.32% no están recibiendo información sobre cáncer de mama por parte del instituto u otros medios, el resto con 73.68% indica que sí ha recibido información sobre la enfermedad como lo indica la figura 4.28.

## 20. ¿Ha recibido información sobre la enfermedad por parte del instituto u otros medios?

Opción	Cuenta	Porcentaje
Sí (Y)	14	73.68%
No (N)	5	26.32%
Sin respuesta	0	0.00%



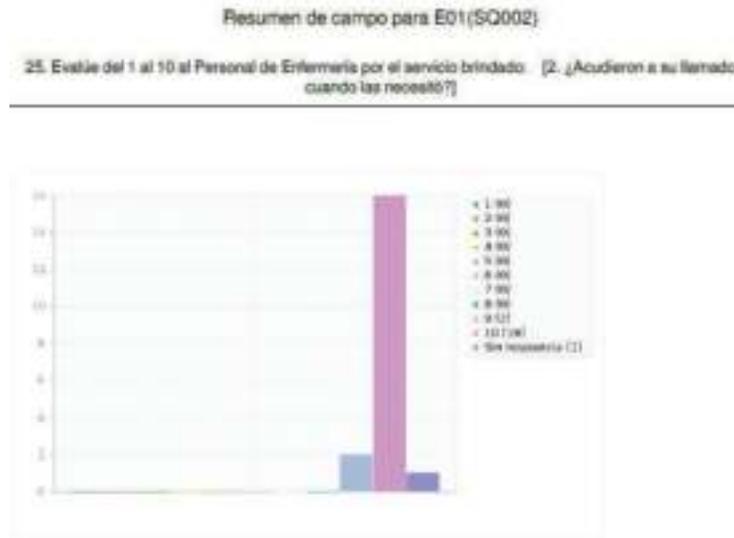
**Figura 4.28.** Resumen de campo para I01, ¿Ha recibido información sobre la enfermedad por parte del instituto u otros medios?

- En la sección de personal de enfermería, se encontraron algunos resultados relevantes, el personal de enfermería se refiere al personal de quimioterapia, las calificaciones brindadas a este personal fueron muy buenas, algunas excelentes, de los comentarios recibidos fueron acerca de la sala, es un espacio reducido para la cantidad de pacientes que hay, de igual forma el comentario de mejora para la institución fue hacia los muebles existentes que pudieran ser cambiados.

En una escala de 1 a 10 de calificación, los resultados obtenidos para enfermería rondan arriba del 9.5. Para el cálculo del promedio de calificación no se tomaron en cuenta a las pacientes que omitieron responder alguna pregunta

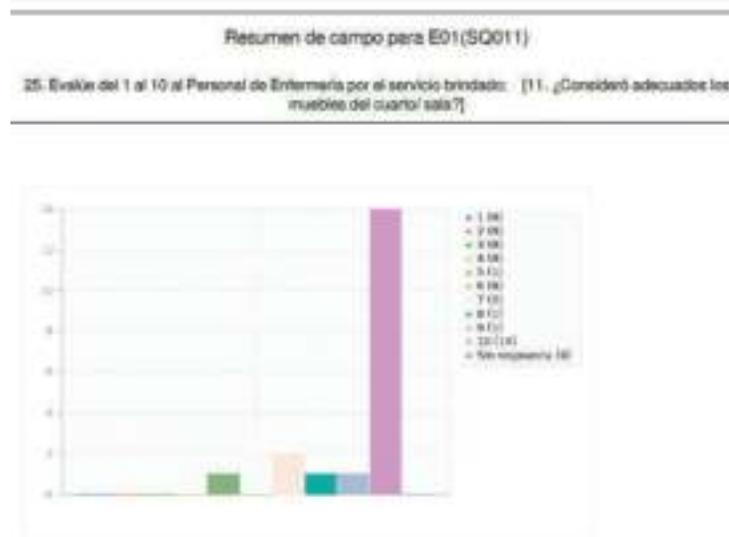
debido a que no aplicaba a su proceso. A continuación, se enlistan algunos resultados:

- Resumen de campo E01, ¿Acudieron a su llamado cuando las necesitó?: el promedio de calificación fue de 9.8, como se muestra en la figura 4.28.



**Figura 4.29.** Resumen de campo E01, ¿Acudieron a su llamado cuando las necesitó?

- Resumen de campo E01, ¿Consideró adecuados los muebles del cuarto/sala?: el promedio de calificación fue de 9.26, como se muestra en la figura 4.30.



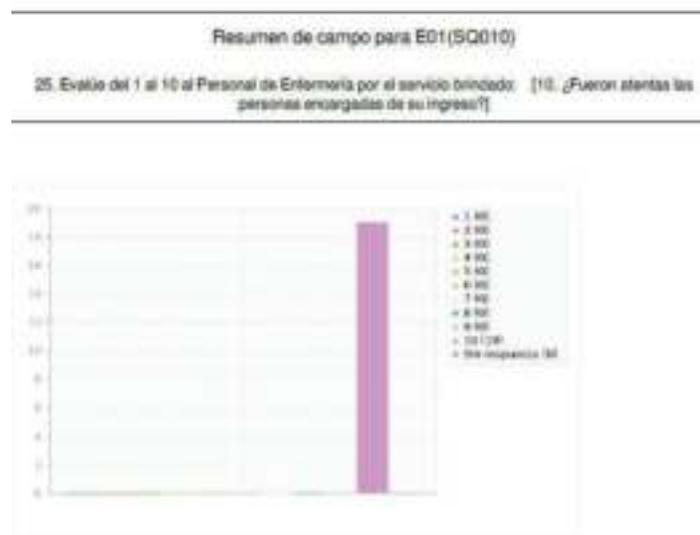
**Figura 4.30.** Resumen de campo E01, ¿Consideró adecuados los muebles del cuarto/sala?

- Resumen de campo E01, en general, ¿diría que el personal de enfermería le ha demostrado interés?: obtuvo una calificación perfecta de 10, como se muestra en la figura 4.31.



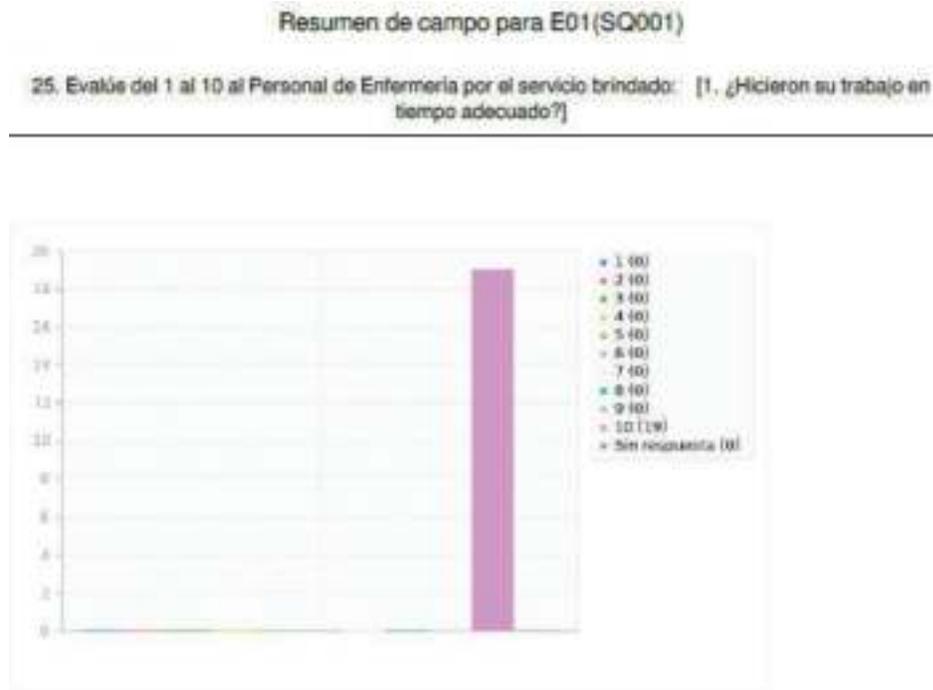
**Figura 4.31.** Resumen de campo E01 en general, ¿diría que el personal de enfermería le ha demostrado interés?

- Resumen de campo E01, ¿fueron atentas las personas encargadas de su ingreso?: obtuvo una calificación perfecta de 10, como se muestra en la figura 4.32.



**Figura 4.32.** Resumen de campo E01, ¿fueron atentas las personas encargadas de su ingreso?:

- Resumen de campo E01, ¿hicieron su trabajo en tiempo adecuado?: obtuvo una calificación perfecta de 10, como se muestra en la figura 4.33.



**Figura 4.33.** Resumen de campo E01, ¿hicieron su trabajo en tiempo adecuado?

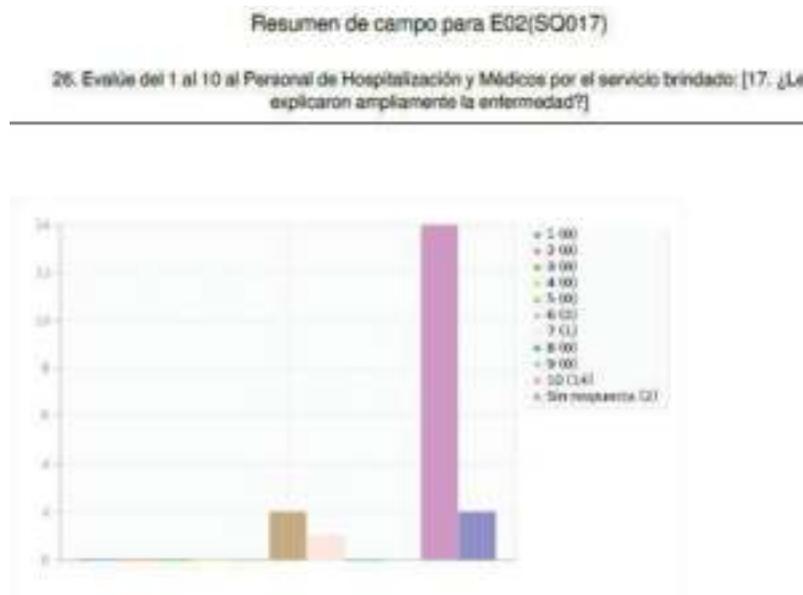
- Personal de Hospitalización: este rubro tuvo algunos comentarios para mejorar como: tuvo 7.7 en promedio de calificación de comodidad la cama de hospitalización, otra de las calificaciones bajas fue la limpieza con 8.5. Sin embargo, los puntos en que las pacientes consideraron arriba de 9 de calificación fueron: la explicación que se les dio al ingresar sobre la enfermedad fue considerada, así como también la aplicación que se les da a las pacientes sobre el procedimiento o aplicación del tratamiento fue realizada de manera correcta, el trato que se les dio a las pacientes el personal de hospitalización. A continuación, se muestran estos resultados.

- Resumen de campo E03, ¿fue cómoda la cama?: obtuvo una calificación promedio de 9.43, como se muestra en la figura 4.34.



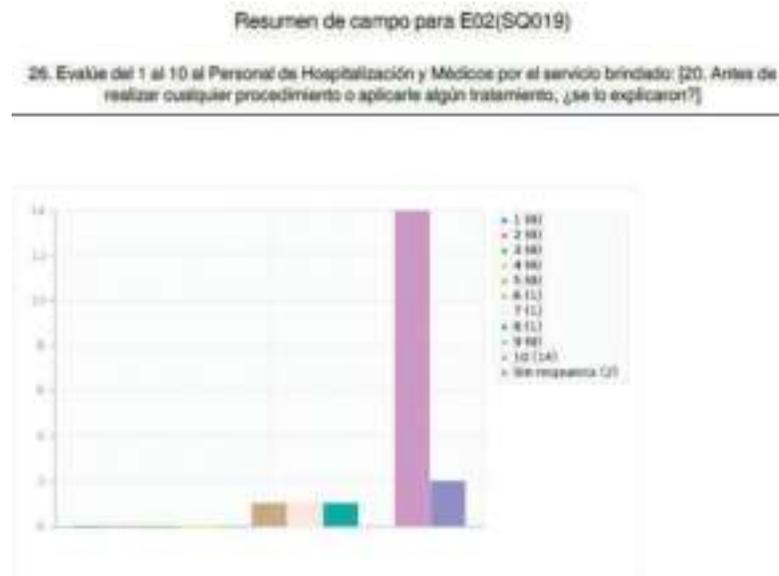
**Figura 4.34.** Resumen de campo E03, ¿fue cómoda la cama?

- Resumen de campo E02, ¿le explicaron ampliamente la enfermedad?: obtuvo una calificación promedio de 9.35, como se muestra en la figura 4.35.



**Figura 4.35.** Resumen de campo E02, ¿Le explicaron ampliamente la enfermedad?

- Resumen de campo E02, antes de realizar cualquier procedimiento o aplicarle algún tratamiento, ¿se lo explicaron?: obtuvo una calificación promedio de 9.47, como se muestra en la figura 4.36.



**Figura 4.36.** Resumen de campo E02, antes de realizar cualquier procedimiento o aplicarle algún tratamiento, ¿se lo explicaron?

- Resumen de campo E02, ¿cómo fue el trato demostrado por el personal de las comidas?: obtuvo una calificación promedio de 9.29, como se muestra en la figura 4.37.



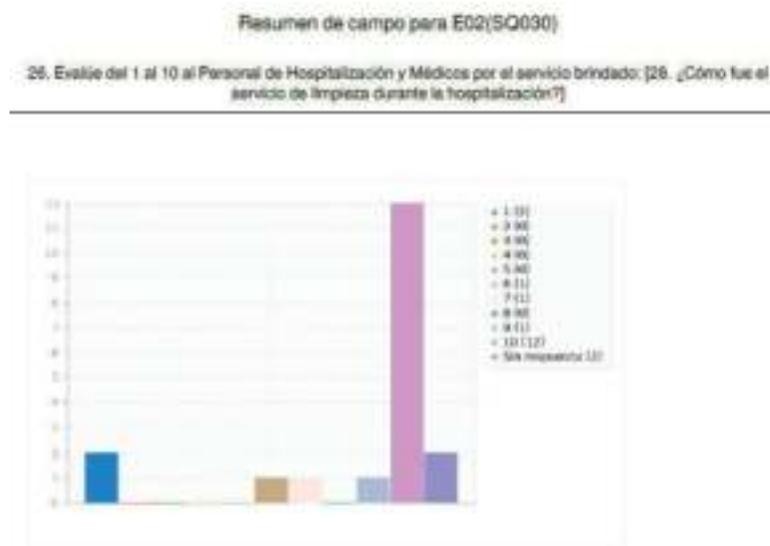
**Figura 4.37.** Resumen de campo E02, ¿cómo fue el trato demostrado por el personal de las comidas?

- Resumen de campo E02, ¿le dijeron qué artículos de uso personal necesitaría traer usted y cuáles le proporcionaría el hospital?: obtuvo una calificación promedio de 9, como se muestra en la figura 4.38.



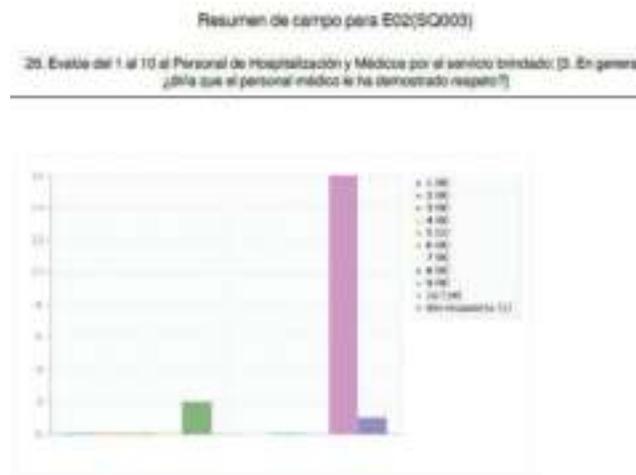
**Figura 4.38.** Resumen de campo E02, ¿le dijeron qué artículos de uso personal necesitaría traer usted y cuáles le proporcionaría el hospital?

- Resumen de campo E02, ¿cómo fue el servicio de limpieza durante la hospitalización?: obtuvo una calificación promedio de 8.47, como se muestra en la figura 4.39.



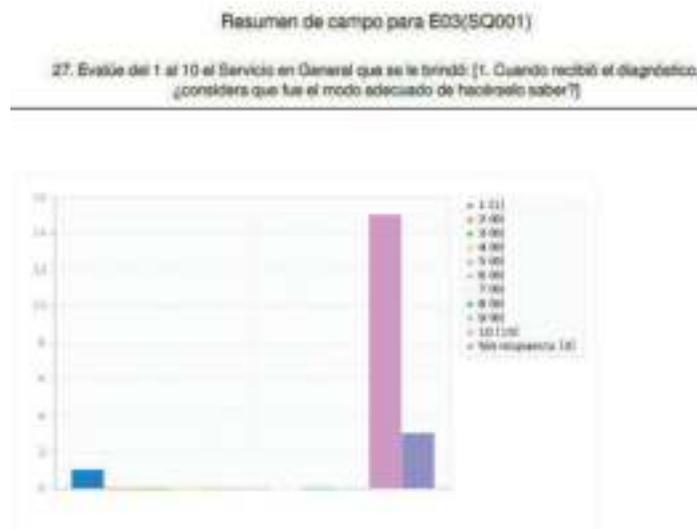
**Figura 4.39.** Resumen de campo E02, ¿cómo fue el servicio de limpieza durante la hospitalización?

4. Personal médico: en esta sección las preguntas más relevantes se encuentran:
- Resumen de campo E02. ¿diría que el personal médico le ha demostrado respeto?: obtuvo una calificación promedio de 9.44, como se muestra en la figura 4.40.



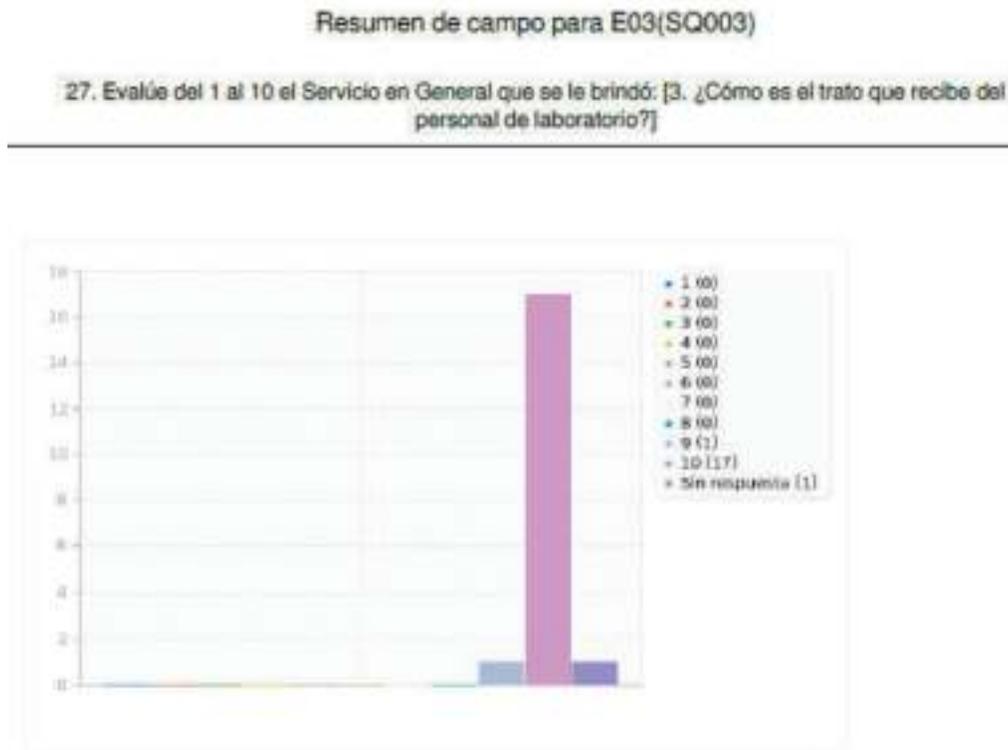
**Figura 4.40.** Resumen de campo E03. ¿diría que el personal médico le ha demostrado respeto?

- Resumen de campo E02, cuando recibió el diagnóstico, ¿Considera que fue el modo adecuado de hacérselo saber?: obtuvo una calificación promedio de 9.43, como se muestra en la figura 4.41.



**Figura 4.41.** Resumen de campo E03, cuando recibió el diagnóstico, ¿Considera que fue el modo adecuado de hacérselo saber?

5. Personal de laboratorio: en los que se refiere al trato percibido por el personal de laboratorio se evaluó con 9.94 de promedio, como se muestra en la figura 4.42.



**Figura 4.42.** Resumen de campo E03, ¿cómo es el trato que recibe del personal de laboratorio?

6. En la sección donde se cuestionaba de manera abierta las necesidades y recomendaciones hacia la institución arrojó algunos puntos de mejora como: se requiere mayor información o asesoría sobre regímenes alimenticios a seguir durante el proceso, así como estrategias de supervivencia, estilos de vida y rutinas de ejercicios permitidos, así como información en general sobre la enfermedad; se solicita modificar los criterios para realizar estudios de mamografía a menores de 40 años cuando hay riesgo por factor genético, servicio de tanatología, disponibilidad de medicamentos, así como mayor sensibilidad por parte de los médicos al dar la noticia (tabla 4.23).

**Necesidades y recomendaciones**

Información sobre dieta adecuada y estrategias de supervivencia (estilo de vida, dieta, ejercicio)

Modificar criterio para enviar a estudios de mamografía a menores de 40 años cuando hay riesgo por factor genético

Información sobre la enfermedad

Servicio de tanatología

Disponibilidad de Medicamentos

Simplificar trámites

Mayor sensibilidad a la hora de dar la noticia por parte de los médicos

**Tabla 4.23.** Necesidades y recomendaciones que realizó la paciente hacia la institución

Adicionalmente encontrará en el anexo 7.3 otros resultados que arrojó la encuesta.

## 4.6. Implementación

Se presentaron ante las autoridades de la institución una serie de reportes: utilizando gráficas, se encontraba lo más destacado del análisis descriptivo de registros, así como los resultados de la aplicación del algoritmo K-Means (georeferenciación), visual miner (proceso), así como también se presentaron los resultados obtenidos del análisis cualitativo obtenidos de la encuesta; a continuación, se enlistan algunos de los resultados por cada rubro.

- Existe incidencia de padecimientos de la mama en infantes menores de 1 año, la mayoría del sexo femenino.
- La distribución geográfica de los casos en la ciudad de Hermosillo muestra un foco rojo en el sur y poniente de la ciudad, especialmente en la colonia Altares y Villa Bonita.
- La mayor frecuencia de padecimientos de la mama por grupo de edad en el
- sexo femenino se concentra en las edades de los 35-49 años, mientras que el sexo masculino se da en edades tempranas de los 10-18 años contrario al sexo femenino.
- El número de casos de cáncer va al alza.

- El análisis de procesos arrojó sus principales actividades, siendo una de ellas el tiempo que tarda la paciente en asistir al médico desde que se observa alguna anomalía hasta que va al médico algunas pacientes dejaron pasar días o semanas antes de ir con el médico después de palpar una anomalía en la mama, mientras que otras dejaron pasar años, que puede indicar un nuevo enfoque de prevención no sólo para explorar, sino que también acudan al médico con regularidad; mientras los días que se tardan desde que el paciente acude por primera vez a consulta hasta la realización de la mamografía tuvo una desviación estándar de 10.6 días que se consideran razonables.
- En el análisis cualitativo se evaluó el trato que recibe el paciente del personal de la institución del cual se obtuvieron algunas retroalimentaciones que posibilitan la mejora del proceso de atención de cáncer de mama.

## **5. CONCLUSIONES RECOMENDACIONES Y TRABAJOS FUTUROS**

Según las experiencias vividas a lo largo del proceso de implantación de la metodología, surgen las conclusiones finales y, recomendaciones que buscan mejorar los procesos de atención de cáncer de mama en la institución de salud.

### **5.1. Conclusiones**

Este proyecto forma parte de uno integral para dar soporte a la toma de decisiones y planificación de acciones en términos de medicina preventiva en el Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado de Sonora (ISSSTESON), que proporciona atención médica a más de 180,000 derechohabientes.

El objetivo principal de esta investigación fue desarrollar un modelo que evaluara el proceso de atención de cáncer de mama utilizando los datos disponibles en el HIS de la institución que permitiera identificar áreas de oportunidad para establecer señalamientos de mejora en el proceso, se logró desarrollar una metodología compuesta por 6 fases; se inició con la identificación de la situación actual de la institución por medio de un análisis descriptivo, posteriormente para obtener datos del proceso se diseñó una herramienta para encuestar a pacientes activas en tratamiento, esos datos fueron procesados para descubrir el modelo de proceso; se pasó a la interpretación de los datos arrojados y, finalmente se realizaron reportes a la institución de los resultados obtenidos para la toma de decisiones y las estrategias de prevención.

Es importante el seguimiento a las guías clínicas establecidas por instituciones, sin embargo, los resultados reflejan que en la práctica el proceso de atención de cáncer de mama es flexible ya que depende de las características de cada paciente para saber qué acción realizar.

La ventaja de utilizar la metodología de minería de procesos para cualquier proceso hospitalario que requiera una descripción formal del mismo, además ayudará a evaluarlo para mejorar su eficiencia y eficacia, así como aumentar la satisfacción de las necesidades y expectativas del paciente y, del personal de atención hospitalaria; pudiera extenderse por ejemplo a otros tipos de cáncer y tener un alcance mayor con un impacto más amplio en la población. Para realizar este cambio será necesario ajustar las variables de cada modelo, para el modelo basado en guías clínicas será necesario identificar las actividades del proceso a analizar y, para el diseño del modelo de proceso es necesario recolectar la información.

Para conocer la situación actual de la institución se realizó con la muestra proporcionada por la institución que representó consultas médicas (generales y especializadas) del proceso de atención de cáncer de mama., sin embargo, es importante que se utilice la muestra completa que incluya todos los pacientes de cáncer de mama.

El proyecto presentó varios retos debido a cambios internos en la institución, se incluyó un análisis cualitativo que no se tenía contemplado realizar, contando con la participación de 19 pacientes de cáncer de mama en tratamiento activo, que gracias a la apertura que se tuvo por parte de la institución, pacientes, así como del personal administrativo, médico y enfermería pudo realizarse. Fue interpretado en conjunto generando meta inferencias; los datos obtenidos de la entrevista requieren de ser validada por un especialista y contrarrestarla con las recomendaciones de la norma.

## **5.2. Recomendaciones**

Para realizar un análisis de minería de procesos, se recomienda utilizar los datos de los registros del sistema de información que se tenga en la institución, con ellos se tendrá un análisis completo que arrojen datos más precisos del proceso. Se recomienda a la institución almacenar cada paso que da el paciente por cada área (llamado trace en inglés), esto ayudaría para tener el camino completo de cada

paciente; incluir fechas de solicitud y fechas en las que acude a citas al médico general, al laboratorio, ginecólogo, oncólogo, tratamientos; entre más detalle se tenga del proceso más completa será la evaluación. Además de que con este tipo de análisis se pueden obtener indicadores utilizados para evaluar los procesos, por ejemplo: tiempo de espera, recíproco, tiempo de permanencia, que ayudan a gestionar el flujo de pacientes.

Si no se cuenta con datos completos en el sistema de la institución del proceso de atención de cáncer de mama, se recomienda un análisis cualitativo utilizando el cuestionario propuesto para obtener datos de las variables relevantes de los pacientes. Para la aplicación del cuestionario se recomienda especial cuidado con el manejo de los datos por ser datos sensibles, se debe contar con una carta consentimiento del paciente donde se protejan los datos personales, con un previo permiso de la institución. El documento de consentimiento debe estar avalado por una persona que este registrado bajo la Junta de Revisión Institucional (IRB por sus siglas en inglés), siendo el órgano administrativo establecido para proteger los derechos y el bienestar de las personas reclutadas para la investigación (el ejemplo de documento de consentimiento se encuentra en el anexo 7.2).

Estos comentarios surgen debido a que no fue posible tener acceso a la muestra total de pacientes de cáncer de mama de la institución y se recomienda utilizar la muestra completa para obtener resultados a mayor detalle.

Adicionalmente al descubrimiento de proceso se recomienda realizar la verificación de conformidad, que ayuda a comparar el proceso real contra las guías o normas clínicas establecidas por el sistema de salud o institución, en el que se aprecian las similitudes, diferencias entre ambos modelos.

### 5.3. Trabajos futuros

Como trabajos futuros se encuentran:

- Se propone extender esta metodología a otros tipos de cáncer y tener un alcance mayor con un impacto más amplio en la población. Se puede hacer en otro proceso para encontrar desviaciones e identificar a los derechohabientes con alto riesgo o casos prioritarios dando seguimiento a cada derechohabiente de acuerdo a su perfil y, promoviendo prácticas de autoexploración y chequeos periódicos para la detección precoz de la enfermedad en la población afiliada a la institución, utilizando la aplicación móvil de acceso a derechohabientes por medio del código QR de su nueva credencial, difundiendo información importante que ayude a sus suscriptores a tomar decisiones que mejoren su salud, elevando la efectividad de la estrategia de medicina preventiva, procurando educar, informar y concientizar a los derechohabientes.
- KNN es una primera aproximación para indicar si el paciente desarrollará o no la enfermedad. Con el fin de obtener indicadores estadísticos, quizá sería interesante valorar técnicas tradicionales de predicción estadística o más actuales como Deep Neuronal Networks.
- Utilizar algoritmos de aprendizaje supervisado como KNN que calculen la probabilidad de que una persona llegue a desarrollar cáncer de mama en un futuro.
- Realizar un análisis de la incidencia de casos para emitir pronósticos sobre el estado físico de una paciente de cáncer de mama en tratamiento y detectar a tiempo una crisis por una alteración en su fisiología hepática u otro aspecto que la ponga en riesgo.

## 6. REFERENCIAS

Allemani, C., Weir, H. K., Carreira, H., Harewood, R., Spika, D., Wang, X.-S., Bannon, F., Ahn, J.V., Johnson, C.J., Bonaventure, A., Marcos-Gragera, R., Stiller, C., Azevedo e Silva, G., Chen, W.Q., Ogunbiyi, O.J., Rachet, B., Soeberg, M.J., You, H., Matsuda, T., Bielska-Lasota, M., Storm, H., Tucker, T.C., Coleman, M.P., the CONCORD Working Group., 2015. Global surveillance of cancer survival 1995–2009: analysis of individual data for 25 676 887 patients from 279 population-based registries in 67 countries (CONCORD-2). *Lancet (London, England)*, 385(9972), 977–1010.

American Cancer Society, 2018. *Síndromes de cáncer en las familias*. [online]. Available from <<https://www.cancer.org/es/cancer/causas-del-cancer/genetica/sindromes-de-cancer-familiar.html>> [6 de junio 2018]

Anderson, B.O., Yip, C.H., Smith, R.A., Shyyan R., Sener, S.F., Eniu, A., Carlson RW. y Azavedo E, Harford J., 2008. Guideline implementation for breast healthcare in low-income and middle-income countries: overview of the Breast Health Global Initiative Global Summit 2007. *Cancer*, 113, 2221–43.

Ángeles-Llerenas, A., Torres-Mejía, G., Lazcano-Ponce, E., Uscanga-Sánchez, S., Mainero-Ratchelous, F., Hernández-Ávila, J. E., Hernández-Ávila, M., 2016. Effect of care-delivery delay on the survival of Mexican women with breast cancer TT - Efecto de la demora en la atención sobre la supervivencia de mujeres mexicanas con cáncer de mama. *Salud Pública de México*, 58(2), 237–250.

Benko, A., y Wilson, B., 2003. Online decision support gives plans an edge, *Managed Healthcare Executive*, 13, 5, p. 20, OmniFile Full Text Select (H.W. Wilson), EBSCOhost, viewed 8 January 2018.

Breast Cancer, 2015. *Estadios del cáncer de mama*. [online]. Available from <<http://www.breastcancer.org/es/sintomas/diagnostico/estadios>> [6 de junio 2018]

Bottrighi, A., Canensi, L., Leonardi, G., Montani, S., Terenziani, P., 2018. Interactive mining and retrieval from process traces. *Expert Systems with Applications*. Volume 110. p. 62-79. ISSN 0957-4174. <https://doi.org/10.1016/j.eswa.2018.05.041>.

Camacho-Centellas, S.S., 2016. Método heurístico para el diagnóstico de cáncer de mama basado en minería de datos. Tesis Maestría. Universidad Mayor de San Andres.

Chandra, S., Patel, M., y Chandra, S., 2017. Prediction Of Oral Cancer Recurrence Using Data Mining. *Journal Of Cancer Research & Therapeutics*, 13, p. S347, Academic Search Complete, EBSCOhost, viewed 3 May 2018.

Chiotti, O., Monserrat, S., (nd). Minería de Datos en Base de Datos de Servicios de Salud. *Estudiante Maestría en Ingeniería en Sistemas de Información UTN – FRSF Cidisi – UTN – FRSF, Ingar UTN- CONICET*

Coleman, MP., Quaresma, M., Berrino, F., Lutz, JM., De Angelis, R., Capocaccia, R., Bail, P., Rachet, B., Gatta, G., Hakulinen, T., Micheli, A., Sant, M., Weir, HK., Elwood, JM., Tsukuma, H., Koifman, S., E Silva, GA., Francisci, S., Santaquilani, M., Verdecchia, A., Storm, HH., Young, JL., 2008. Cancer survival in five continents: a worldwide population-based study (CONCORD). *CONCORD Working Group. Lancet Oncol.* 2008 Aug;9(8):730-56. doi: 10.1016/S1470-2045(08)70179-7.

de la Vara-Salazar, E., Suárez-López, L., Ángeles-Llerenas, A., Torres-Mejía, G. y Lazcano-Ponce, E., 2011. Tendencias de la mortalidad por cáncer de mama en México, 1980-2009. *Salud Pública de México*, 53(5), 385–393.

Fayyad, U., Piatestky-Shapiro, G. y Smyth, P., 1996. The kdd Process for Extracting Useful Knowledge from Volumes of Data. *Communications of the acm*, 39(11), 27-34

Flores-Luna, L., Salazar-Martinez, E., Duarte-Torres, RM., Torres-Mejía, G., Alonso-Ruiz, P., Lazcano-Ponce, E., 2008. Prognostic factors related to breast cancer survival. *Salud Publica Mex* 2008;50(2):119-125.

García, A.O., Armenteros, O. U. L., Ramírez Y. E. P. y Alfonso D.P., 2016. Inductive Visual Miner Plugin Customization for the Detection of Eventualities in the Processes of a Hospital Information System. *IEEE Latin America Transactions*, vol. 14, no. 4, pp. 1930-1936. doi: 10.1109/TLA.2016.7483536.

Gheorghe, M. y Petre, R. 2014. Integrating Data Mining Techniques into Telemedicine Systems. *Informatica Economica*, Vol.18, No.1, pp.120–130.

Gómez-Dantés, O., Sesma, S., Becerril, VM., Knaul, FM., Arreola, H., Frenk, J., 2011. Sistema de salud de México. *Salud Púb Méx*, 53, pp. s220-s232.

Guler, E. N., 2017. Gene Expression Profiling in Breast Cancer y Its Effect on Therapy Selection in Early-Stage Breast Cancer. *European Journal of Breast Health*, 13(4), 168–174.

Gutiérrez, E., Ramos, W., Uribe, M., Ortega-Loayza, A.G. Torres, C., Montesinos, D., León, O., Galarza, C., 2009. Tiempo de espera y su relación con la satisfacción de los usuarios de la farmacia central de un hospital general de Lima. *Rev. perú. med. exp. salud publicada*, vol.26, n.1, p.61-65.

Han, J., Kamber, M., y Pei, J., 2011. *Data Mining: Concepts and Techniques*. 3rd ed. *Morgan Kaufmann Publishers*.

Instituto Nacional de Estadística y Geografía, 2015. Conjunto de datos: Mortalidad general en México. Disponible en: <[http://www.inegi.org.mx/lib/olap/consulta/general\\_ver4/MDXQueryDatos.asp?#Regreso&c=11144](http://www.inegi.org.mx/lib/olap/consulta/general_ver4/MDXQueryDatos.asp?#Regreso&c=11144)> [29 de octubre 2017].

Instituto Nacional de Salud Pública, 2015. ¿Quiénes somos? [online]. Available from <<https://www.insp.mx/el-instituto/iQuienes-somos.html>> [5 July 2018]

Kreuze, D., 2001. Debugging Hospitals. *Technology Review*, 104, 2, p. 32, Business Source Complete, EBSCOhost, viewed 8 January 2018.

Koh, H.C., Tan G., 2005. Data mining applications in healthcare. *J Healthc Inf Manag*.19(2): 64–72.

Koh, H.C. y Leong, S.K., 2001. Data mining applications in the context of casemix. *Annals, Academy of Medicine*, 30(4, Supplement), 41-49.

López-Ramírez, J.A., 2017. Optimización del proceso de reclutamiento en una empresa de desarrollo de software con técnicas de minería de datos. Tesis Maestría. Universidad de Sonora.

Lorente, R. R., Hernández, M. C. R., Durán, D. H. y Lien, T. G. T., 2011. Experiencia en mujeres con cáncer de mama durante un trienio. *Revista Cubana de Obstetricia Y Ginecología*, 37(3), 349–358.

Mans, R. S., Schonenberg, M. H., Song, M., van der Aalst, W. M. P., & Bakker, P. J. M. (2008). Application of Process Mining in Healthcare – A Case Study in a Dutch Hospital. *Communications in Computer and Information Science*, 425–438. doi:10.1007/978-3-540-92219-3\_32

McLaughlin, J.M., Anderson, R.T., Ferketich, A.K., Seiber, E.E., Balkrishnan, R., Paskett, E.D., 2012. Effect on survival of longer intervals between confirmed diagnosis and treatment initiation among low-income women with breast cancer. *J Clin Oncol* 2012;30(36):4493-4500.

Morgado-García, T., Antonio Ponce-de-León-Lima, D. y Rosete-Suárez, A., 2017. Descubrimiento de conocimiento en bases de datos históricas de una empresa comercializadora. *Ingeniería Industrial*, 38(3), pp. 289–297. Available at: <http://search.ebscohost.com/login.aspx?direct=true&db=zbh&AN=125797674&lang=es&site=ehost-live> (Accessed: 1 December 2018)

NOM-041-SSA2-2011, 2011. Norma Oficial Mexicana NOM-041-SSA2-2011, Para la prevención, diagnóstico, tratamiento, control y vigilancia epidemiológica del cáncer de mama. *Secretaría de la Salud*. 9 de junio de 2011.

Oded M., y Lior R., 2005. *Data Mining and Knowledge Discovery Handbook*. Springer-Verlag New York, Inc., Secaucus, NJ, USA

Orellana-García, A., Armenteros, O.U.L., Ramírez Y.E.P. y Pérez-Alfonso, D., 2016. Inductive Visual Miner Plugin Customization for the Detection of Eventualities in the Processes of a Hospital Information System. *IEEE Latin America Transactions*, vol. 14, no. 4, pp. 1930-1936. doi: 10.1109/TLA.2016.7483536.

Orellana-García, A., Pérez-Alfonso, D. y Estrada-Sentí, V., 2016. Review of Models for Applying Process Mining Techniques. *GECONTEC: Revista Internacional de Gestión del Conocimiento y la Tecnología*. ISSN 2255-5684. Vol.4 (1).

Organización Mundial de la Salud, nd. *Cáncer de mama: prevención y control*. [online]. Available from <<http://www.who.int/topics/cancer/breastcancer/es/index3.html>> [30 de octubre 2017]

Ozkan, M., Yildirim, N., Disci, R., Ilgun, A. S., Sarsenov, D., Alco, G. y Ozmen, V., 2017. Roles of Biopsychosocial Factors in the Development of Breast Cancer. *European Journal of Breast Health*, 13(4), 206–212.

Peng, L., Tingting, B., Jiuling, H. y Siben, L., 2014. Breast cancer early diagnosis based on hybrid strategy. *Bio-Medical Materials & Engineering*, 24, 6, pp. 3397-3404, Academic Search Complete, EBSCOhost, viewed 26 December 2017.

Piñeros, M., Sánchez, R., Perry, F., García, O.A., Ocampo, R., Cendales, R. Demoras en el diagnóstico y tratamiento de mujeres con cáncer de mama en Bogotá, Colombia. *Salud Publica Mex* 2011; 53: 478-485.

Porta i Sales, J., 2017. Cuidados paliativos y derechos humanos. *Medicina Paliativa*, 24(4), 45–47. <https://doi.org/10.1016/j.medipa.2014.03.001>

Richards, MA., Westcombe, AM., Love, SB., Littlejohns, P., Ramirez, AJ., 1999. Influence of delay on survival in patients with breast cancer: a systematic review. *Lancet*1999;353(9159):1119-1126.

Rovani, M., Maggi, F.M., de Leoni, M., van der Aalst, Wil M.P., 2015. Declarative process mining in healthcare. *Expert Systems with Applications*, Volume 42, Issue 23, Pages 9236-9251, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2015.07.040>.

Rudnitckaia, J., 2015. Process Mining. Data science in action. *University of Technology, Faculty of Information Technology*: 1-11. Volume 42, Issue 23, Pages 9236-9251, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2015.07.040>.

Secretaría de Salud. 2013. Defunciones de mujeres de 25 años y más por tumor maligno de mama (C50) 2000-2013. Disponible en: <<https://www.gob.mx/cms/uploads/attachment/file/12950/MortalidadCaMa2000a2013.pdf>> [12 de noviembre 2017].

Silachan, K. y Tantasanawong, P. 2011. Evaluation Information Extraction for Health Text Categories Using C4.5 and Naïve Bayes. *Proceedings Of The European Conference On Information Management & Evaluation*, pp. 403-411, Business Source Complete, EBSCOhost, viewed 26 December 2017.

Singh, N.K., 2015. Prediction of Breast Cancer using Rule Based Classification. *Applied Medical Informatics*, 37, 4, pp. 11-22, Academic Search Complete, EBSCOhost, viewed 26 December 2017.

Stankov, A., Bargallo-Rocha, J.E., Silvio, A.Ñ., Ramirez, M.T., Stankova-Ninova, K., Meneses-Garcia, A., 2012. Prognostic factors and recurrence in breast cancer:

experience at the national cancer institute of Mexico. ISRN Oncol. 2012;2012:825258. doi:10.5402/2012/825258. Epub 2012 Jul 5. PubMed PMID: 22830047; PubMed Central PMCID: PMC3399427.

Thompson, W. y Roberson, M., 2000. Making Predictive Medicine Possible. (cover story). *R&D Magazine*, 42, 6, p. E4, Business Source Complete, EBSCOhost, viewed 10 January 2018.

Torres-Sakipova, D.Y., 2014. Algorithms and discovering techniques for unstructured processes: state of art. *Revista Cubana de Ciencias Informática*. ISSN: 2227-1899. Vol. 8, No. 3. pp. 114-130.

Uma, O. y Savita, G., 2017. A study on prediction of breast cancer recurrence using data mining techniques. 7th International Conference on Cloud Computing, *Data Science & Engineering - Confluence*, Noida, pp. 527-530. doi: 10.1109/CONFLUENCE.2017.7943207.

Van der Aalst, W.M., 2016. *Process mining: data science in action*. Springer.

Veletsos, A., 2003. Getting to the Bottom of Hospital Finances. *Health Management Technology*, 24, 8, pp. 30-31, OmniFile Full Text Select (H.W. Wilson), EBSCOhost, viewed 9 January 2018.

Yip, C.H., Anderson, B.O., Smith, R.A., Shyyan, R., Sener, S.F., Eniu, A., Carlson, R.W., Azavedo, E. y Harford, J., 2008. Guideline implementation for breast healthcare in low- and middle-income countries: early detection resource allocation. *Cancer*, 113, 2244–56.

Zaki, M.J. y Meira, J.W., 2014. *Contents, In Data Mining and Analysis: Fundamental Concepts and Algorithms*, toc, Cambridge, Cambridge University Press, pp. v-viii.

## 7. ANEXOS

Este apartado se encuentra el apoyo utilizado para la metodología e implementación.

### 7.1. Entrevistas

Para evaluar los procesos de atención a pacientes con cáncer de mama en ISSSTESON, se diseñó la siguiente encuesta que extrae información de los pacientes atendidos en la institución; consta de cinco secciones (general, estatus actual, información sobre la enfermedad, evaluaciones y vivir con la enfermedad).

En la figura 7.1 se muestra la bienvenida a participar en el estudio de investigación, se describe el propósito de la encuesta, se le hace saber al paciente que le tomará aproximadamente de 30-45 min y da consentimiento informado sobre su participación.

**Título del estudio: Minería de datos y análisis de georeferenciación para dirigir esfuerzos de medicina preventiva y cuidado del cáncer de mama en Sonora**

**Estimado(a) paciente:**

*Esta es una invitación a participar en un estudio de investigación sobre el proceso de atención de cáncer de mama en el Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado de Sonora (ISSSTESON).*

*El propósito del estudio es evaluar los procesos de atención a pacientes con cáncer de mama en ISSSTESON y el Hospital Oncológico para determinar mejoras que acorten los tiempos entre citas médicas y servicios, así como eliminar o agilizar trámites.*

*Le tomará unos 30-45 minutos, su aportación es anónima y ayudará a mejorar los servicios de salud.*

*Responder a las preguntas también significa dar su consentimiento informado sobre la participación en este estudio.*

**Gracias por su amable cooperación.**

*Si tiene alguna duda o comentario, puede escribir a los siguientes correos electrónicos:*

**Investigador principal:** Dra. Raquel Torres Peralta ([raquel.torresp@t@gmail.com](mailto:raquel.torresp@t@gmail.com))

**Estudiante de investigación:** Ing. Angélica Enríquez Amaya ([angelica.enriquez.88@gmail.com](mailto:angelica.enriquez.88@gmail.com))

**Registro en Universidad de Sonora: USO 316003073**

*Hay 20 preguntas en la encuesta.*

**Siguiente**

*Figura 7.1. Bienvenida a la encuesta.*

General

En esta sección se le pregunta al paciente sobre su proceso de atención, cuestionando fechas sobre las consultas realizadas, ya sea al médico general, oncólogo, fecha de mamografía, biopsia, operación, de tratamientos, como se muestra en la figura 7.2.

The screenshot shows a form titled "General" with the following sections:

- 1. ¿Cuándo llegó por última vez (dentado) a algún tipo de institución de salud?** (Empty text input)
- 2. ¿Fecha(s) de la primera consulta?** (Date input)
- 3. ¿Fecha(s) de mamografía (sin imagen) después de la primera consulta?** (Date input)
- 4. ¿Fecha(s) de inicio de radioterapia?** (Date input)
- 5. ¿Fecha(s) de cirugía?** (Date input)
- 6. ¿Fecha(s) de la última cirugía?** (Date input)
- 7. ¿Fecha(s) de inicio y/o tratamiento?** (Date input)
- 8. ¿Fecha(s) de inicio con algún tipo de agente anti-cancerígeno?** (Date input)
- 9. ¿Fecha(s) de inicio de cirugía y/o radioterapia?** (Date input)
- 10. ¿Fecha(s) de inicio de otros tipos de tratamientos?** (Date input)
- 11. Síntomas**
  - 11.1. ¿Qué síntomas o molestias experimenta?**
    - Dolor(es) (¿Dónde? ¿Cómo? ¿Cuánto?)
    - No sé
    - Pérdida de peso involuntaria
    - Anemia (¿cómo se sintió?)
    - Otros (¿cómo se sintió?)
- 12. ¿Qué síntomas o molestias experimenta (después de la cirugía o radioterapia)?** (Buttons: Sí, No, No responde)
- 13. ¿Qué síntomas o molestias experimenta (después de los tratamientos)?** (Buttons: Sí, No, No responde)

Figura 7.2. Sección general.

## Estatus general

En la sección de estatus general se le pregunta al paciente datos actuales e historial de estancia como se muestra en la figura 7.3.

**Estatus Actual**

14. Cabeza:

15. Ciudad de residencia:

16. ¿Cuánto tiempo tienen viviendo ahí?  
Comenzar sólo si escoge una respuesta.

Años:

Meses:

17. Lugar de nacimiento:

18. Localidad/colonias dónde creció, especifique edades (ejemplo: de los 0-20 años en colonia Jardines en Hermosillo, de los 21-51 en colonia Castro en Hermosillo):

19. En su trabajo o en su residencia, ¿se encuentra como de taller de carpintería o tiene contacto con fertilizantes, pesticidas, artículos de limpieza, etc.?

Sí  No  Sin respuesta

**Siguiente**

**Figura 7.3.** Sección estatus general.

## Información sobre la enfermedad

En esta sección se le pide al paciente que plasme la información que la institución le brindó sobre la enfermedad como se muestra en la figura 7.4.

### Información sobre la enfermedad

**20. ¿Ha recibido información sobre la enfermedad por parte del Instituto u otros especialistas?**

Sí
  No
  Sin respuesta

---

**21. Ha recibido información sobre:**

Comentar sólo si escoge una respuesta:

Apoyo emocional:   
 Dieta:   
 Recursos externos al Instituto:   
 Otro:

---

**22. ¿Ha buscado información por su parte?**

Sí
  No
  Sin respuesta

---

**23. ¿Tiene alguna duda sobre la enfermedad y no ha encontrado información?**

Sí
  No
  Sin respuesta

---

**24. ¿Cuáles son sus principales necesidades?**

Las necesidades pueden ser por ejemplo: información sobre la enfermedad, tratamientos, ayuda psicológica, nutricional, etc.

**Figura 7.4.** Sección Información sobre la enfermedad.

Evaluaciones

En esta sección se evalúa la labor y actitud del personal de enfermería (figura 7.5), evaluación de la labor y actitud del personal médico, evaluación de las condiciones ambientales y sociales de hospitalización y evaluación de la mejoría percibida (figura 7.6), evaluación en general del servicio brindado (figura 7.7).

**Evaluaciones**

25. Evalúe del 1 al 10 al Personal de Enfermería por el servicio brindado.

	1	2	3	4	5	6	7	8	9	10	Sin respuesta
1. ¿Hicieron su trabajo en tiempo adecuado?	<input type="radio"/>	<input checked="" type="radio"/>									
2. ¿Asidieron a su llamado cuando los necesitó?	<input type="radio"/>	<input checked="" type="radio"/>									
3. ¿Lo escucharon lo suficiente?	<input type="radio"/>	<input checked="" type="radio"/>									
4. ¿Hicieron bien su trabajo?	<input type="radio"/>	<input checked="" type="radio"/>									
5. En general, ¿diría que el personal de enfermería le ha demostrado simpatía?	<input type="radio"/>	<input checked="" type="radio"/>									
6. ¿Lo estimulaban para preguntar lo que no entendía?	<input type="radio"/>	<input checked="" type="radio"/>									
7. En general, ¿diría que el personal de enfermería le ha demostrado interés?	<input type="radio"/>	<input checked="" type="radio"/>									
8. Antes de realizar cualquier procedimiento o aplicar algún medicamento, ¿se le explicaron?	<input type="radio"/>	<input checked="" type="radio"/>									
9. En general, ¿diría que el personal de enfermería le ha demostrado respeto?	<input type="radio"/>	<input checked="" type="radio"/>									
10. ¿Fueron atentos las personas encargadas de su ingreso?	<input type="radio"/>	<input checked="" type="radio"/>									
11. ¿Consideró adecuado los muebles del cuarto?	<input type="radio"/>	<input checked="" type="radio"/>									

Figura 7.5. Evaluación personal enfermería.

26. Evalúe del 1 al 10 al Personal de Hospitalización y Médicos por el servicio brindado:

	1	2	3	4	5	6	7	8	9	10	Sin respuesta
1. En general, ¿diría que el personal médico le ha demostrado simpatía?	<input type="radio"/>	<input checked="" type="radio"/>									
2. En general, ¿diría que el personal médico le ha demostrado interés?	<input type="radio"/>	<input checked="" type="radio"/>									
3. En general, ¿diría que el personal médico le ha demostrado respeto?	<input type="radio"/>	<input checked="" type="radio"/>									
4. Los médicos, ¿hicieron su trabajo en tiempo adecuado?	<input type="radio"/>	<input checked="" type="radio"/>									
5. Los médicos, ¿acudieron a su llamado cuando los necesitó?	<input type="radio"/>	<input checked="" type="radio"/>									
6. ¿Lo escucharon lo suficiente?	<input type="radio"/>	<input checked="" type="radio"/>									
7. Durante la hospitalización, ¿se sintió rodeado de personas amistosas?	<input type="radio"/>	<input checked="" type="radio"/>									
8. ¿Lo estimularon para preguntar lo que no entendía?	<input type="radio"/>	<input checked="" type="radio"/>									
9. Durante la hospitalización, ¿se sintió solo?	<input type="radio"/>	<input checked="" type="radio"/>									
10. ¿Le molestó el ruido durante la noche?	<input type="radio"/>	<input checked="" type="radio"/>									
11. ¿El ambiente fue de calma y tranquilidad?	<input type="radio"/>	<input checked="" type="radio"/>									
12. ¿Le molestó el ruido durante el día?	<input type="radio"/>	<input checked="" type="radio"/>									
13. ¿Fue cómoda su cama?	<input type="radio"/>	<input checked="" type="radio"/>									
14. ¿Fue rígida la atención del personal de hospitalización encargadas de su ingreso?	<input type="radio"/>	<input checked="" type="radio"/>									
15. ¿Considera que la duración de su hospitalización fue adecuada?	<input type="radio"/>	<input checked="" type="radio"/>									
16. Durante la hospitalización, ¿se sintió ansioso o temeroso?	<input type="radio"/>	<input checked="" type="radio"/>									
17. ¿Le explicaron ampliamente la enfermedad?	<input type="radio"/>	<input checked="" type="radio"/>									
18. ¿Entendió esta información?	<input type="radio"/>	<input checked="" type="radio"/>									
19. ¿Lo mantuvieron siempre al tanto de su progreso?	<input type="radio"/>	<input checked="" type="radio"/>									
20. Antes de realizar cualquier procedimiento o aplicarle algún tratamiento, ¿se lo explicaron?	<input type="radio"/>	<input checked="" type="radio"/>									
21. ¿Le dijeron por qué estaba ingresando?	<input type="radio"/>	<input checked="" type="radio"/>									
22. ¿Conoce mejor su enfermedad y cómo manejarla?	<input type="radio"/>	<input checked="" type="radio"/>									
23. ¿Cómo fue el trato demostrado por el personal de las camillas?	<input type="radio"/>	<input checked="" type="radio"/>									
24. ¿Le explicaron rutinas y procedimientos del servicio?	<input type="radio"/>	<input checked="" type="radio"/>									
25. ¿Le dijeron qué artículos de uso personal necesitaría traer antes y cuáles le proporcionaría el hospital?	<input type="radio"/>	<input checked="" type="radio"/>									
26. ¿Le informaron el tiempo aproximado de hospitalización?	<input type="radio"/>	<input checked="" type="radio"/>									
27. ¿Fue adecuada la iluminación del cuarto durante la noche?	<input type="radio"/>	<input checked="" type="radio"/>									

Figura 7.6. Evaluar personal médico y hospitalización.

27. Evalúa del 1 al 10 el Servicio en General que se le brindó:

	1	2	3	4	5	6	7	8	9	10	Sin respuesta
1. Cuando recibió el diagnóstico, ¿considera que fue el modo adecuado de haberlo sabido?	<input type="radio"/>	<input checked="" type="radio"/>									
2. ¿Cómo es el trato que recibe del personal de quimioterapia?	<input type="radio"/>	<input checked="" type="radio"/>									
3. ¿Cómo es el trato que recibe del personal de laboratorio?	<input type="radio"/>	<input checked="" type="radio"/>									
4. ¿Cómo es el trato que recibe del personal administrativo en general?	<input type="radio"/>	<input checked="" type="radio"/>									

28. ¿Algún comentario o sugerencia?

Figura 7.7. Evaluar servicio en general.

### Vivir con la enfermedad

En la última sección se le cuestiona al paciente que le ayuda a sentirse mejor físicamente y emocionalmente como se ve en la figura 7.8.

Vivir con la enfermedad

29. ¿Qué le ayuda a sentirse mejor físicamente?

30. ¿Qué le ayuda a sentirse mejor emocionalmente?

Figura 7.8. Sección vivir con la enfermedad.

## 7.2. Documento de consentimiento

A continuación, se presenta el documento de consentimiento informado sobre la participación en este estudio donde el paciente garantiza que ha expresado voluntariamente su intención de participar en el estudio, además su aportación es anónima y ayudará a mejorar los servicios de salud.



### UNIVERSIDAD DE SONORA DOCUMENTO DE CONSENTIMIENTO INFORMADO PARA PROYECTO DE INVESTIGACIÓN

**Título del estudio:** Minería de datos y análisis de georeferenciación para dirigir esfuerzos de medicina preventiva y cuidado del cáncer de mama en Sonora.

**Investigador principal:** Dra. Raquel Torres Peralta

**Estudiante de investigación:** Ing. Angélica Enríquez Amaya

**Registro en Universidad de Sonora: USO 316003073**

Esta es una invitación a participar en un estudio de investigación sobre el proceso de atención de cáncer de mama en el Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado de Sonora (ISSSTESON).

El propósito del estudio es evaluar los procesos de atención a pacientes con cáncer de mama en ISSSTESON y el Hospital Oncológico para determinar mejoras que acorten los tiempos entre citas médicas y servicios, así como eliminar o agilizar trámites.

#### ¿Qué haré si elijo estar en este estudio?

- Contestar una encuesta en línea donde se le preguntará sobre los tiempos de atención, información que se le haya proporcionado sobre la enfermedad y la calidez humana por parte del personal y requerimientos de su parte o necesidades que considere se deberían cubrir.

**Ubicación del estudio:** todos los procedimientos del estudio se llevarán a cabo en el ISSSTESON-Centro Médico Dr. Ignacio Chávez ubicado en Aguascalientes S/N, Centro, 83000 Hermosillo, Sonora, aunque se puede realizar en diferentes ubicaciones la que se pueda convenir con el participante.

La entrevista podría ser grabada (sólo audio) para asegurar la fidelidad de la información, pero si no se siente cómoda podemos limitarnos a tomar nota de sus respuestas. Los audios serán utilizados sólo por Angélica Enríquez y Raquel Torres. En ningún momento habrá una identificación de los archivos con las participantes. Se garantiza el anonimato de las participantes durante todo el proceso y después de terminada la investigación.

En caso de citar sus comentarios en presentaciones o artículos resultantes de este trabajo se utilizará un seudónimo para proteger su identidad, **a menos que solicite específicamente que lo identifiquen con su nombre verdadero.**

### **¿Cuáles son los posibles riesgos o incomodidades?**

Su participación en este estudio no implica ningún riesgo físico o emocional para usted más allá de la vida cotidiana.

En todo momento podrá detener la entrevista. En caso de sentirse emocional o incómoda al responder algunas de las preguntas puede simplemente no contestarla y pasar a la siguiente.

Para minimizar los riesgos de confidencialidad, no compartiremos la identidad de aquellos que participen en la investigación. La información que recojamos por este proyecto de investigación se mantendrá confidencial. La información acerca de usted que se recogerá durante la investigación será puesta fuera de alcance y nadie sino los investigadores tendrán acceso a verla. Cualquier información acerca de usted tendrá un número en vez de su nombre. Solo los investigadores sabrán cuál es su número y se mantendrá con acceso limitado a los registros del estudio.

Podemos compartir los datos que recopilamos de usted para su uso en futuros estudios de investigación o con otros investigadores; si compartimos los datos que recopilamos sobre usted, eliminaremos cualquier información que pueda identificarlo antes de compartirla.

### **Información financiera**

La participación en este estudio no implicará ningún costo para usted. No se le pagará por participar en este estudio. Sin embargo, tendremos un detalle de agradecimiento al finalizar la entrevista.

### **¿Con quién me puedo comunicar si tengo preguntas o inquietudes sobre este estudio de investigación?**

Si tiene preguntas, puede hacerlas ahora. Si tiene preguntas más adelante, puede comunicarse con los investigadores a:

Nombre: Angélica Enríquez Amaya

Cel.: 6623-61-98-27

Email: [angelica.enriquez.88@gmail.com](mailto:angelica.enriquez.88@gmail.com)

**UNIVERSIDAD DE SONORA**  
**DOCUMENTO DE CONSENTIMIENTO INFORMADO PARA PROYECTO DE**  
**INVESTIGACIÓN**

**Título del estudio:** Minería de datos y análisis de georeferenciación para dirigir esfuerzos de medicina preventiva y cuidado del cáncer de mama en Sonora

**Investigador principal:** Dra. Raquel Torres Peralta

**Estudiante de investigación:** Ing. Angélica Enríquez Amaya

**Registro en Universidad de Sonora: USO 316003073**

**Consentimiento**

He leído este formulario y el estudio de investigación me ha sido explicado. Se me ha dado la oportunidad de hacer preguntas y mis preguntas han sido respondidas. Me han proporcionado un contacto por si tengo preguntas adicionales. Estoy de acuerdo en participar en el estudio de investigación descrito anteriormente y recibiré una copia de este formulario de consentimiento.

**Elementos opcionales de estudio**

**Consentimiento para el uso de la información de contacto para ser contactado sobre la participación en otros estudios**

Marque uno de los siguientes para indicar su elección:

\_\_\_\_\_ (Marque S/N) Estoy de acuerdo en permitir que los investigadores utilicen mi información de contacto recopilada durante este estudio para contactarme sobre la participación en futuros estudios de investigación.

\_\_\_\_\_ (Marque S/N) Autorizo a los investigadores de este estudio a utilizar mis respuestas como material para publicaciones, conferencias y divulgación en general guardando mi anonimato en todo momento.

\_\_\_\_\_  
Nombre del participante

\_\_\_\_\_  
Firma del participante

\_\_\_\_\_  
Fecha

### 7.3. Resultados de la encuesta

En esta sección se presenta el análisis de las entrevistas a pacientes de cáncer de mama ISSSTESON, se realizaron 19 entrevistas que contiene 30 reactivos.

#### Resumen de campo para G14

12. a) Si hubo algo que retrasara su tratamiento que no tenga que ver su estado de salud (ejemplo: falta de medicamento, trámites administrativos), especifíquelos y con cuánta frecuencia se presentó:

Opción	Cuenta	Porcentaje
Opción	1	100.00%
Sin respuesta	0	0.00%

ID	Respuesta
67	porque no hay citas esta ferno aquí

20. ¿Ha recibido información sobre la enfermedad por parte del instituto u otros medios?

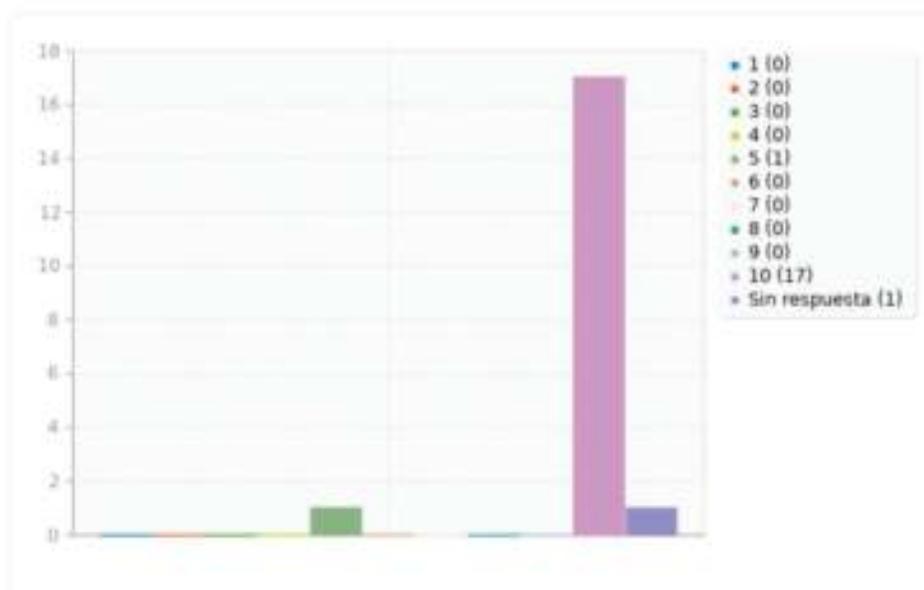
Opción	Cuenta	Porcentaje
Si (Y)	14	73.68%
No (N)	5	26.32%
Sin respuesta	0	0.00%

## 24. ¿Cuáles son sus principales necesidades?

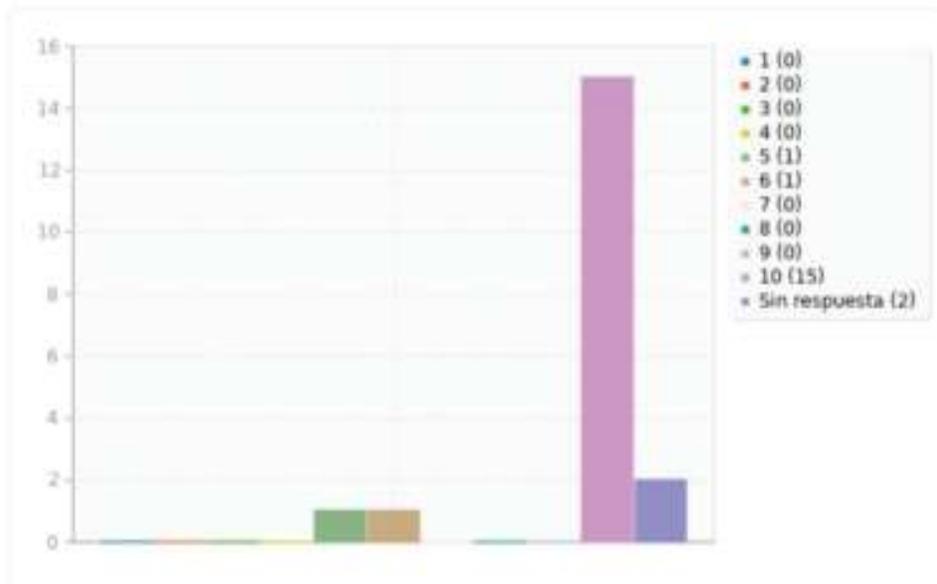
Opción	Cuenta	Porcentaje
Opción	19	100.00%
Sin respuesta	0	0.00%

ID	Respuesta
60	tanatólogo, saber si la comida en sonora influye en el cáncer de mama
61	medicamentos que no las encuentra; aunque tienes antecedentes de familiares con Ca de mama no te pasan para que te realicen la mamografía
62	Estudios se los tuvo que hacer por fuera momografía, ultrasonido, estudios de inmunohistoquímica
63	nada
64	se soló tiene 2 hermanos médicos que le ayudan
65	no hasta el momento
66	medicamento capecitabina
67	saber; le dicen que no tiene nada por qué aun le siguen poniendo quimios
68	medicamentos no ha fallado, siempre ha encontrado
69	nada
70	ninguna hasta el momento
71	ninguna hasta el momento
72	no por el momento
73	algunas medicinas tardan
74	la información sobre la alimentación, desconoce y tmb que ejercicios hacer (no ha preguntado)
75	sugerencias: ultrasonido inmediatamente la biopsia, no mamografía ya que solo lastima, vueltas para conseguir el medicamento
76	tuvieron que cambiar medicamento por el desabasto a nivel nacional
77	ninguna
78	en cuestión de estudios detección más profundas no tan básicas

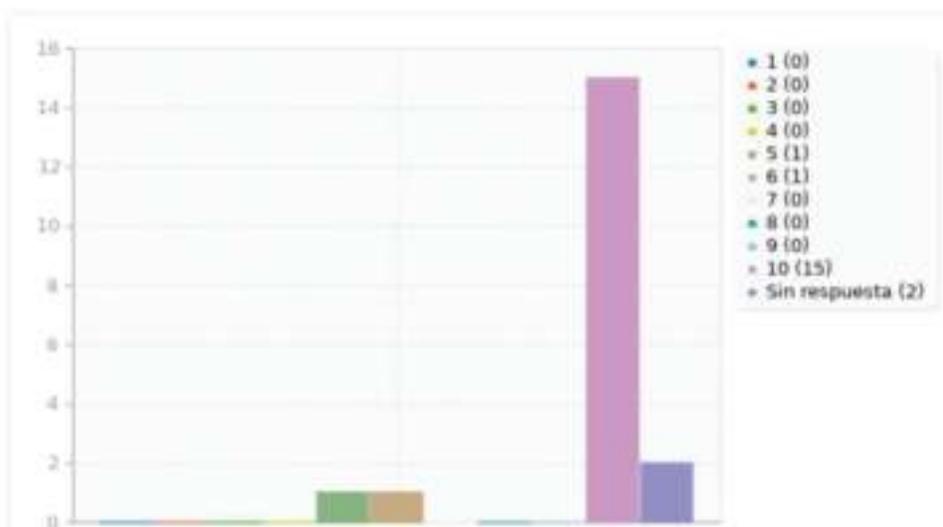
26. Evalúe del 1 al 10 al Personal de Hospitalización y Médicos por el servicio brindado: [1. En general, ¿diría que el personal médico le ha demostrado simpatía?]



26. Evalúe del 1 al 10 al Personal de Hospitalización y Médicos por el servicio brindado: [8. ¿Lo estimularon para preguntar lo que no entendía?]

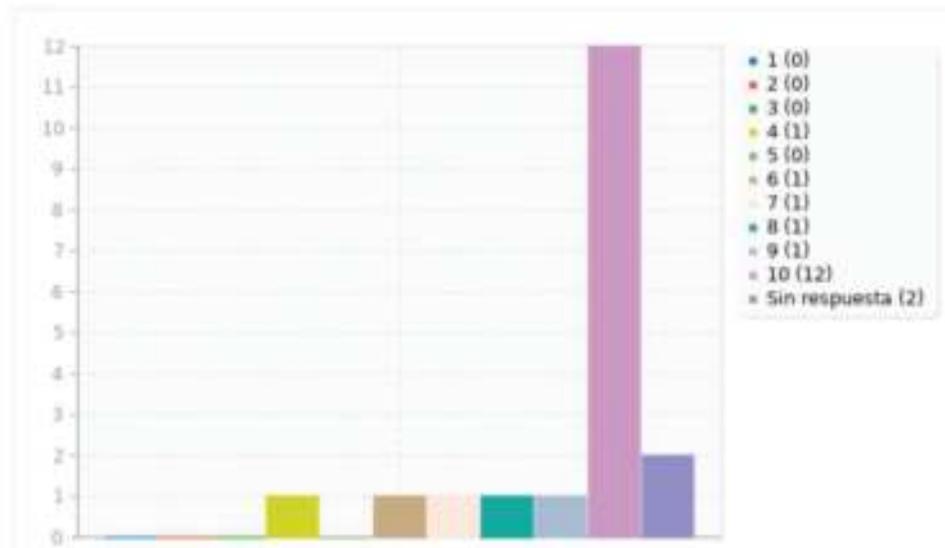


26. Evalúe del 1 al 10 al Personal de Hospitalización y Médicos por el servicio brindado: [10. ¿Le molestó el ruido durante la noche?]



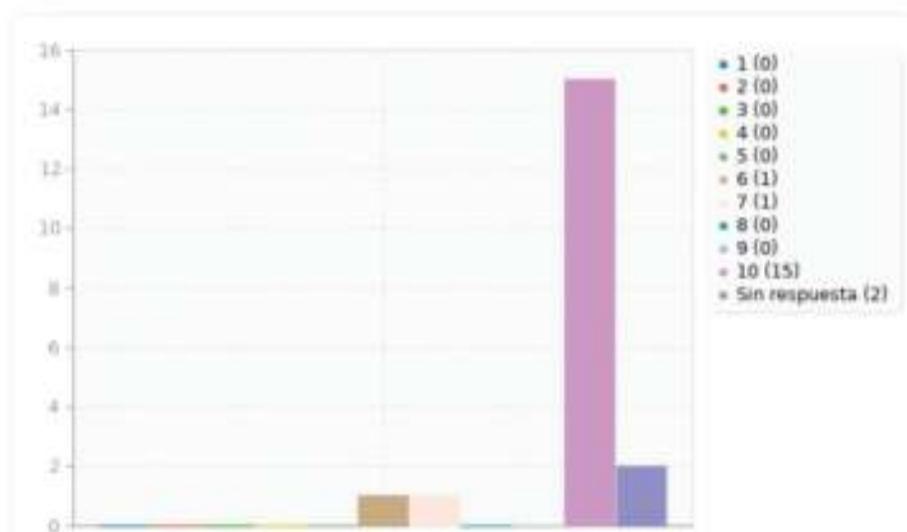
26. Evalúe del 1 al 10 al Personal de Hospitalización y Médicos por el servicio brindado: [13. ¿Fue cómoda su cama?]

---

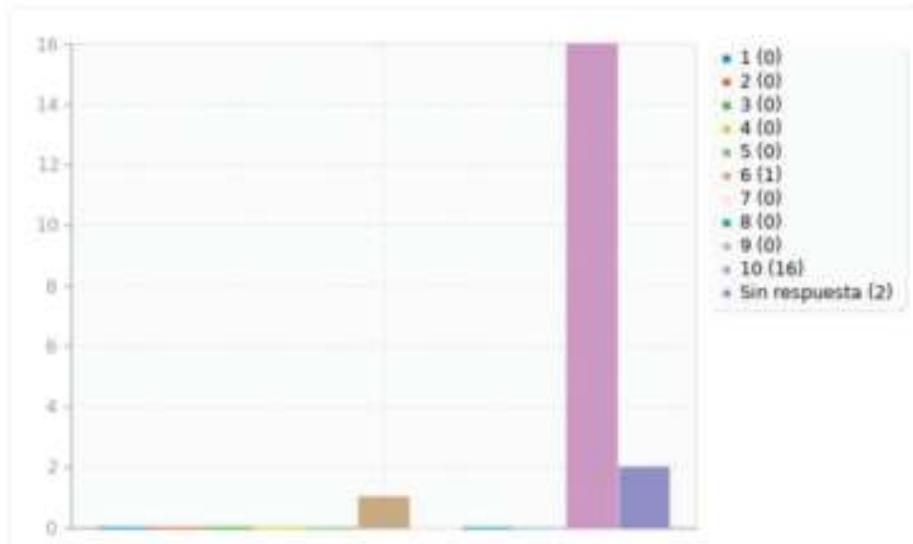


26. Evalúe del 1 al 10 al Personal de Hospitalización y Médicos por el servicio brindado: [18. ¿Entendió esta información?]

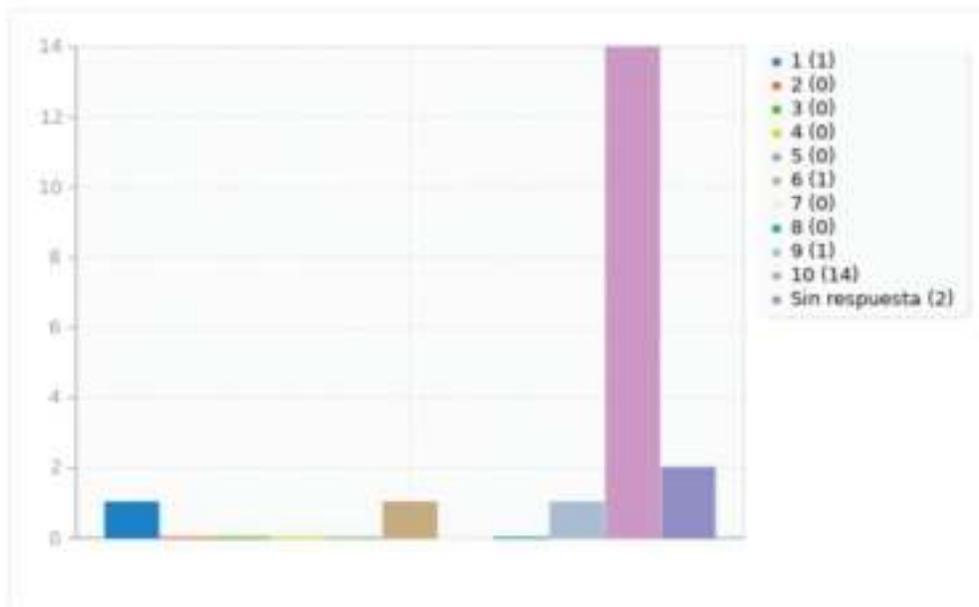
---



26. Evalúe del 1 al 10 al Personal de Hospitalización y Médicos por el servicio brindado: [21. ¿Le dijeron por qué estaba ingresando?]



26. Evalúe del 1 al 10 al Personal de Hospitalización y Médicos por el servicio brindado: [24. ¿Le explicaron rutinas y procedimientos del servicio?]



## 28. ¿Algún comentario o sugerencia?

Opción	Cuenta	Porcentaje
Opción	19	100.00%
Sin respuesta	0	0.00%

ID	Respuesta
60	Como paciente vas cansada te piden sellos firmas vigencias para poder autorizarte un medicamento, la doctora no subió cuando la operaron, sensibilidad tiene más los enfermeros que los doctores; no hay higiene, tapado las duchas, fatales los baños de afuera; comida mala pero el trato esta bien
61	si hay antecedentes no esperar a los 40 para realizar estudios más afondo, pq no lo hacen, medicamento que no lo encuentra vacuna para subir las defensas antes de la quimio, vueltas administrativas
62	revalorar del personal y queden los que llenen los perfiles que las profesiones exigen
63	Estudios inmunohistoquímica, ecocardiograma tuvo que pagarlos ella, tuvo que comprar medicamento, batalló con las citas, vueltas administrativas
64	aquí en hermosillo hay mas gente capacitada, que en caborca, que les paguen bien a los del personal ;batallaron con el aire en el aislado, tuvo 2 oncólogos
65	la consienten mucho, todo muy bien
66	no vueltas con medicamentos, son muy meticulosos
67	medicamento escaso, le dan receta y no hay
68	medicamentos no ha fallado, siempre ha encontrado;sala de quimio es un espacio pequeño;cama 6 fallaba aire;mejorar todo para bien de nosotros, que la gobernadora cumpla;
69	;5 de calificación a la sala por pequeña;
70	ninguna hasta el momento;; quimio muy bien tiene muy buen estilo;
71	ninguna hasta el momento;; Gracias por todo el trato;
72	no por el momento;; Muy bien;
73	algunas medicinas tardan;escasos de sábanas y pijamas; gracias ellos puede dar esta entrevista;
74	la información sobre la alimentación, desconoce y tmb que ejercicios hacer (no ha preguntado);sala pequeña; me ha ido muy bien;
75	sugerencias: ultrasonido inmediatamente la biopsia, no mamografía ya que solo lastima, vueltas para conseguir el medicamento;; que en el isssteson haya un lugar especial para las pacientes de CA, consultas tardadas, en ocasiones se compran medicamentos y no están en tiempo; atención a cualquier duda;
76	tuvieron que cambiar medicamento por el desabasto a nivel nacional;pequeña la sala, ampliarla;; tener la precaución del medicamento en tiempo, uno no puede esperar hasta que el medicamento llegue;
77	ninguna;; no citas, atención en farmacia como tratas te tratan ;
78	en cuestión de estudios detección más profundas no tan básicas;la sala ha mejorado con el paso del tiempo;;la enfermera se asustó al momento del diagnóstico; ; la atención ha mejorado de 1 pésimo cuando inició, pero ahora que tiene médico de cabecera oncólogo un 5 ( en una cali de 1 al 5) ; invita a ver la realidad ver lo mejor de lo que se tiene;