

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL



**FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS
DEPARTAMENTO DE POSGRADO**

PROYECTO DE TITULACIÓN

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

“MAGÍSTER EN ESTADÍSTICA APLICADA”

TEMA:

CARACTERIZACIÓN EPIDEMIOLÓGICA DEL CÁNCER DE
ESTÓMAGO A TRAVÉS DE UN ANÁLISIS MULTIVARIANTE

AUTOR:

MARIUXI DEL ROCÍO QUIMIS MARCILLO

Guayaquil - Ecuador

2024

RESUMEN

El cáncer de estómago es considerado un problema de salud pública mundial, por lo que, Ecuador no es ajeno a este problema, ya que el cáncer gástrico es la primera causa de mortalidad en el país. Por lo tanto, este trabajo propone determinar las características epidemiológicas del cáncer de estómago en pacientes del hospital de la Sociedad de Lucha contra el Cáncer en el Ecuador - SOLCA Guayaquil, entre los años 2014 - 2023, mediante los métodos multivariantes HJ-Biplot y Análisis Triádico Parcial, de los cuales se obtuvieron insights como que el antro es el lugar topográfico donde más se encuentra el tumor maligno, las edades con mayor incidencia está en mayormente en los hombres de entre 50 a 79 años y en las mismas edades afecta a las mujeres en menor cantidad. Los años con mayor incidencia también presentaron que gran parte de esos pacientes a fecha de corte han fallecido.

Palabras Claves: Cáncer de estómago, HJ-Biplot, análisis multivariante, Análisis Triádico Parcial, Incidencia

ABSTRACT

Stomach cancer is considered a global public health issue, and Ecuador is no exception, as gastric cancer is the leading cause of mortality in the country. Therefore, this work aims to determine the epidemiological characteristics of stomach cancer in patients at the Sociedad de Lucha contra el Cáncer hospital in Ecuador - SOLCA Guayaquil, between the years 2014 and 2023, using the multivariate methods HJ-Biplot and Partial Triadic Analysis. Insights obtained from these methods include the finding that the antrum is the most common topographic location for the malignant tumor. The ages with the highest incidence are predominantly men between 50 and 79 years old, and it affects women in the same age range to a lesser extent. The years with the highest incidence also showed that a significant portion of these patients had passed away by the cut-off date.

Keywords: Stomach cancer, HJ-Biplot, multivariate analysis, Partial Triadic Analysis, incidence

DEDICATORIA

A Dios, por ser mi guía y fortaleza en cada momento, iluminando mi camino y brindándome la sabiduría necesaria para alcanzar este logro.

A mis padres y hermanas, por su amor incondicional, compañía, palabras de aliento y apoyo constante. Gracias por enseñarme el valor del esfuerzo, la perseverancia y la educación.

AGRADECIMIENTO

A mis mentores y profesores, por su guía, enseñanzas y por inculcarme la pasión por el conocimiento. Gracias por ser ejemplos a seguir y por su dedicación a la formación académica.

Finalmente, dedico este trabajo a todos aquellos que, de una u otra forma, contribuyeron a que esta meta se convirtiera en realidad. A todos ustedes, mi más profundo agradecimiento.

DECLARACIÓN EXPRESA

La responsabilidad por los hechos y doctrinas expuestas en este Proyecto de Titulación, me corresponde exclusivamente y ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría. El patrimonio intelectual del mismo, corresponde exclusivamente a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.



Mariuxi del Rocío Quimis Marcillo

TRIBUNAL DE GRADUACIÓN



PAÚL HERRERA SAMANIEGO, Ph.D.
PRESIDENTE



LUIS PILACUAN BONETE, Ph.D.
TUTOR



OMAR RUIZ BARZOLA, Ph.D.
DOCENTE EVALUADOR

ABREVIATURAS O SIGLAS

ACP	Análisis de Componentes Principales
ADK	Adenocarcinoma
ADSQ	Adenoescamoso
CA	Carcinoma
CIE-10	Clasificación estadística internacional de enfermedades y problemas relacionados con la salud
CIE-0	Clasificación Internacional de Enfermedades para Oncología
CRFE	Contribuciones relativas del factor al elemento
EPID	Epidermoide
GLOBOCAN	Global Cancer Observatory
IND	Indiferenciado
LINF	Linfoma
NED	Neuroendocrino
NIH	Instituto Nacional del Cáncer
PTA	Análisis Triádico Parcial
OMS	Organización Mundial de la Salud
SAR	Sarcoma
SOLCA	Sociedad de Lucha Contra el Cáncer

TABLA DE CONTENIDO

CAPÍTULO 1	1
1. INTRODUCCIÓN	1
1.1. Antecedentes	1
1.2. Descripción del Problema	1
1.3. Objetivos	2
1.3.1. Objetivo General	2
1.3.2. Objetivos Específicos	2
1.4. Hipótesis	3
1.5. Alcance	3
CAPÍTULO 2	4
2. MARCO TEÓRICO	4
2.1. Cáncer	4
2.1.1. Estómago	5
2.1.2. Desarrollo del Cáncer de Estómago	6
2.1.3. Morfología del Cáncer de Estómago	6
2.1.4. Codificación Morfológica y Topográfica Según CIE-O	9
2.2. Análisis Multivariante	10
2.2.1. Análisis Multivariante de Matrices de Dos Dimensiones	10
2.2.1.1. Análisis de Componentes Principales	11
2.2.1.2. Biplots	11
2.2.1.3. HJ-Biplot	12
2.2.2. Análisis Multivariante de Matrices de Tres – Vías	12
2.2.2.1. Análisis Triádico Parcial o X-STATIS	13
2.2.3. Aplicaciones de los Métodos Multivariantes	14
CAPÍTULO 3	16
3. METODOLOGÍA	16
3.1. Base de Datos	16
3.2. Diseño	16
3.3. Procesamiento de Base	16
3.4. Metodología Estadística	17
3.4.1. Método HJ-Biplot	18
3.4.2. Método Análisis Triádico Parcial o XSTATIS	19
3.5. Tecnologías Necesarias	22
CAPÍTULO 4	23
4. RESULTADOS	23

4.1.	Análisis Multivariante	25
4.1.1.	HJ-Biplot	25
4.1.1.1.	Topografía del Cáncer de Estómago	25
4.1.1.2.	Estado Vital Según el Año de Incidencia	27
4.1.2.	Análisis Triádico Parcial	29
CAPÍTULO 5		38
5.	CONCLUSIONES Y RECOMENDACIONES	38
5.1.	Conclusiones.....	38
5.2.	Recomendaciones	39
6.	Referencias	1
7.	Apéndices y anexos.....	5
7.1.	Apéndice A	5
7.2.	Apéndice B	6
7.3.	Apéndice C.....	7
7.4.	Apéndice D.....	8
7.5.	Apéndice E	9

LISTADO DE FIGURAS

Figura 2.1 Comportamiento normal y anormal de las células	4
Figura 2.2 Anatomía del estómago	5
Figura 2.3 Clasificación del cáncer de estómago	7
Figura 2.4 Estructura del código morfológico	9
Figura 3.5 Procesamiento de la base de datos.....	17
Figura 3.6 Posibles estructuras de la información para el análisis de datos. (a) 2 vías y (b) 3 vías	18
Figura 3.7 Matriz Z obtenida mediante Análisis de Componentes Principales.....	20
Figura 3.8 Esquema del Análisis de Datos en el Análisis Triádico Parcial.....	21
Figura 4.9 Residencia pacientes con cáncer de estómago que son atendidos en SOLCA-Guayaquil.....	23
Figura 4.10 Incidencia por Año	24
Figura 4.11 Representación HJ-Biplot de la matriz de individuos CIE topográfico según sexo y grupo morfológico como variable.....	26
Figura 4.12 Sexo y grupo etario de pacientes fallecidos según su año de incidencia.....	28
Figura 4.13 Interestructura y Valor Tipológico del Análisis PTA	29
Figura 4.14 Análisis por separado la información de las	30
Figura 4.15 Proyección de los ejes principales del ACP sobre el PTA.....	31
Figura 4.16 Proyección de los ejes principales año 2014	32
Figura 4.17 Proyección de los ejes principales año 2015 y 2018	33
Figura 4.18 Proyección de los ejes principales año 2016	34
Figura 4.19 Proyección de los ejes principales año 2017	34
Figura 4.20 Proyección de los ejes principales año 2020	35
Figura 4.21 Proyección de los ejes principales año 2021	35
Figura 4.22 Proyección de los ejes principales año 2022	36
Figura 4.23 Proyección de los ejes principales año 2023	37
Figura A1.24 Estructuras y Arreglos Multivariantes	5
Figura E1.25 Estructuras y Arreglos Multivariantes	9

LISTADO DE TABLAS

Tabla 2.1 Clasificación morfológica en base a los diagnósticos según OMS	8
Tabla 2.2 Código estructural para el estómago	10
Tabla 4.3 Distribución de casos de cáncer de estómago según su topografía	25
Tabla B1.4 Valores propios y varianza explicados	6
Tabla B2.5 Contribuciones relativas del factor al elemento filas	6
Tabla C1.6 Valores propios y varianza explicados.....	7

CAPÍTULO 1

1. INTRODUCCIÓN

1.1. Antecedentes

El cáncer gástrico es considerado un problema de salud pública mundial. Según el Global Cancer Observatory (GLOBOCAN) el cáncer de estómago se encuentra en el quinto lugar tanto en incidencia como en mortalidad a nivel mundial (Ferlay et al., s. f.). La mayoría de los diagnósticos morfológicos son adenocarcinomas, el cual se desarrolla a partir de células que forman la capa más interna del estómago (Chao, 2022).

La incidencia del cáncer de estómago muestra una variación significativa, con más del 50% de los nuevos casos emergentes en países en desarrollo debido a factores como la infección por *helicobacter pylori* y dietas ricas en alimentos salados o procesados. Las regiones señaladas con mayor riesgo de incremento del cáncer son Asia oriental, Europa oriental y América central y del sur (Iwu & Iwu-Jaja, 2023).

Ecuador no es ajeno a este problema, ya que el cáncer gástrico es la primera causa de mortalidad en el país, representando el 13.6% del total de defunciones por cánceres (Ferlay et al., s. f.). Además de que este tipo de cáncer ha venido incrementando su impacto en la edad de los pacientes, teniendo que el 59.79% de los fallecidos son mayormente hombres de entre 60 a 84 años y el 51.2% son mujeres en el mismo rango de edad (Real-Cotto et al., 2023). No obstante, aún en la actualidad hay mucho que se desconoce o que falta relacionar con respecto a las características epidemiológicas del cáncer gástrico.

1.2. Descripción del Problema

La caracterización epidemiológica del cáncer de estómago es crucial debido a su alta tasa de mortalidad, lo que resalta la necesidad urgente de una investigación más profunda. Aunque se dispone de una descripción general de la distribución de los casos y su evolución en el tiempo como se discute en el artículo "Mortalidad del Cáncer de

Estómago en Guayaquil, según tendencias históricas”; es fundamental avanzar hacia un enfoque multivariante y detallado del cáncer de estómago (Real-Cotto et al., 2023). Esto implica identificar relaciones entre variables mediante la implementación del método multivariante HJ-Biplot y el análisis triádico parcial.

Esta caracterización no solo aumentaría potencialmente la eficacia de los tratamientos disponibles, sino que también optimizaría las campañas de prevención, lo que a su vez podría llevar a una reducción significativa de la carga de la enfermedad. Por lo tanto, este trabajo propone determinar las características epidemiológicas del cáncer de estómago en pacientes del hospital de la Sociedad de Lucha contra el Cáncer en el Ecuador - SOLCA Guayaquil, entre los años 2014 - 2023, con el fin de establecer mejores estrategias de prevención y control de esta enfermedad.

1.3. Objetivos

1.3.1. Objetivo General

Determinar las características epidemiológicas del cáncer de estómago mediante los métodos de análisis multivariante de última generación en pacientes atendidos un hospital de tratamiento de cáncer en la ciudad de Guayaquil periodo 2014 – 2023.

1.3.2. Objetivos Específicos

- Realizar revisión bibliográfica de aplicación del método multivariante HJ-Biplot en el campo de la salud.
- Caracterizar la epidemiología del cáncer de estómago aplicando técnicas multivariantes.
- Aplicar los métodos multivariantes HJ-Biplot y PTA para relacionar las diferentes caracterizaciones del cáncer de estómago.

1.4. Hipótesis

La aplicación de los métodos de análisis multivariante HJ-Biplot permitirá visualizar si existen variaciones o agrupamientos entre las variables epidemiológicas del cáncer de estómago en pacientes del hospital SOLCA Guayaquil (2014 - 2023)

1.5. Alcance

El estudio se realizará según la incidencia hospitalaria¹ de pacientes diagnosticados con cáncer de estómago en el hospital SOLCA Guayaquil en el periodo del 2014 al 2023. Se excluirán los pacientes que tengan otro tipo de cáncer de órganos digestivos y que no sean atendidos en el hospital.

Se realizará un análisis descriptivo de las variables disponibles para obtener una panorámica clara de la población de estudio y las características generales de la enfermedad. A su vez, se analizarán patrones y relaciones entre variables mediante los métodos de análisis multivariante HJ-Biplot y análisis triádico parcial.

Este trabajo se enfocará en un análisis retrospectivo para obtener estrategias de prevención y control, a fin de contribuir a las políticas de salud pública e información para futuras investigaciones en Ecuador, por lo cual no abordará la efectividad de los tratamientos actuales ni realizará seguimiento clínico de los pacientes involucrados.

¹ La incidencia hospitalaria es cuando un trabajador del hospital revisa o realiza algún procedimiento por primera vez a un paciente nuevo en la institución oncológica.

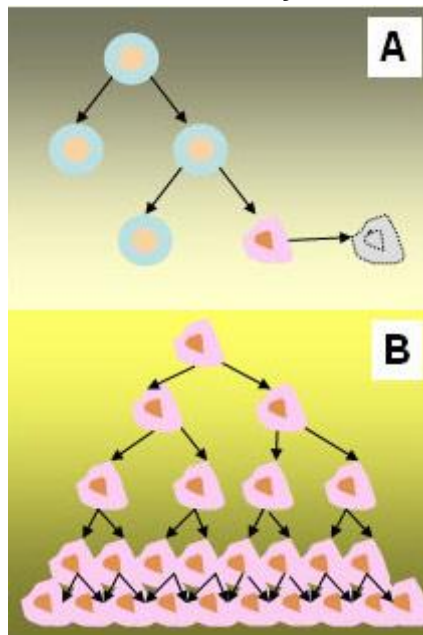
CAPÍTULO 2

2. MARCO TEÓRICO

2.1. Cáncer

El cáncer se caracteriza por el crecimiento de células anormales, las cuales se dividen, crecen y diseminan sin control en cualquier lugar del cuerpo humano, el cáncer es un término que engloba un número grande de enfermedades.

Figura 2.1 Comportamiento normal y anormal de las células

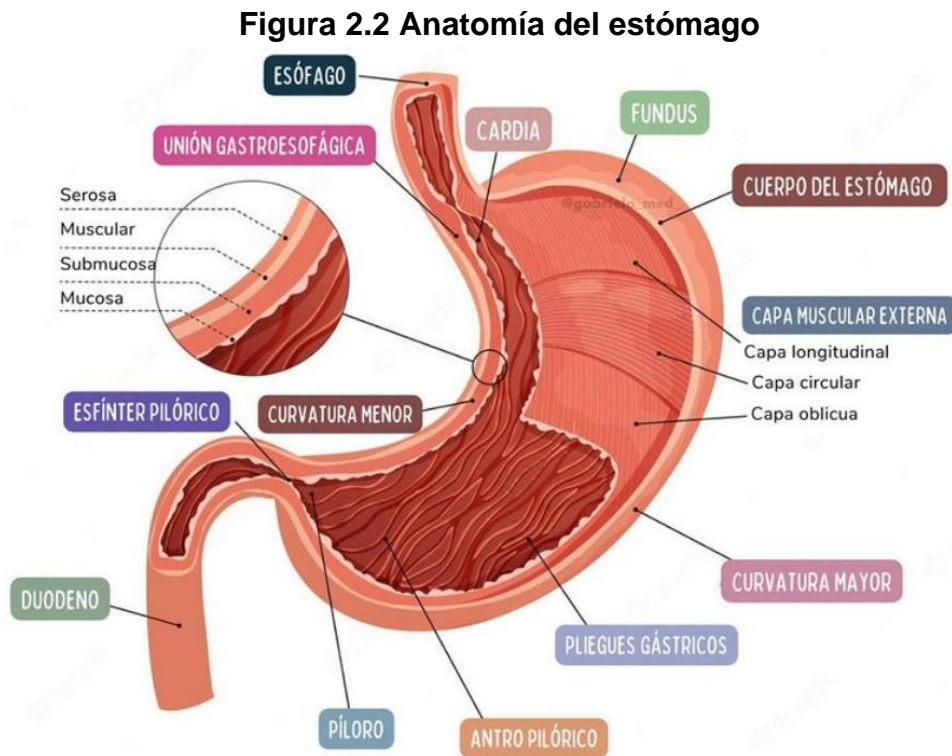


Fuente: Puente, J., & de Velasco, G. (2019). ¿Qué es el cáncer y cómo se desarrolla? - SEOM: Sociedad Española de Oncología Médica

El comportamiento normal de una célula es que se divida una cantidad programada de veces y esta muera una vez cumplido su proceso como muestra la figura 2.1 parte A, por el contrario, las células cancerosas o tumorales pierden esta capacidad de manera que llegan a formar masas como se visualiza en la parte B donde las mutaciones se sobreponen y no son reparadas perdiendo la capacidad de morir, esto se conoce como tumores o neoplasias (Puente & de Velasco, 2019).

2.1.1. Estómago

El estómago es un órgano muscular que pertenece al sistema digestivo, ubicado en la parte superior del abdomen del lado izquierdo, es un órgano que simula la forma de un saco cuya función es recibir los alimentos mezclarlos con los jugos gástricos, descomponerlos y pasarlos al intestino delgado.



Fuente: Med,G (2023). Anatomía del estómago.

Son cinco las partes que le dan forma al estómago las cuales son divididas en estómago proximal y distal. El cardias, fondo y cuerpo forman parte de las tres primeras partes del estómago proximal, en estas partes las células producen ácido y una enzima digestiva llamada pepsina que ayuda a digerir los alimentos. Además de producir una proteína llamada factor intrínseco que se necesita para la absorción de B12. El área inferior formada por el antro y píloro conocido como estómago distal cuenta con dos curvaturas que se señalan en la figura 2 (ACS, 2017).

2.1.2. Desarrollo del Cáncer de Estómago

Para el desarrollo del cáncer de estómago se consideran factores extrínsecos, que con el pasar de los años producen mutaciones genéticas reconocidas como lesiones premalignas, las más comunes gastritis, displasia y adenomas. También se suman factores intrínsecos o factores genéticos, ya que son susceptibles a violentos cambios en el ambiente o en la dieta (Barboza Besada, 1996).

Esta neoplasia maligna inicia en la capa más interna que es la mucosa, a medida que crece se extiende a las otras cuatro capas de la pared del estómago: submucosa, músculo, subserosa y serosa. Cabe resaltar que el cáncer de estómago puede ubicarse y crecer en cualquiera de las partes del órgano como muestra la figura 2.2.

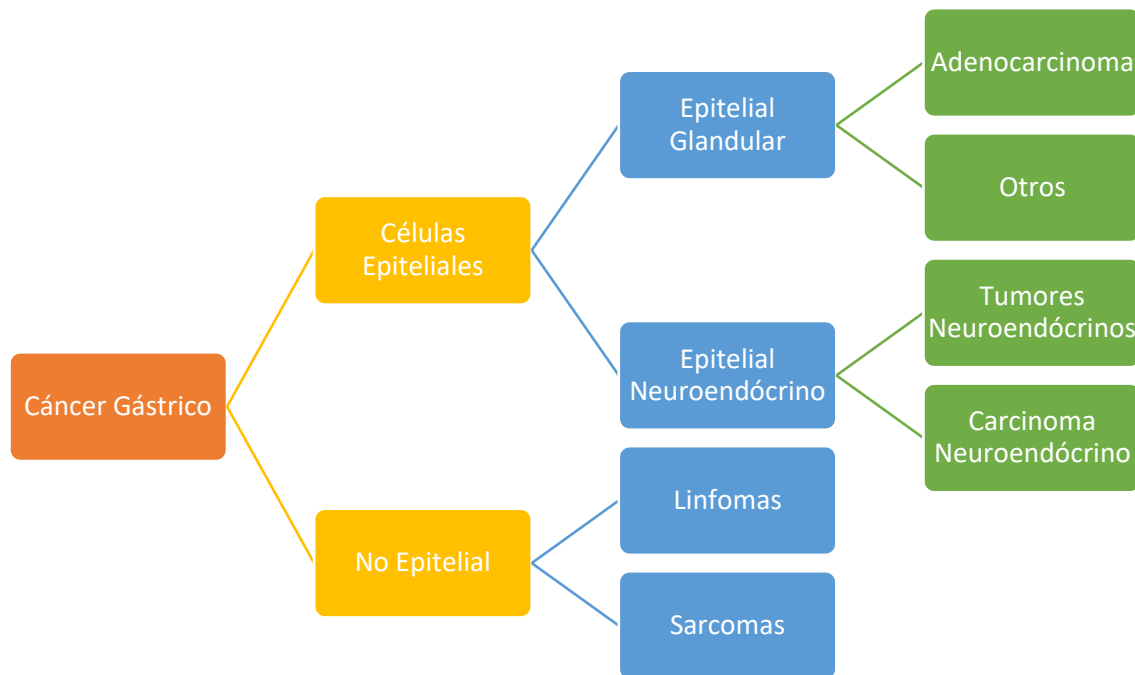
2.1.3. Morfología del Cáncer de Estómago

Según el Instituto Nacional del Cáncer (NIH) el cáncer puede describirse según el tipo de célula que lo formó, por ejemplo, las células epiteliales o las no epiteliales, a esto se lo conoce también como histología² del cáncer.

Según el estado del arte, el adenocarcinoma es una neoplasia maligna epitelial glandular a la cual le corresponde del 90% al 95% de los tumores malignos diagnosticados (Lordick et al., 2022), es el morfológico más frecuente del cáncer de estómago el cual ha sido clasificado de diferentes maneras por variedad de autores, en este caso la clasificación se realiza como se presenta en la figura 2.3 en base a la información de la clasificación de los tumores online de la Organización Mundial de la Salud (OMS). Esta clasificación que se visualiza en la tabla 2.1 se realizó con la supervisión de un patólogo trabajador del hospital SOLCA, accediendo a la 5ta edición de la clasificación de los tumores (Board, 2019).

² Rama de la anatomía que permite analizar como son las células que componen el tumor mediante el estudio microscópico de la estructura, composición y función de los tejidos.

Figura 2.3 Clasificación del cáncer de estómago



Fuente: Elaboración basada en Board, W. C. of T. E. (2019). Digestive System Tumours (5.a ed., Vol. 1). <https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/Digestive-System-Tumours-2019>

La clasificación de la OMS se basa en las características histológicas, como ejemplo se explica la clasificación solo del grupo morfológico adenocarcinoma, es decir, lo clasifica en: Papilar, tubular³, células en anillo de sello, indiferenciado propiamente y mucinoso de manera general. Mientras que la clasificación según Lauren toma en consideración el patrón de crecimiento por lo que lo comprende el tipo intestinal, el difuso y el indiferenciado no productor de moco. Sin olvidar una pequeña proporción de linfomas como diagnóstico morfológico, el cual según el trabajo de (Espejo Romero & Navarrete Siancas, 2003) está relacionada con la infección por *Helicobacter pylori*.

En Perú mediante el uso de la clasificación de Lauren dentro del Instituto de Enfermedad Neoplásica determinaron que el adenocarcinoma de tipo intestinal representa el 52.8%, el tipo difuso el 43.3% el cual afecta más a jóvenes del sexo femenino y el tipo mixto el 3.8%; además de que la localización más común donde se aloja el tumor maligno es en el antro el 80% de casos (Barboza Besada, 1996).

³ Tub1, tub2, tub3 según el grado de diferenciación

Tabla 2.1 Clasificación morfológica en base a los diagnósticos según OMS

Clasificación	Abreviatura	Morfología
Adenocarcinoma	ADK	Adenocarcinoma Adenocarcinoma, intestinal Carcinoma, difuso Adenocarcinoma tubular Carcinoma de células parietales Adenocarcinoma tubuloveloso Adenocarcinoma de células mixtas Adenocarcinoma mucinoso Adenocarcinoma productor de mucina Carcinoma de células en anillo de sello Carcinoma medular con estroma limfoide Adenocarcinofibroma mucinoso
Neuroendocrino	NED	Carcinoma neuroendocrino de células grandes Carcinoma de células pequeñas Gastrinoma maligno Tumor carcinoide Carcinoma neuroendocrino
Indiferenciado	IND	Carcinoma indiferenciado
Adenoescamoso	ADSQ	Carcinoma adenoescamoso
Epidermoide	EPID	Carcinoma epidermoide Carcinoma epidermoide, tipo queratinizante
Linfomas	LINF	Linfoma maligno Linfoma maligno no Hodgkin Linfoma de Hodgkin, celularidad mixta Linfoma maligno difuso de tipo B de células grandes Linfoma de Burkitt Linfoma de células B de zona marginal
Sarcomas	SAR	Leiomioma Sarcoma de estroma gastrointestinal Sarcoma de Kaposi
Carcinomas	CA	Neoplasia maligna Carcinoma

Fuente: Elaboración basada en Board, W. C. of T. E. (2019). Digestive System Tumours (5.a ed., Vol. 1). <https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/Digestive-System-Tumours-2019>

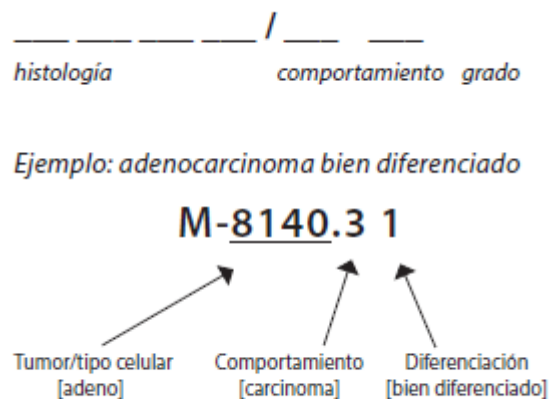
En Ecuador según la caracterización clínica del Hospital de Especialidades Dr. Albert Gilbert Portón el sexo masculino representó el 72.6% siendo mayormente afectado a comparación con el femenino, la localización más frecuente fue el antro (50%) y el tipo morfológico más usual dentro de los adenocarcinomas fue el de tipo intestinal (64.5%), seguido del difuso (29%). El intestinal se presentó en una edad más avanzada, es más

frecuente en el cuerpo del estómago y su localización fue proximal; mientras que el tipo difuso se presentó más en edad temprana, en su mayoría se lo encontró en el antro y su localización fue distal(Cedeño et al., 2021).

2.1.4. Codificación Morfológica y Topográfica Según CIE-O

La Clasificación Internacional de Enfermedades para Oncología (CIE-O) ha sido utilizada cerca de 35 años para la codificación de la localización (topografía) y la histología (morfológica) de la neoplasia, el cual señala la estructura de codificación de la morfología.

Figura 2.4 Estructura del código morfológico



Fuente: Fritz, A., Percy, C., Jack, A., Shanmugaratnam, K., Sobin, L., Parkin, D. M., & Whelan, S. (2003) Clasificación Internacional de enfermedades para oncología (CIE-O)

Como se aprecia en la figura 2.3 los términos morfológicos cuentan con códigos de cinco dígitos que inician en 8000/0 hasta 9992/3. Los primeros cuatro dígitos indican el término histológico específico, el quinto dígito numérico indica el comportamiento, de manera que permite identificar si es un tumor es maligno (1), benigno (0) o in situ (2), o en caso de haber incertidumbre (9) sobre si es benigno o maligno (Fritz et al., 2003).

La tabla 2.2 permite conocer el código estructural para el estómago, además de las descripciones alternativas permitiendo que al momento de codificar se pueda ser lo más específico posible de tener la información necesaria.

Tabla 2.2 Código estructural para el estómago

Código	Área del Estómago	Descripciones Alternativas
C16.0	Cardias, SAI	Unión cardioesofágica, Unión esofagogástrica, Unión gastroesofágica
C16.1	Fondo del estómago	Fondo gástrico, Fórnix gástrico, Fundus gástrico
C16.2	Cuerpo del estómago	Cuerpo gástrico
C16.3	Antro gástrico	Antro del estómago, Antro pilórico
C16.4	Píloro	Canal pilórico, Prepíloro
C16.5	Curvatura menor del estómago, SAI	(no clasificable entre C16.1-C16.4)
C16.6	Curvatura mayor del estómago, SAI	(no clasificable entre C16.0-C16.4)
C16.8	Lesión de sitios continuos del estómago	Pared anterior del estómago, SAI, Pared posterior del estómago, SAI (no clasificable entre C16.0-C16.4)
C16.9	Estómago, SAI	Gástrico, SAI

Fuente: Fritz, A., Percy, C., Jack, A., Shanmugaratnam, K., Sobin, L., Parkin, D. M., & Whelan, S. (2003) Clasificación Internacional de enfermedades para oncología (CIE-O)

2.2. Análisis Multivariante

Análisis multivariante es un conjunto de métodos estadísticos y matemáticos, destinados a describir e interpretar los datos que provienen de la observación de varias variables estadísticas estudiadas conjuntamente (Cuadras, 2012).

A continuación, se contextualiza los conceptos que dan apertura al desarrollo de las metodologías a usar como HJ-Biplot y Análisis Triádico Parcial (PTA).

2.2.1. Análisis Multivariante de Matrices de Dos Dimensiones

Los métodos que forman parte de esta subdivisión del análisis multivariante se centran en estudiar el comportamiento de datos contenidos dentro de una matriz de dos

dimensiones $X_{[N,P]}$ con una estructura formada por N filas también conocidas como individuos y P columnas que representan las variables indicado en (1), una mejor representación en la figura A1.a.

$$X_{[N,P]} = \{x_{ij} | i = 1, \dots, N; j = 1, \dots, P\} \quad (1)$$

2.2.1.1. Análisis de Componentes Principales

Es un método de análisis estadístico multivariante desarrollado por Pearson (1901) el cual permite la simplificación o reducción de la dimensionalidad de los datos conservando la mayor cantidad de información posible, es decir variabilidad. Este trabajo fue enriquecido con las aportaciones de Hotelling (1936) en el cual mantiene un enfoque orientado a la ordenación de filas (individuos) de una matriz en un espacio de baja dimensión, cuyo análisis se basa en la Descomposición de Valores Singulares (DVS) partiendo de una matriz de covarianzas o de correlaciones descrita en (2).

$$X = U\Sigma V^T \quad (2)$$

2.2.1.2. Biplots

El método Biplot se origina de la aportación de Gabriel (1971) el cual proporciona una herramienta útil de análisis de datos que consiste en la representación gráfica estructural de grandes matrices de datos. Lo que caracteriza al método Biplot es la capacidad de mostrar distancias y agrupaciones entre individuos, así como mostrar variaciones y correlaciones entre las variables.

Se basa en la Descomposición de Valores Singulares de la matriz de partida, realizando factorización en matrices de marcadores de filas y columnas. Lo inicialmente propuesto por Gabriel (1971) consta de dos métodos: el primero es GH-Biplot el cual representa de la manera más factible las variables mermando la calidad de representación de los individuos; por el contrario, el JK-Biplot representa mucho mejor los individuos, pero la calidad de representación de las variables es menor.

Teniendo la matriz definida en (1) de rango r , se puede factorizar de la manera presentada a continuación:

$$X \cong GH^T \quad (3)$$

Donde G es una matriz cuya dimensión es $(N \times r)$ y H tiene una dimensión $(P \times r)$. En dicha matriz cada elemento $x_{ij} = g_i^T h_j$ puede ser representado por el i -ésimo producto de la fila G y la j -ésima columna de H^T . Los vectores g_i y h_j son considerados los efectos de las filas y columnas respectivamente. Sea X como la ecuación definida en (2) se determinan las matrices JK (4) y GH (5), en el tercer capítulo se revisará más a detalle.

$$JK : J = U\Sigma \text{ y } K = V \quad (4)$$

$$GH : G = U \text{ y } H = V\Sigma \quad (5)$$

2.2.1.3. HJ-Biplot

En base a lo descrito por Gabriel en el método Biplot se desarrolla la metodología (Galindo-Villardón, 1986), la cual propone una representación simultánea procurando una alta calidad representativa para los marcadores de filas $j_1 \dots j_N$ y para los marcadores de columnas $h_1 \dots h_p$ en un mismo conjunto de referencia de ejes factoriales, método conocido como HJ.Biplot.

$$HJ : H = U\Sigma \quad \text{y} \quad J = V\Sigma \quad (6)$$

Mediante métricas se puede interpretar relaciones entre los ejes factoriales; capturando así las estructuras de covariación multivariantes entre las variables. Llevando esta técnica de análisis multivariante al ámbito médico se podría comprender las diversas interacciones entre los diagnósticos y otros factores demográficos, como el grupo etario y el sexo, para brindar una atención más personalizada y eficiente, específico para un hospital oncológico.

2.2.2. Análisis Multivariante de Matrices de Tres – Vías

En ocasiones existe la necesidad de estudiar el comportamiento de un evento, considerando las características de las unidades de observación (filas) explicadas por múltiples variables (columnas) y su evolución en el tiempo o su comportamiento ante diferentes situaciones (condición). Dicho de otra manera, los datos se estructuran como un conjunto de matrices de tres dimensiones o cubo de datos (Ver Figura A 1.c).

$X[N, P, K]$ es una matriz de tres dimensiones compuesta por los elementos descritos en (7):

$$\{X_{ijk} | i = 1, \dots, N; j = 1, \dots, P; k = 1, \dots, K\} \quad (7)$$

En 1976 aparecen los métodos que estudiarían la estabilidad de la estructura del conjunto de matrices, se desarrollaron con la aportación de Douglas Carroll y Richard Harshman a la matriz de origen propuesta por Ledyard R. Trucker y son reconocidos como la familia STATIS: STATIS, STATIS DUAL y X-STATIS o también conocido como Análisis Triádico Parcial (PTA) (Narváez et al., 2020).

2.2.2.1. Análisis Triádico Parcial o X-STATIS

El análisis triádico parcial (PTA), también conocido como X-STATIS fue propuesto por Jaffrenou (1978). Es un método aplicable a K matrices de datos, con igual N individuos y P variables en las K matrices formando un cubo de datos (Ver Figura A 1.f).

A diferencia del STATIS que trabaja con operadores el X-STATIS hace uso directamente de la matriz. Vectoriza cada matriz y las concatena convirtiéndolas en columnas, generando una tabla bidimensional, a la cual se aplica análisis de componentes principales (ACP) para bajar la dimensión, resultando en una matriz consenso formada por la primera componente principal. Desde el punto de vista matemático, inicia con una matriz de tres dimensiones, la cual contiene igual N individuos y P variables en K condición.

$$X[N, P, K] \quad (8)$$

Cada matriz cuenta con los tripletes que contienen la matriz diagonal D_N con los pesos para las filas y D_P los pesos para las columnas de cada K -matriz.

$$(X_1, D_N, D_P), (X_2, D_N, D_P), \dots, (X_K, D_N, D_P) \quad (9)$$

El siguiente paso es el calcular el producto interno entre las K -matrices, tal como:

$$COVV(X_K, X_N) = Trace(X_K^T D_N X_K D_P) = Trace(X_K^T D_N X_K D_P) \quad (10)$$

Seguido del cálculo de la Varianza vectorial:

$$VARV (X_K, X_K) = Trace (X_K^T D_N X_K D_P) \quad (11)$$

Coeficiente de correlación vectorial (RV):

$$RV (X_K, X_N) = \frac{COVV(X_K, X_N)}{\sqrt{VAR(X_K)}\sqrt{VAR(X_N)}} \quad (12)$$

2.2.3. Aplicaciones de los Métodos Multivariantes

En la actualidad, existe un incremento de la aplicación de los diferentes métodos de análisis multivariante en diferentes ramas de la ciencia incluyendo la medicina, un ejemplo claro es lo que resalta Arenaza (2022) en su trabajo donde explica que hay estudios en los que herramientas estadísticas como las técnicas de análisis multivariante ayudan a determinar la contribución relativa de diferentes causas a un determinado evento o resultado. Los investigadores clínicos necesitan este tipo de análisis, dado que las enfermedades presentan múltiples causas y el pronóstico se determina, usualmente, por un gran número de factores.

En un estudio publicado por Patino-Alonso et al. (2015) en Cambridge University Press, comenta que es poco el conocimiento acerca de los patrones de agrupamiento en el estilo de vida de los adultos de España, por lo que hace uso de un método de agrupación basado en las coordenadas HJ-Biplot para comprender las variables subyacentes a estas agrupaciones. El HJ-Biplot permite un análisis exhaustivo de todas las variables y revela tres conglomerados diferentes entre sí, en una muestra aleatoria de centros sanitarios de un amplio rango de edad (20-80 años) incluyendo sujetos sanos y con diferentes niveles de riesgo cardiovascular, los cuales son caracterizados por edad, sexo y variables clínicas posteriormente.

En Ecuador se realizó un estudio cuyo objetivo fue realizar un análisis multivariado HJ-Biplot de la ocurrencia de *Helicobacter pylori* como riesgo para cáncer gástrico, el cual ha permitido generar una nueva interpretación y conocimiento facilitando la interpretación entre las similitudes y disimilitudes de los diferentes individuos (puntos de muestreo) y la covariación entre las variables hábitos de higiene y afecciones que produce el bacilo *H. pylori*. Lo más importante de la publicación es haber determinado qué variables caracterizan las diferencias entre los distintos puntos de muestreo. De

manera que se obtuvo dos grupos de puntos de muestra que coinciden satisfactoriamente con las estaciones de la región (Cisneros et al., 2020).

La tesis de Hernández (2015) analiza el uso del tiempo de los colombianos con una edad superior a los 10 años mediante el método HJ-Biplot y X-STATIS, en el cual se obtiene que existen diferencias entre regiones tanto en la variabilidad como en la covariación de las variables. Asimismo, a nivel país se captan diferencias entre las variables que explican y caracterizan cada uno de los grupos de edades, además de percibir que, dentro de estos grupos etarios, el comportamiento de los géneros y estratos socioeconómicos no es similar. Y el trabajo recalca que los resultados evidencian complementariedad entre los métodos de análisis multivariantes aplicados.

CAPÍTULO 3

3. METODOLOGÍA

En el presente capítulo se presentan las propuestas de técnicas multivariantes a aplicar de manera efectiva, por lo que se consultará bases de datos relevantes, filtrando fecha de publicación y aplicabilidad al contexto de este estudio.

3.1. Base de Datos

Los datos son obtenidos de la base anonimizada de incidencia hospitalaria de SOLCA Guayaquil y atendidos durante el periodo 2014 - 2023. Se excluirán los pacientes que tengan otro tipo de cáncer de órganos digestivos y que no sean atendidos en el hospital, también será parte de la exclusión aquellos que no cuente con información demográfica completa.

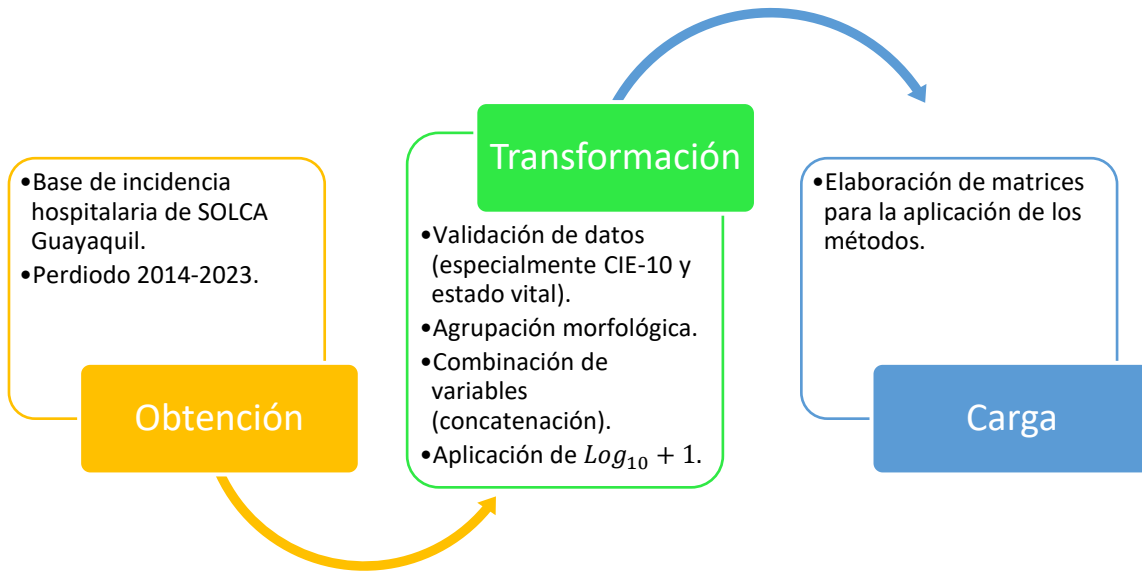
3.2. Diseño

El estudio tendrá un diseño observacional de corte transversal, de tipo descriptivo multivariante y con un enfoque cuantitativo.

3.3. Procesamiento de Base

La base de datos inicial era de 2029 filas y 29 variables, el cuál después de un proceso de transformación, es decir filtrar, limpiar, deduplicar, validar y autenticar los datos, en especial de la topografía del cáncer y el estado vital del paciente cuya validación tiene fecha de corte al 17/05/2024; el total de filas quedó en 1999 y 33 variables que fueron resultado de la concatenación de diferentes variables (véase figura 3.5). A su vez dentro del proceso de transformación se realizó la agrupación morfológica según la clasificación OMS (figura 2.3) en base al diagnóstico patológico de la base con la guía del especialista véase tabla 2.1.

Figura 3.5 Procesamiento de la base de datos



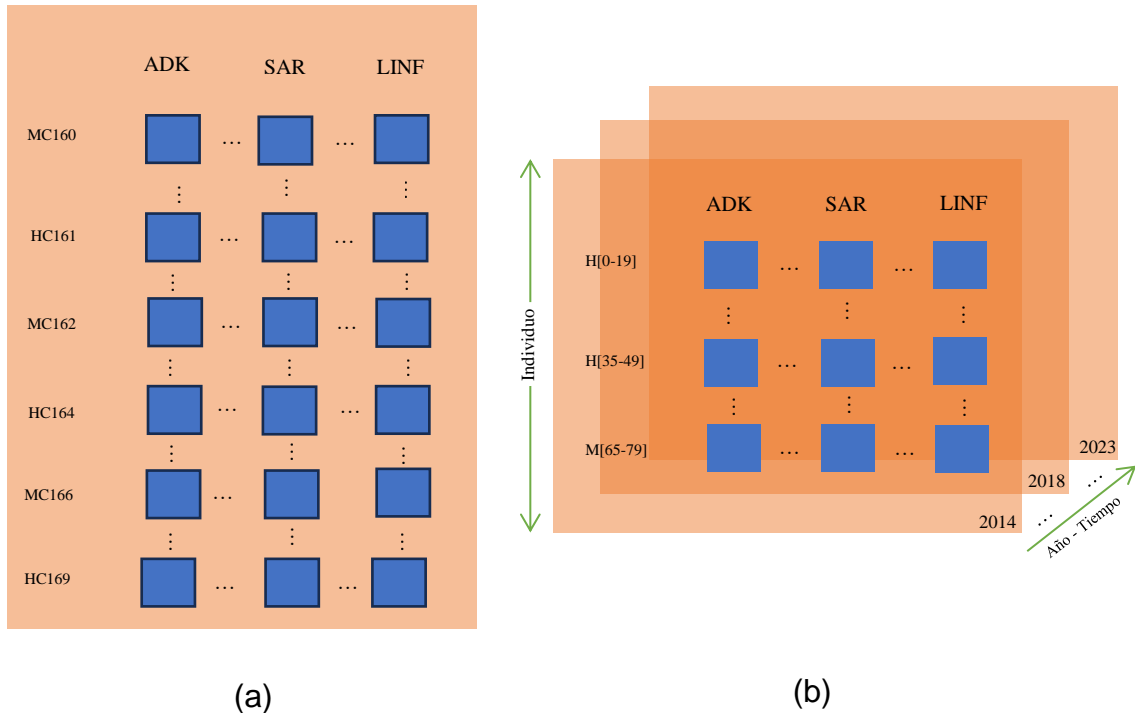
Fuente: Elaboración propia

En este trabajo se obtuvo que el 92% de los casos tuvieron un diagnóstico morfológico dentro del grupo de adenocarcinoma, presentando gran abundancia de datos en este grupo morfológico, por lo que se aplicó $\text{Log}_{10} + 1$ ya que comprime los valores grandes más que los pequeños, reduciendo así la asimetría y estabilizando la varianza (Caspers, 1981).

3.4. Metodología Estadística

En este trabajo se procesan los datos de la siguiente forma: se realizó la concatenación de variables como CIE topográfico y sexo para formar a los individuos que serán parte de la estructura de la matriz para el análisis HJ-Biplot (figura 3.6 a) y de variables como sexo, grupo etario y año de incidencia para formar los individuos que estructuran la matriz para el análisis PTA (figura 3.6 b) en ambas matrices las agrupaciones morfológicas de la tabla 2.1 son las variables, con excepción de indiferenciado (IND) y adenoescamoso (ADSQ) para la matriz de PTA, ya que estas dos presentan poca variabilidad (muy pocos datos), por lo que no aportan suficiente información.

**Figura 3.6 Posibles estructuras de la información para el análisis de datos.
(a) 2 vías y (b) 3 vías**



Fuente: Elaboración propia

3.4.1. Método HJ-Biplot

En el presente trabajo se decidió por la aplicación de la técnica multivariante llamada HJ-Biplot propuesto por (Galindo-Villardón, 1986) debido a la maximización de la representación de calidad para el mismo sistema de coordenadas y también lo versátil que es como herramienta gráfica multivariante de los datos de una matriz $X_{(N,P)}$; donde N presenta las filas (la topografía del cáncer y sexo) mediante $i = 1, \dots, N$ y P para las columnas (variables de grupo morfológico) mediante $j = 1, \dots, P$; así mismo se aplicaría para el segundo análisis donde N pasa a ser estado vital del paciente según el año de incidencia y P es el grupo etario por sexo.

Los métodos Biplot parten de los mismos principios del análisis de componentes principales y DVS de una matriz de datos. El algoritmo comienza con la descomposición de valores y vectores propios de la matriz definida previamente:

$$X = U\Sigma V^T$$

Donde,

- X es la matriz de datos.
- U es la matriz de datos cuyas columnas contienen los vectores propios de XX^T .
- V es la matriz de datos cuyas columnas contienen los vectores propios de $X^T X$.
- Σ es la matriz diagonal que contiene los valores propios de X .

U y V deben ser ortogonales, es decir, $U^T U = I$ y $V^T V = I$, para garantizar la unicidad de la factorización. Considerando que la matriz X está centrada, los marcadores para las columnas en el HJ-Biplot se emparejan con las columnas marcadas del GH Biplot; a su vez, los marcadores para las filas se hacen coincidir con las filas de marcadores del JK Biplot, como se muestra en la ecuación (6). Según Cubilla-Montilla et al. (2021) las reglas para la interpretación del HJ-Biplot son:

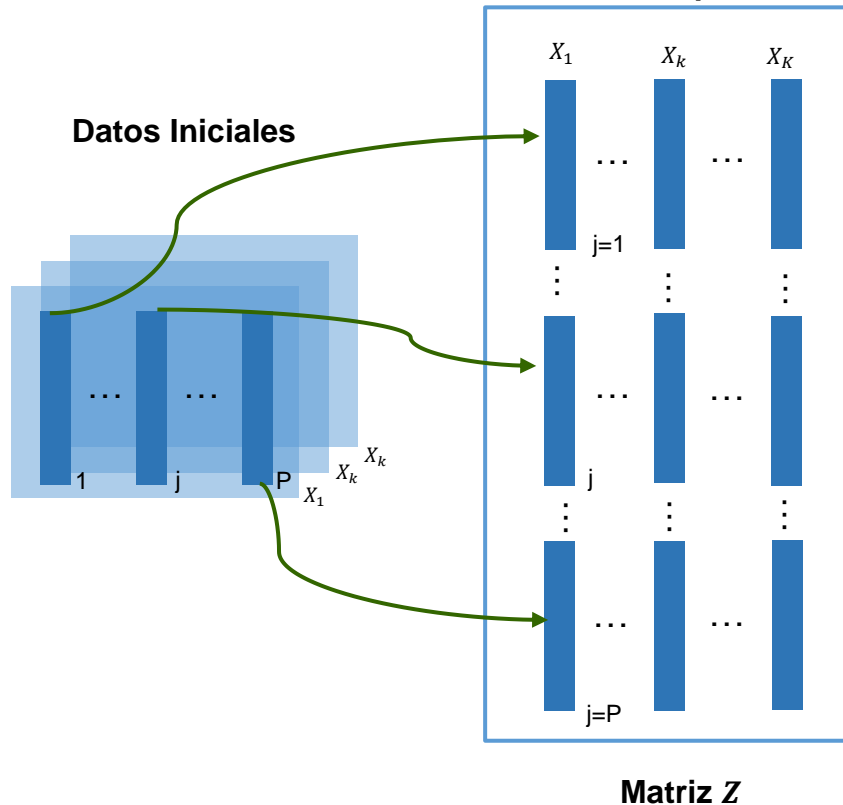
- La proximidad entre los puntos representa los marcadores de filas que se interpreta como la similitud entre ellos, dando paso a la identificación de grupos de individuos con perfiles similares.
- Mediante la longitud de los vectores se puede estimar la desviación típica de una variable.
- A partir de los ángulos entre los vectores se puede capturar la correlación entre las variables. Si dos variables están correlacionadas, tendrán un ángulo agudo; por el contrario, si es obtuso las variables presentarán una correlación negativa; y si el ángulo es recto indica que las variables no están correlacionadas.
- Los puntos proyectados ortogonalmente sobre una variable aproximan la posición de los valores de muestra en esa variable.

3.4.2. Método Análisis Triádico Parcial o XSTATIS

El Análisis Triádico Parcial o XSTATIS es una técnica de análisis de datos multivariante que permite estudiar simultáneamente las sub-matrices de datos cuantitativos, con la finalidad de analizar la posibilidad de patrones similares de comportamiento. En este contexto sería detectar cualquier patrón común entre los diferentes grupos de edad y

sexo, en relación con las variables de clasificación morfológica, justificando el uso de una técnica de análisis de datos de tres vías, ya que se analiza un cubo de datos. Para el análisis de la información de los diez años, es adecuado usar esta técnica de análisis multivariante.

Figura 3.7 Matriz Z obtenida mediante Análisis de Componentes Principales



Fuente: Hernández, E. J. M. (2015). Análisis Multivariante del Uso del Tiempo de la Población Colombiana Mediante el HJ Biplot y el Análisis Triádico Parcial

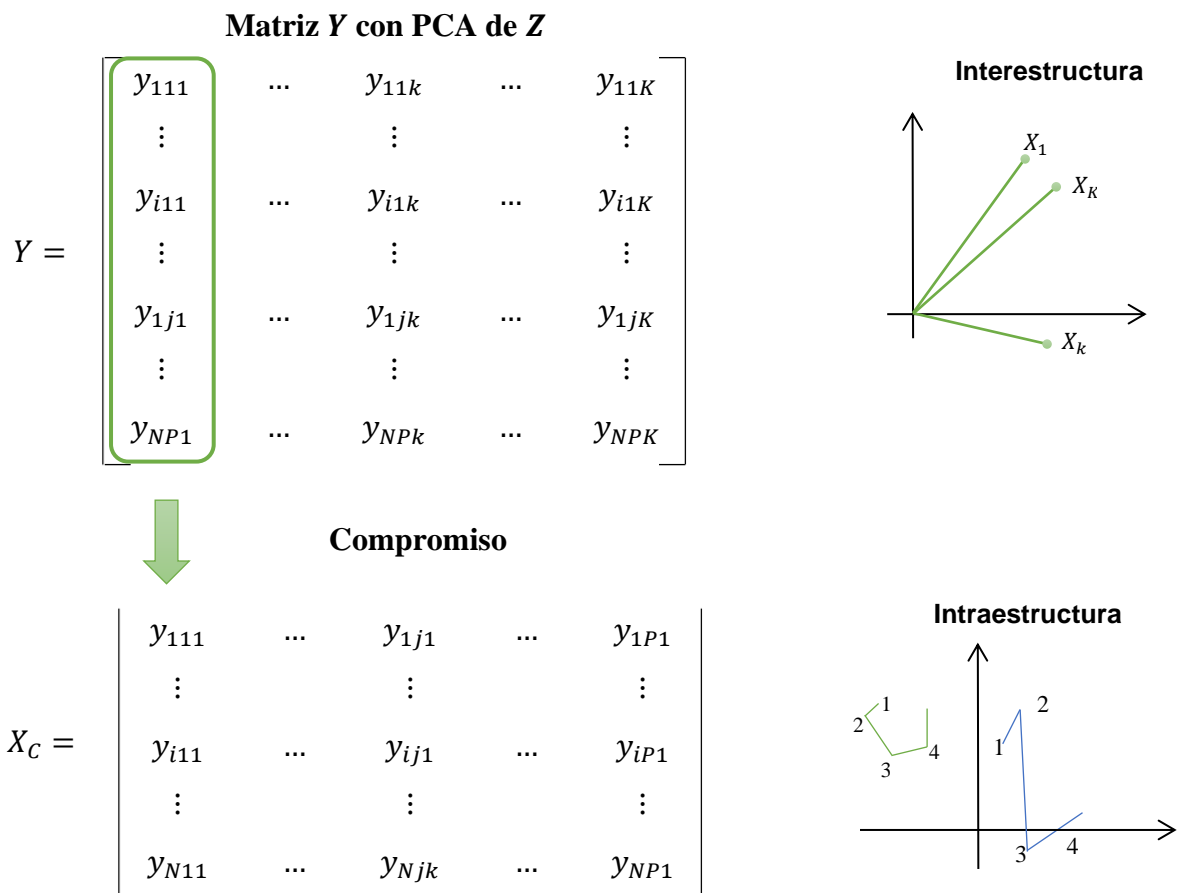
El PTA no supone operadores lo que da una ventaja frente a otros métodos, debido a que no pierde información original; además es mayormente restrictiva ya que supone las mismas P variables medidas en los mismos N individuos en todas las ocasiones permitiendo más representaciones.

El procedimiento del ATP puede resumirse en tres etapas principales. Como primer paso se analiza la **Interestructura**, por lo que se construye una matriz Z a partir de varias matrices de datos ver ecuación (7); las cuales están formadas por los mismos N individuos (filas) y mismas P variables (columnas). Se construye la matriz Z , donde cada vector columna corresponde con una de las K matrices de manera extendida. Es por ello que la matriz Z está formada por tantas columnas como matrices en análisis, y puede visualizarse como una tabla bidimensional ver figura 3.7. A continuación, se

realiza un Análisis de Componentes Principales (PCA) sobre Z para obtener una nueva matriz Y , que simplifica y resume la información a través de las principales direcciones de variabilidad.

El segundo paso es la construcción y análisis de la estructura de la matriz compromiso la cual tiene la tarea de resumir la información de las K matrices de partida en una sola. A partir del primer componente de la matriz Y , se construye la matriz de compromiso X_c , que es una especie de media ponderada de las matrices originales. Los pesos para esta ponderación provienen de los primeros vectores propios del ACP de Z . Esta matriz de compromiso busca maximizar la correlación media entre sus variables y las de cada matriz original, proporcionando una representación que resalta las estructuras comunes a todas las matrices véase figura 3.8.

Figura 3.8 Esquema del Análisis de Datos en el Análisis Triádico Parcial



Fuente: Hernández, E. J. M. (2015). Análisis Multivariante del Uso del Tiempo de la Población Colombiana Mediante el HJ Biplot y el Análisis Triádico Parcial

Finalmente, el último paso consiste en el estudio de la **Intraestructura** el cual implica analizar la reproducibilidad y la estabilidad de la matriz de compromiso. Se examina cómo cada individuo o variable mantiene su posición relativa a lo largo del tiempo o entre condiciones diferentes. Se pueden identificar trayectorias estables o inestables: las trayectorias estables muestran poca variación a lo largo del tiempo, indicando que la variable o individuo es consistente a través de las diferentes condiciones; las trayectorias inestables, por otro lado, muestran una gran variación, señalando cambios significativos en la estructura o comportamiento del individuo a lo largo del tiempo.

3.5. Tecnologías Necesarias

Para la obtención de los resultados de ambos métodos multivariantes se hará uso de dos programas estadísticos de libre uso. El primero es Multbiplot, programa desarrollado por Vicente Villardon (2010), diseñado para realizar análisis Biplot, entre los cuales se encuentra el método de interés HJ-Biplot. El programa fue concebido para llenar el vacío entre las imágenes estáticas y una interpretación visual más dinámica.

RStudio se presenta como el segundo programa a usar, el cual es un software conocido por su amplio abanico de bibliotecas y herramientas que permiten la limpieza, preparación de datos, visualización, entre otras opciones variadas. No obstante, el paquete elaborado por Thioulouse et al. (1997) es el de interés para aplicar y obtener los resultados del método XTATIS. ADE-4 es un paquete de análisis multivariante y visualización gráfica que dispone de métodos habituales de una sola tabla, como el análisis de componentes principales y el análisis de correspondencias y varios métodos de análisis de tabla de tres vías (tabla k). La interfaz de usuario es sencilla y homogénea; esto contribuye a que el uso de ADE-4 sea muy fácil para los no especialistas en estadística o análisis de datos.

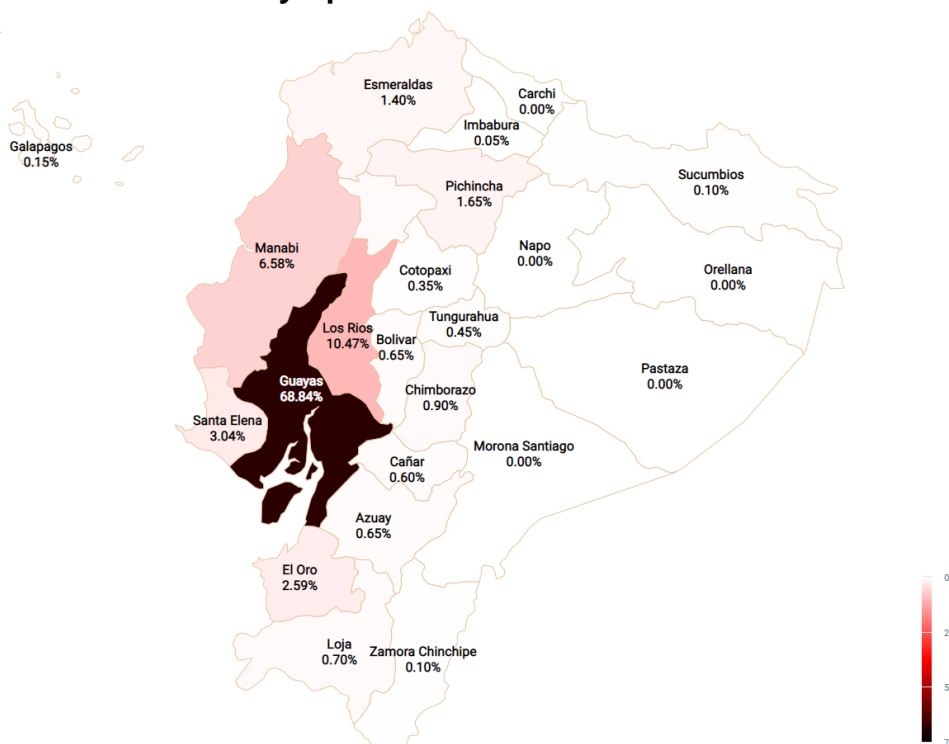
CAPÍTULO 4

4. RESULTADOS

En este capítulo se presentan los resultados del estudio para determinar las características epidemiológicas del cáncer de estómago mediante la aplicación de los métodos multivariantes HJ-Biplot y Análisis Triádico Parcial o X-Statís.

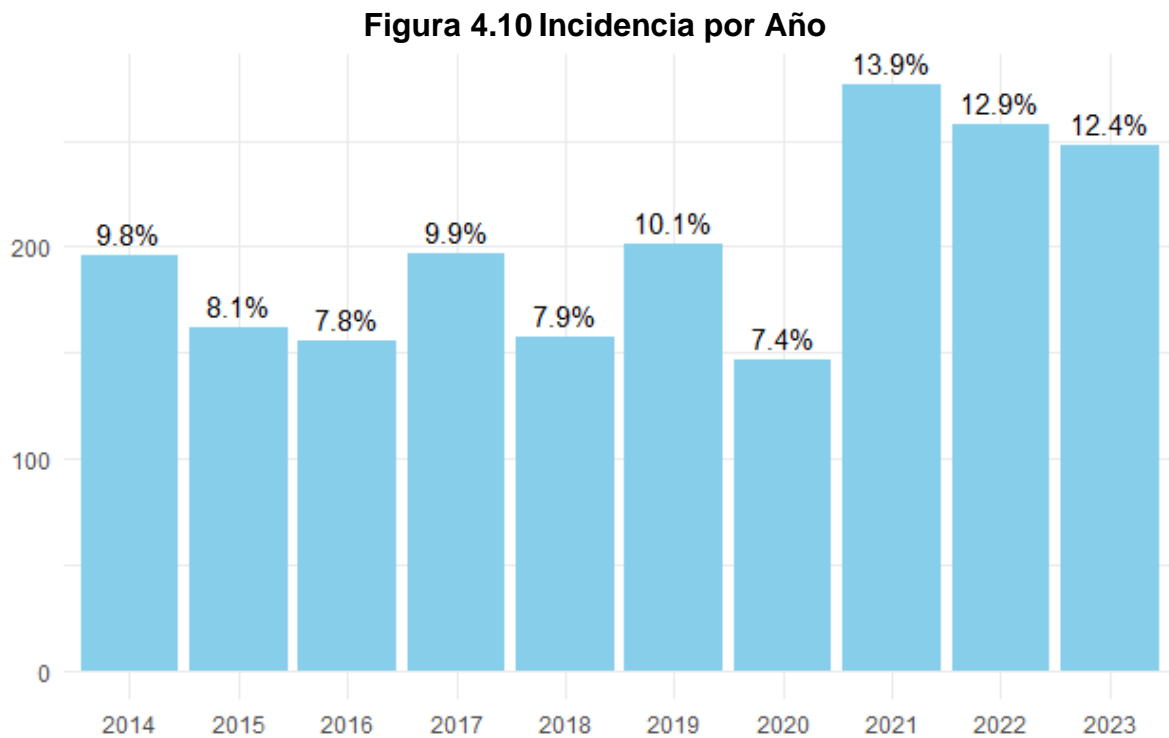
La muestra del estudio consistió en 1.999 casos de incidencia hospitalaria con cáncer de estómago entre los años 2014 al 2023. La distribución demográfica de los datos fue del 58.53% en los hombres y el 41.47% en las mujeres, donde gran cantidad de los casos de ambos sexos se agrupan entre las edades de 50 a 79 años, afectando más a hombres que mujeres.

Figura 4.9 Residencia pacientes con cáncer de estómago que son atendidos en SOLCA-Guayaquil



Fuente: Elaboración propia

A continuación, el gráfico 4.9 presenta un mapa del Ecuador donde señala que el 68.84% de los casos llegan de la provincia del Guayas, seguido de Los Ríos y Manabí en mayores concentraciones (10.47% y 6.58% respectivamente), y en menor porcentaje provienen de las provincias de Santa Elena, El Oro, Esmeraldas, entre otras. A su vez, se tiene que el 83.15% de los pacientes residen dentro del área de influencia del Hospital SOLCA Guayaquil⁴.



Fuente: Elaboración propia

La variabilidad de los nuevos casos que llegan al hospital es considerable anualmente, aunque desde el año 2014 al 2018 tiende a mantenerse baja, los tres últimos destacan significativamente en comparación con los otros años. Debido a factores externos como la pandemia del COVID-19 el 2020 es el año que presenta una menor cantidad de casos con un 7.4% de los datos de la base, mientras que el año siguiente presenta un vasto crecimiento de pacientes nuevos en la institución oncológica con un 13.9%. Parte de lo que caracteriza al cáncer gástrico es su alta mortalidad, en base a esto se tiene que el 78.49% de los pacientes que conforman la data han fallecido, y solo el 21.51% continúan en su lucha.

⁴ Está formada por las provincias Guayas, Los Ríos, Santa Elena, Bolívar y Galápagos.

4.1. Análisis Multivariante

4.1.1. HJ-Biplot

4.1.1.1. Topografía del Cáncer de Estómago

En el HJ-Biplot obtenido para observar la relación y comportamiento de la topografía según el sexo del paciente con el grupo morfológico del cáncer, se tiene que entre los dos primeros ejes se explica el 77.15% de la varianza total, mayormente concentrada en el eje 1, véase anexo b.

La tabla de Contribuciones relativas del Factor al elemento (CRFE) que se encuentra en el anexo b, primero para las columnas o variables (grupo morfológico) y luego para las filas o individuos (CIE topográfico según sexo) indica los elementos que son característicos del eje 1 o del eje 2. En este caso si se observa el anexo se tiene que en las columnas el grupo morfológico epidermoide se explica mejor en el eje 3, mientras que en las filas los hombres con diagnóstico topográfico en C16.3, C16.2, C16.8 y C16.0, este último incluye al sexo femenino se explican mejor en el eje 2.

Por último, el C16.8 en mujeres está mejor representado en el eje 3; aquí entra la tabla 4.3 que muestra la distribución de los casos donde el antro o C16.3 es el segundo con mayor porcentaje de información, lo cual explica porque no está bien representado en la figura 4.11.

Tabla 4.3 Distribución de casos de cáncer de estómago según su topografía

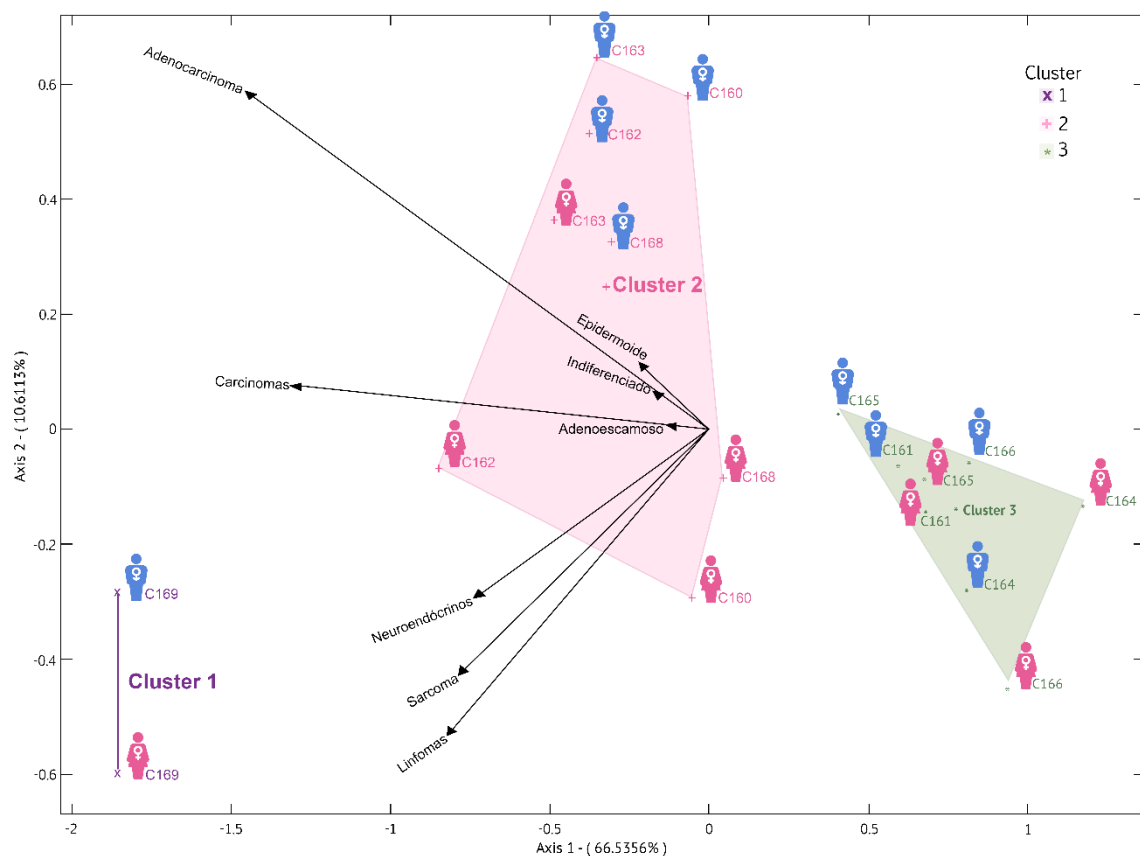
Cie Topográfico	% Total
C16.0	7,25%
C16.1	2,75%
C16.2	17,76%
C16.3	20,16%
C16.4	1,50%
C16.5	4,70%
C16.6	1,90%
C16.8	9,30%
C16.9	34,67%

Fuente: Elaboración propia

Profundizando en la figura 4.11 se tiene al grupo morfológico adenocarcinoma como el vector de mayor longitud lo cual aproxima la variabilidad de la variable, seguida del grupo denominado carcinoma. Por otro lado, el grupo epidermoide y linfoma muestran un ángulo de 90° señalando una relación nula entre ambas variables, esta última al estar cerca de los neuroendocrinos y sarcomas, se considera que tienen un comportamiento similar; mientras que los vectores más cortos presentan poca contribución a la variabilidad como los grupos epidermoide, adenoescamoso e indiferenciado.

El gráfico presenta a su vez tres clústeres donde todos involucran a ambos sexos y se centra más en la ubicación del cáncer al momento del diagnóstico. El primer clúster se lo denominará inespecífico, ya que lo representa el 34.67% de los pacientes con código C16.9 que es estómago SAI, el cual se usa cuando el patólogo no logra identificar el sitio específico de desarrollo del tumor, ya sea porque el estado del cáncer del paciente es avanzado o presentan otros inconvenientes.

Figura 4.11 Representación HJ-Biplot de la matriz de individuos CIE topográfico según sexo y grupo morfológico como variable



Fuente: Elaboración propia

El segundo clúster se llamará frecuentes, en estos casos los patólogos logran la especificidad deseada. Además, este clúster lo conforman las localizaciones que suelen identificarse de manera más frecuente (véase tabla 4.3), donde el líder es el topográfico C16.3 que representa al antro gástrico, seguido del cuerpo gástrico (C16.2), lesión de sitios contiguos⁵ (C16.8) y cardias (C16.0), en estas ubicaciones del estómago predomina el grupo adenocarcinoma.

El clúster verde y último se encuentra por detrás de las flechas, ya que aporta menos información, no obstante, aunque es mínima es interesante la agrupación que se visualiza en el gráfico, ya que son localizaciones poco usuales de identificar, unas más que otras. La ubicación menos usual de este clúster es el píloro (C16.4).

4.1.1.2. Estado Vital Según el Año de Incidencia

El HJ-Biplot a continuación representa la relación de los grupos etarios según el sexo del paciente (variables) y la combinación del estado vital de los pacientes por año de incidencia (individuos). La figura 4.12 muestra como el eje 1 explica el 84.63% y el eje 2 el 4.18% que juntos, estos dos ejes explican el 88.82% de la varianza total. La tabla CRFE columna que se encuentra en el apéndice c, se logra apreciar que el grupo etario pediátrico [0-19] en ambos sexos se explica mejor en el eje 2.

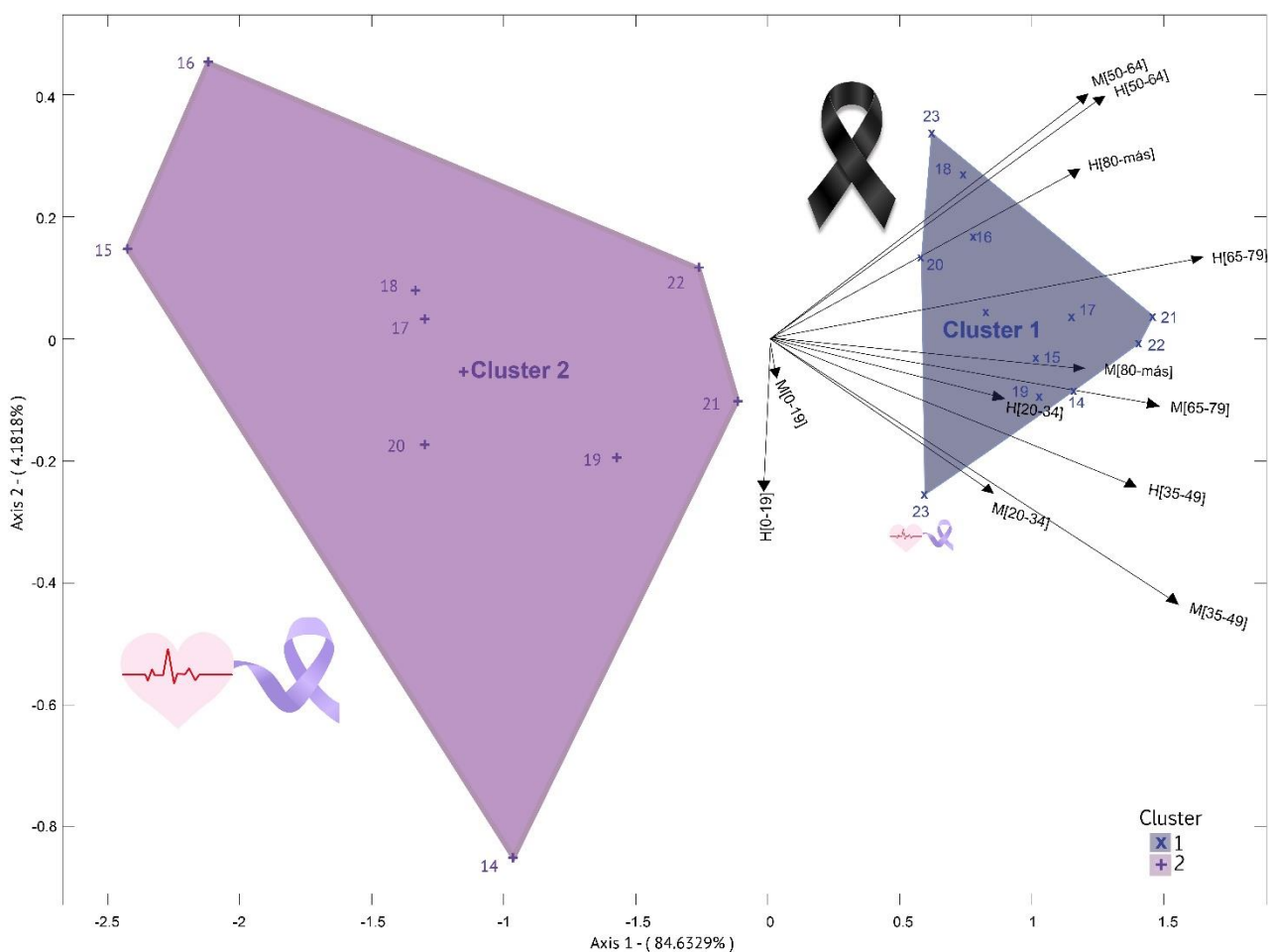
Observando la longitud de las variables la más larga es la de los hombres en las edades de entre 65 a 79 años el cual aporta mayor variabilidad, a su vez se aprecia que el comportamiento del grupo etario pediátrico es opuesto al resto de los grupos. Teniendo que entre mayor sea el paciente indiferente de su sexo aportará más información.

En el gráfico a continuación se visualizan dos clústeres que representan el estado vital de los pacientes. En el primer clúster predomina los pacientes de los diferentes años de incidencia hospitalaria que han fallecido, teniendo solo pacientes vivos del año 2023. Se puede apreciar que los años de incidencia del 2021 y 2022 aportan más información tanto para hombres y mujeres en edad de entre 65 a 79 años, ya que es el rango en el que se diagnostica más el cáncer de estómago y presenta mayor cantidad de muertes

⁵ Lesión de sitios continuos del estómago o C16.8, refiere también a los casos donde existe la combinación de dos o más topografías, ejemplo tumor se sitúa en cuerpo y antro.

por cáncer de estómago primordialmente en el sexo masculino. Para ambos sexos otro grupo etario importante va desde los 50 a 64 años, ya que tienen un comportamiento similar al estar correlacionados y son los segundos en presentar mayor incidencia y mortalidad en cantidad. En general el comportamiento en los demás grupos etarios es similar indiferente del sexo, excepto por las edades de 35 a 49 años donde las mujeres presentan mayor mortalidad que el hombre.

Figura 4.12 Sexo y grupo etario de pacientes fallecidos según su año de incidencia



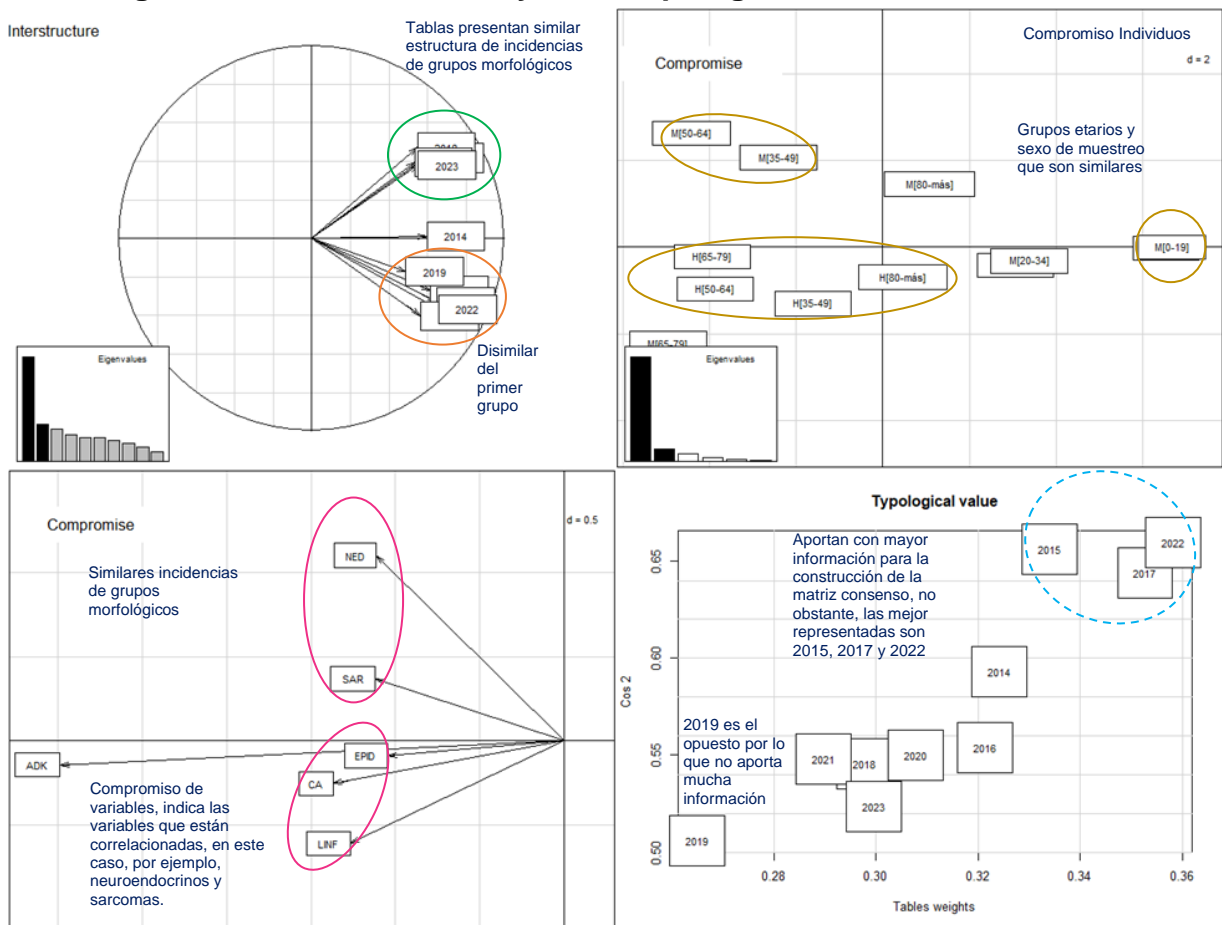
Fuente: Elaboración propia

El segundo clúster son los pacientes vivos de los distintos años de incidencia hospitalaria, los años 2021 y 2022 en este conglomerado son los que más información aportan, lo interesante de este clúster es que de los casos pediátricos, es decir, entre las edades de 0 a 19 años incluidos en el estudio que han presentado cáncer gástrico ninguno ha fallecido a la fecha.

4.1.2. Análisis Triádico Parcial

En el área de la interestructura se puede observar una representación de los años 2014 al 2023, indicando una relación temporal entre los años. Claramente la interestructura se divide en dos grupos donde uno es disimilar del otro; por tanto, se tiene que los años dentro del círculo naranja en la figura 4.13 presentan similares incidencias de grupos morfológicos que lo diferencian de los años del círculo verde.

Figura 4.13 Interestructura y Valor Tipológico del Análisis PTA

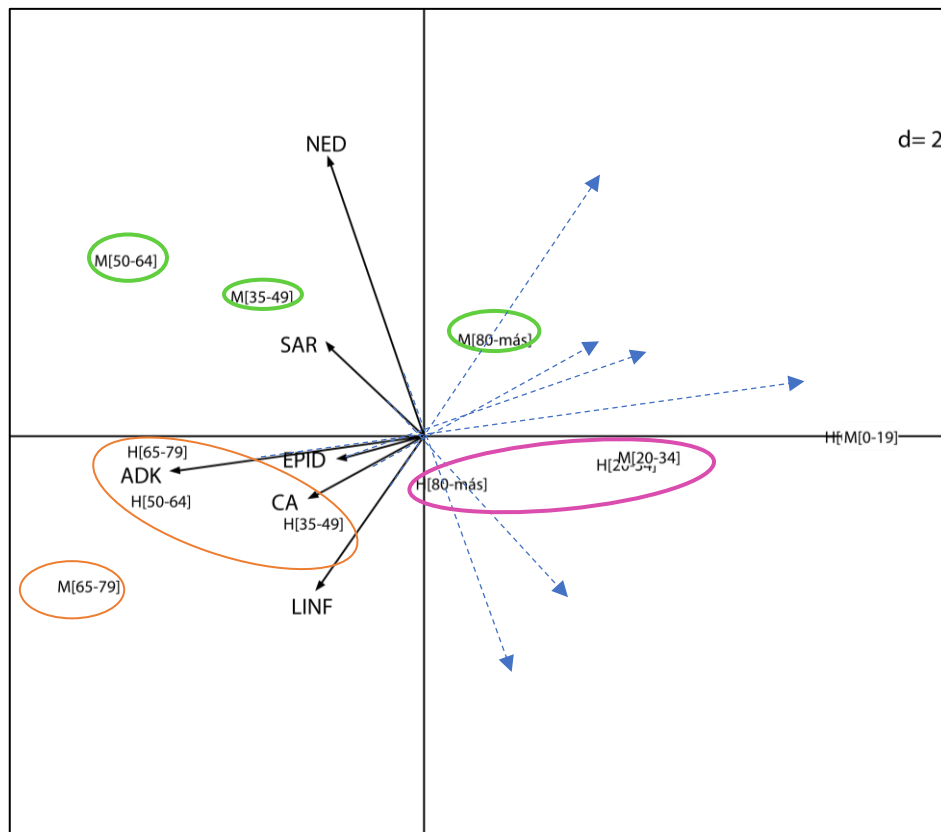


Fuente: Elaboración propia mediante el programa RStudio

El compromiso de individuos indica los grupos etarios según el sexo que son similares en incidencia de grupos morfológicos, como los que están señalados en las diferentes elipses color mostaza. Mientras que el compromiso ubicado bajo el gráfico de interestructura muestra las relaciones entre las variables, indicando correlación significativa como es el caso del grupo morfológico epidermoide (EPID), carcinoma (CA), linfoma (LINF) y adenocarcinoma (ADK) las cuales entre si están cercanas.

En el área de typological value se puede visualizar que las matrices del año 2017, 2015 y 2022 son la de mayor peso, por lo que aportan con más información y son mejor representadas por la matriz consenso, por el contrario, el año 2019 tiene poca representación, véase figura 4.13.

Figura 4.14 Análisis por separado la información de las dos matrices compromiso



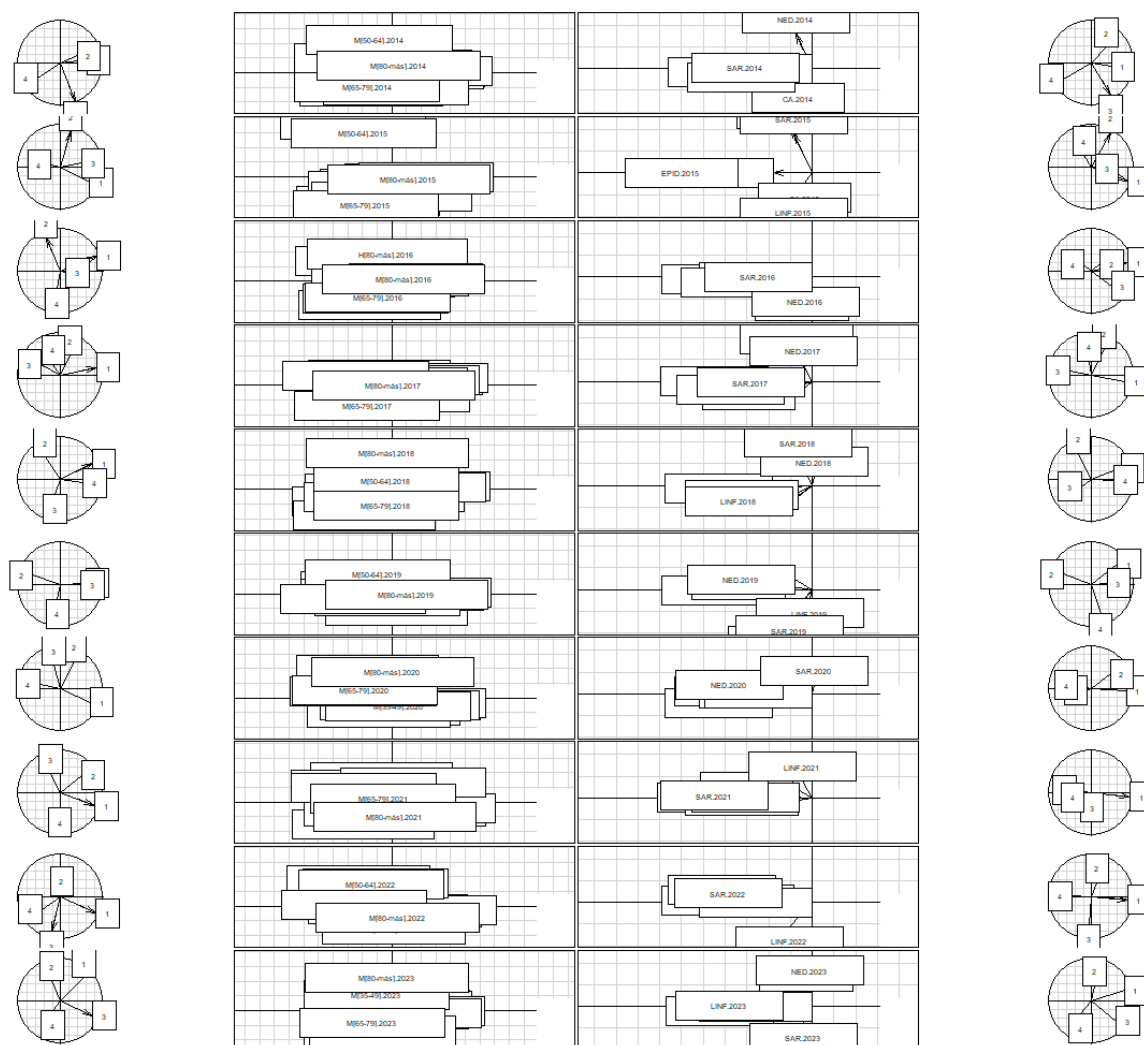
Fuente: Elaboración propia mediante paquete Ade4 - RStudio

Profundizando en las matrices de compromiso estas indican los grupos etarios con su respectivo sexo que son similares en condiciones de grupos morfológicos, como los que están señalados; ejemplo las mujeres en los rangos de edad de 35-49 y 50-64 años presentan valores por encima del promedio de las variables neuroendocrino y sarcoma; mientras que hombres con 80 y más años presentan valores por debajo del promedio de las mismas variables.

Para la obtención del gráfico de proyección de los ejes principales se debe realizar una serie de pasos descritos a continuación; primero revisar los coeficientes de correlación vectorial, el cual es un indicador de la similitud entre las condiciones en el tiempo, en este caso los valores más altos de coeficiente vectorial pertenecen los años 2015-2017 y

2020-2022. Para mayor confirmación de la representatividad de los años 2015, 2017 y 2022 se tiene coseno 2 el cual señala el compromiso que mejor representa la información de incidencias de los distintos grupos morfológicos, el año de menor representación de información en el compromiso es el 2019. También se evalúan los pesos de las k-tablas que confirman los años de mayor y menor aporte. Por último, se tiene que en la interestructura los siete primeros componentes explican el 87% de la varianza y en el compromiso los tres primeros componentes explican un 95%, véase apéndice D.

Figura 4.15 Proyección de los ejes principales del ACP sobre el PTA

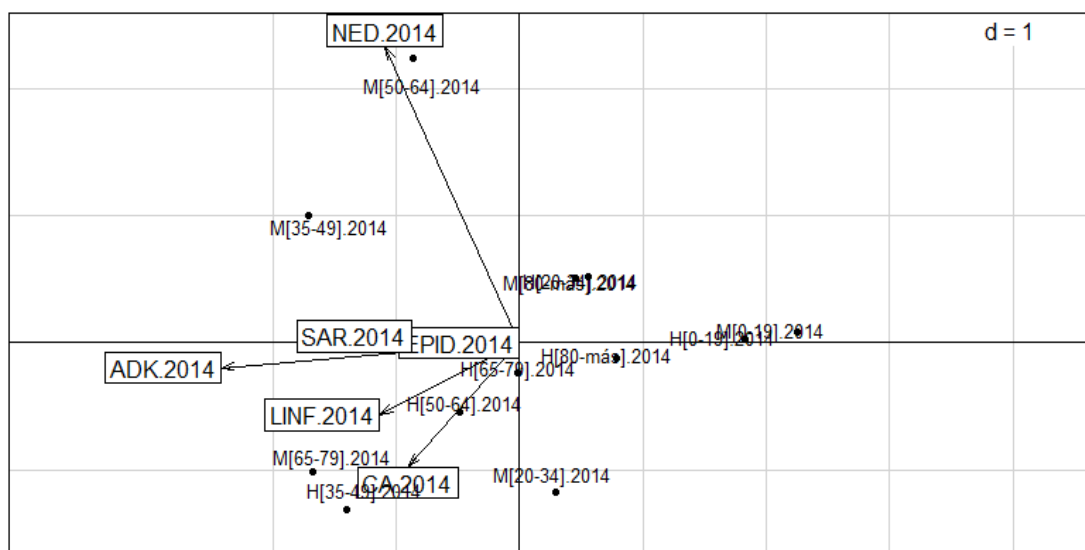


Fuente: Elaboración propia mediante paquete Ade4 - RStudio

La figura 4.15 se la realizó mediante el código `kplot(pta1)`, en el cual se pueden visualizar las proyecciones de los ejes principales, a continuación, se realiza un análisis de los años que aportan con mayor información a la matriz consenso (pesos altos), de este grupo se exceptúa al año 2019 el cuál se encuentra en el apéndice E.

La figura 4.16 muestra una débil correlación entre el grupo morfológico neuroendocrino (NED) con el resto de los grupos que están mayormente correlacionados. También se puede apreciar que las mujeres en los rangos de edad entre 35-49 y 50-64 se encuentra por encima del valor promedio de NED y por debajo M[20-34] y H[80 y más]. Por otro lado, los grupos etarios como H[65-70], H[50-64], M[65-79] y H[35-49] poseen valores por encima del promedio de adenocarcinoma, sarcoma, linfoma y carcinoma, por debajo del promedio se encuentra H[20-34], M[80 y más] y ambos sexos de las edades de 0-19 años.

Figura 4.16 Proyección de los ejes principales año 2014

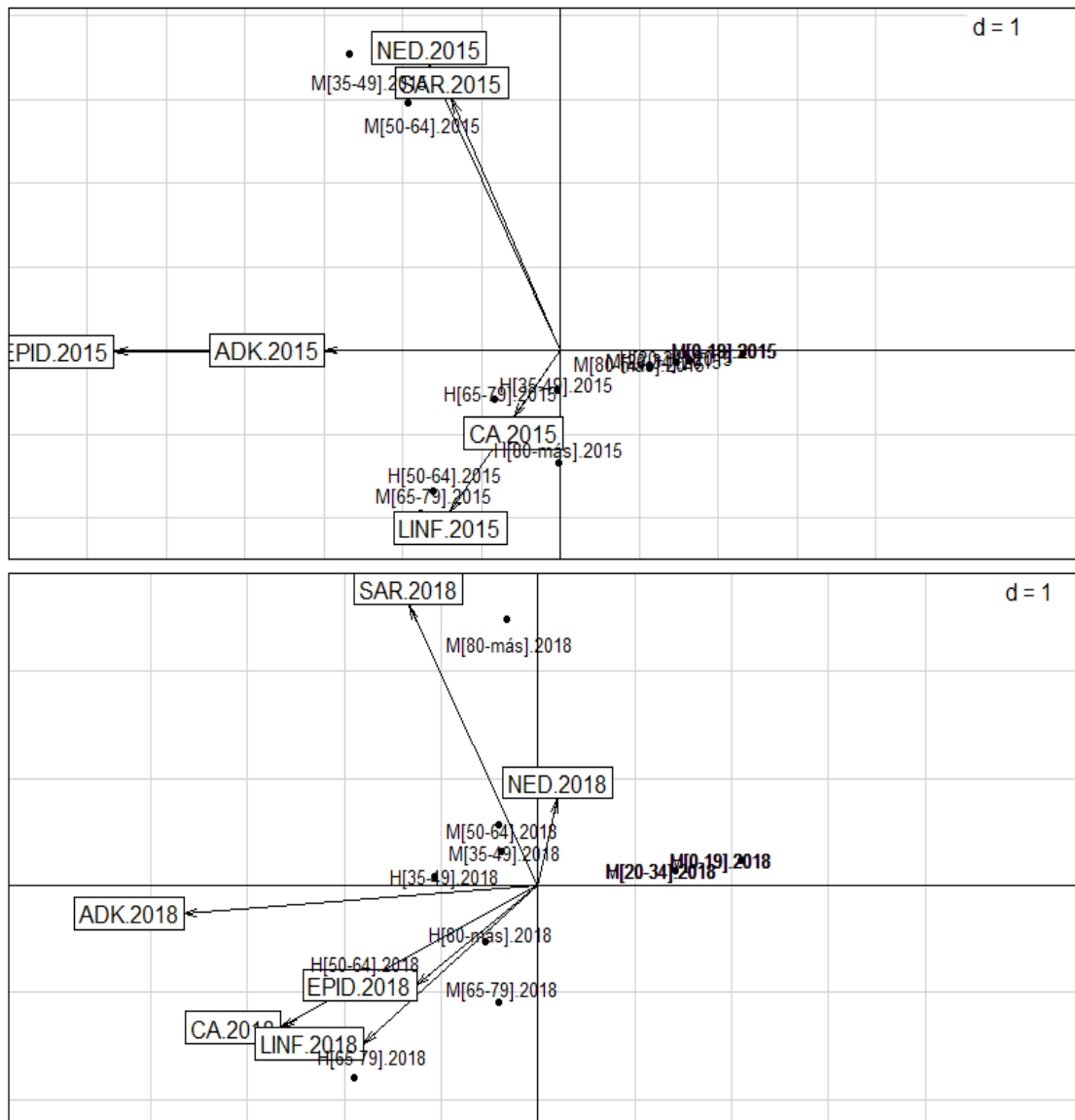


Fuente: Elaboración propia mediante paquete Ade4 - RStudio

A continuación, tenemos los años 2015 y 2018, en el primero los grupos etarios de mujeres de 35-49 y 50-64, se encuentra por encima del valor promedio de los grupos morfológicos SAR Y NED, similar sucede en el año 2018 al cual se le suma los hombres de 35-49 y las mujeres de ochenta y más años, este último en el 2015 tiene valores por debajo del promedio. En el primer año existen tres grupos bien definidos dado su alta correlación, mientras que en el segundo año hay dos grupos con variables directamente correlacionadas.

En ambos años los grupos etarios que poseen valores por encima de ADK, EPID, CA y LINF son los hombres de [35-49], [50-64], [80-más] y de ambos sexos [65-79] (figura 4.17).

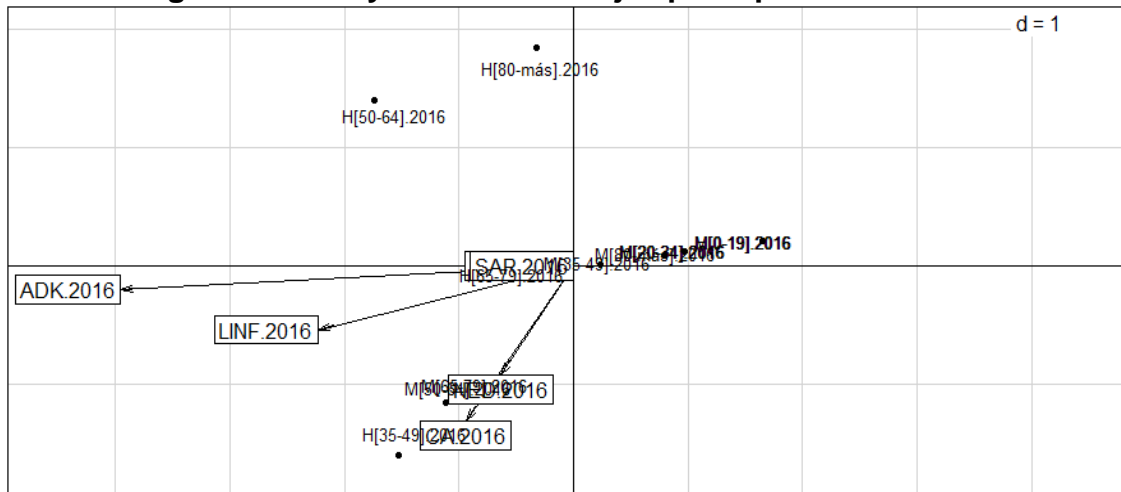
Figura 4.17 Proyección de los ejes principales año 2015 y 2018



Fuente: Elaboración propia mediante paquete Ade4 - RStudio

En el año 2016 el grupo CA y NED están altamente correlacionados, mientras que ADK se encuentra directamente correlacionada con LINF dado que el ángulo no es demasiado grande. Las mujeres en los rangos de edad 50-64 y 65-79 y hombre en edades de 35-49 años presentan valores por encima del promedio de ADK, LINF, CA Y NED; por el contrario los grupos etarios que están por detrás de dichos grupos morfológicos tienen valores por debajo del promedio. Este año SAR y EPID no presentan valores (figura 4.18).

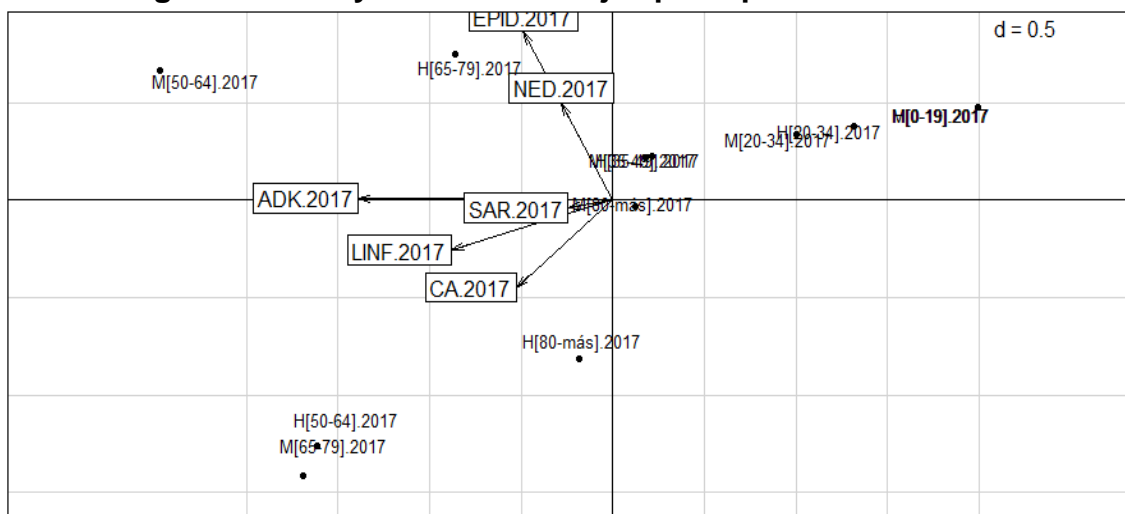
Figura 4.18 Proyección de los ejes principales año 2016



Fuente: Elaboración propia mediante paquete Ade4 - RStudio

Los grupos etarios del sexo masculino [50-64], [80-más] y del sexo femenino [65-79] tienen valores por encima del promedio para ADK, LINF CA y SAR. En cambio, M[50-64] y H[65-79] poseen valores por encima de EPID Y NED. El resto de los grupos etarios presentan valores por debajo del promedio como se observa en la figura 4.19.

Figura 4.19 Proyección de los ejes principales año 2017

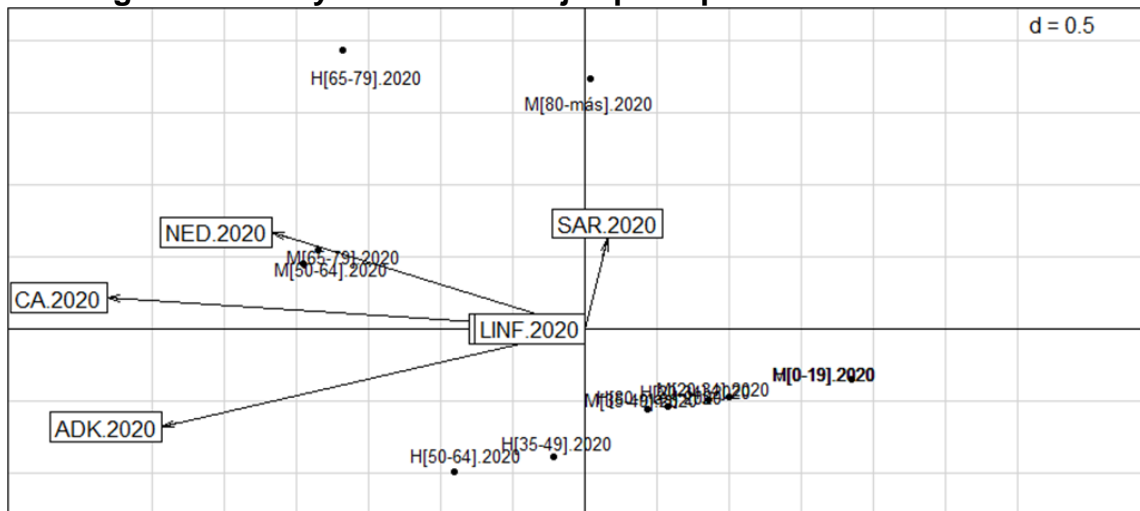


Fuente: Elaboración propia mediante paquete Ade4 - RStudio

En el año 2020 el grupo morfológico SAR tiene la particular de tener una correlación inversa a ADK, esto dado que el único grupo etario que cuenta con valores por encima del promedio es el de las mujeres de 80 y más. Los rangos de edad de mujeres [65-79], [50-64] y hombres [65-79] años poseen valores por encima de CA y NED, los grupos en el cuadrante opuesto son los que tienen valores menores al promedio de ambos grupos

morfológicos. Este año los hombres de [35-49] y [50-64] años son los que tienen valores por encima del promedio. Este año LINF y EPID no presentan valores (figura 4.20).

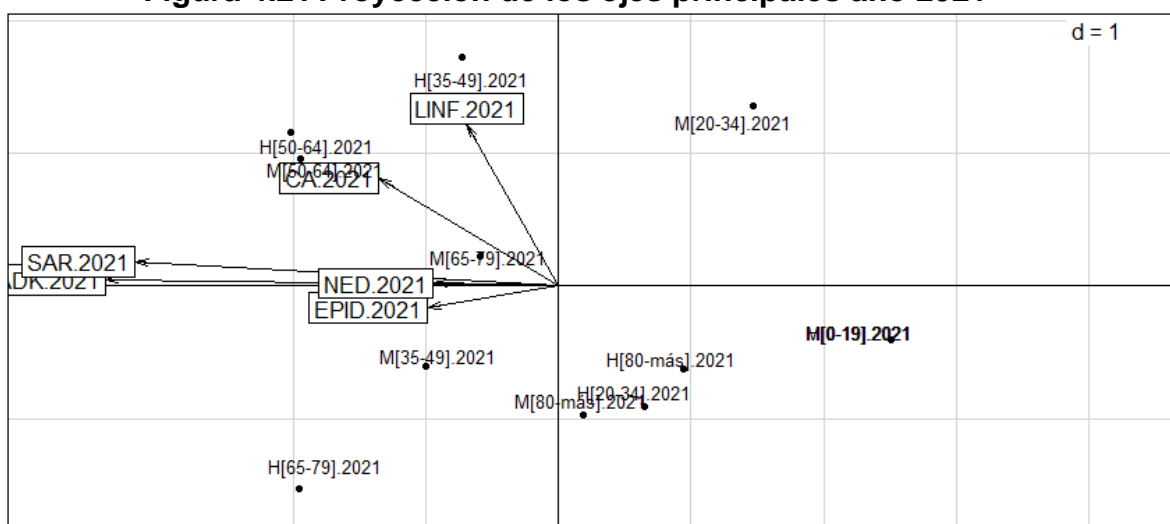
Figura 4.20 Proyección de los ejes principales año 2020



Fuente: Elaboración propia mediante paquete Ade4 - RStudio

En la figura del año 2021 (4.21) se aprecia que ambos sexos en edad de 50 a 64 años y las mujeres de 65 a 79 años poseen valores por encima del promedio de CA. Para LINF por encima del promedio se encuentra H[35-49] y por debajo del promedio se encuentra H[80 y más]. Aunque ADK es el que presenta mayor variabilidad, esta vez se le une el SAR el cual a diferencia de otros años presenta el mayor incremento de casos con ese grupo de morfologías. Los grupos de edades por encima del valor promedio son M[35-49] Y H[65-79]. Por debajo de todos los promedios se encuentran ambos sexos del grupo etario [0-19] años.

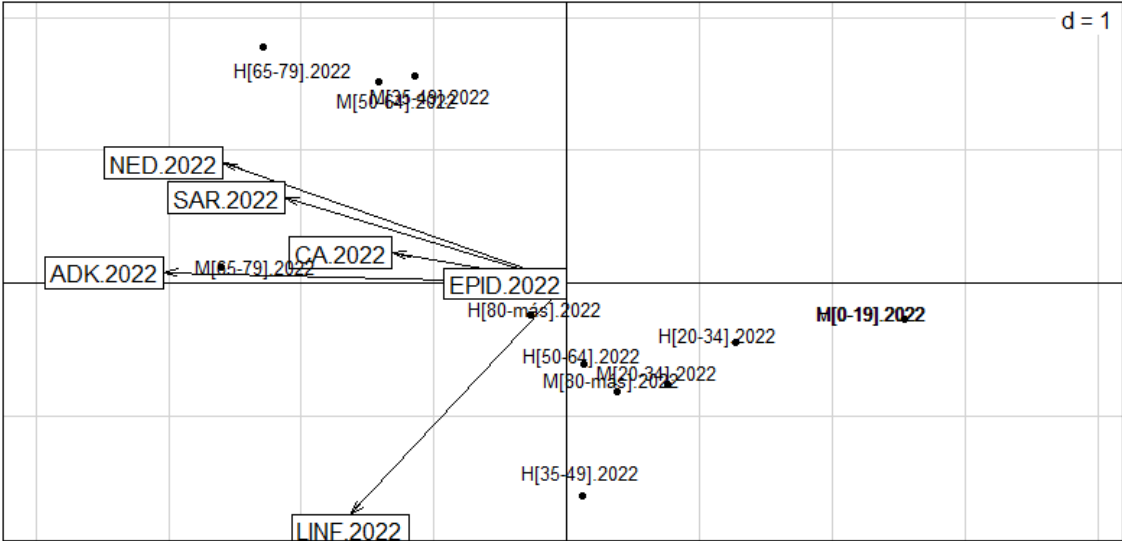
Figura 4.21 Proyección de los ejes principales año 2021



Fuente: Elaboración propia mediante paquete Ade4 - RStudio

En la figura 4.22 los grupos etarios que presentan valores por encima del promedio de la agrupación morfológica ADK, NED, SAR, CA son M[35-49], M[50-64], H-M[65-79], mientras que por debajo del valor promedio se encuentran de ambos sexos menores de 20 años, H[50-64] y M[80-más]. EPID no presenta valores, sin embargo es el año que mayor información aporta dado la variedad de información que tiene de los otros grupos morfológicos.

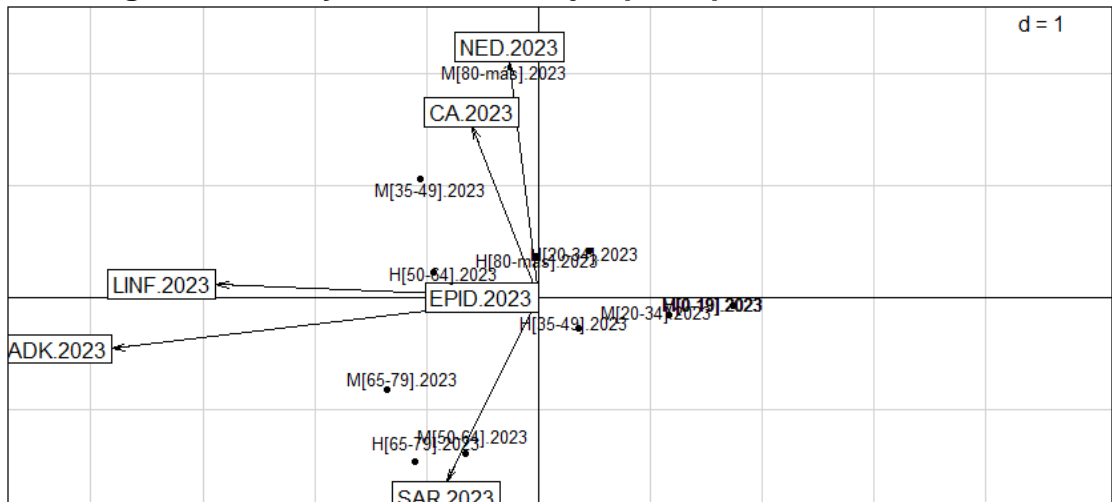
Figura 4.22 Proyección de los ejes principales año 2022



Fuente: Elaboración propia mediante paquete Ade4 - RStudio

Figura 4.23, por encima del valor promedio del grupo morfológico CA y NED, se encuentran las agrupaciones etarias de mujeres de [35-49], [80 y más] y hombres de [80 y más], y aunque el grupo predominante es el adenocarcinoma los rangos etarios como [65-79] se ambos sexos y M[50-64] años ven cercanos a SAR debido a que presentan valores por encima del promedio.

Figura 4.23 Proyección de los ejes principales año 2023



Fuente: Elaboración propia mediante paquete Ade4 - RStudio

CAPÍTULO 5

5. CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

La aplicación de los métodos de análisis multivariantes en el área de salud permiten considerar múltiples variables de manera simultánea, proporcionando un análisis detallado y por tanto más completo en comparación con los métodos tradicionales, dado que las enfermedades presentan múltiples causas y el pronóstico se determina, usualmente, por un gran número de factores, es lo que se aprecia con la investigación de literatura realizada, por lo que, se puede considerar necesario el uso de métodos multivariantes como HJ-Biplot o el método XSTATIS.

A través del análisis realizado, se han obtenido resultados que permiten una comprensión más profunda de los factores epidemiológicos asociados a esta enfermedad catastrófica como lo es el cáncer de estómago, además de su distribución en la muestra estudiada, por lo que mediante el uso del método HJ-Biplot se obtuvo dos gráficos. Del primero se puede concluir que el antro es el lugar topográfico donde más se encuentra el tumor maligno, siendo el grupo morfológico predominante el adenocarcinoma y afectando mayormente a los hombres comprobando el comportamiento similar que mantiene este cáncer al igual que en otros países. El segundo gráfico no solo reconfirma que el sexo masculino presenta mayor incidencia del cáncer gástrico, sino que señala las edades más propensas a diagnosticar esta enfermedad que es en los rangos de entre 50-64 y 65-79 años de edad, y en las mujeres en la edad de 35-49 años su incidencia es mayor al de los hombres por lo que también presenta mayor mortalidad en ese rango de edad, además muestra que los pacientes que llegaron por primera vez a la institución oncológica en el año 2021 y 2022 a la fecha de corte gran parte su estado vital es fallecido.

El análisis triádico parcial aporta con el análisis de los grupos morfológicos seleccionados versus los grupos etarios y su respectivo sexo a lo largo de varios años evidenciando los cambios y relaciones que presentan en el tiempo. Dado que el grupo morfológico adenocarcinoma (ADK) es complejo el análisis de las demás variables, no obstante, lo que se ha logrado visualizar es un leve crecimiento en los años 2021 al 2023 de grupos

morfológicos como neuroendocrino, sarcoma y linfoma, proporcionando insights sobre la evolución de patrones en la muestra estudiada.

5.2. Recomendaciones

Se recomienda incluir más años si se quiere apreciar mejor la evolución del cáncer de estómago en el hospital SOLCA Guayaquil. Además de incluir variables socioeconómicas de ser posible para apreciar un poco el estado financiero de los pacientes que llegan a la institución oncológica.

La aplicación de otra metodología multivariante es una recomendación a tener en cuenta, ya que cabe la posibilidad de que pueda explicar mejor la información de lo que se obtuvo con el análisis triádico parcial.

6. Referencias

- ACS. (2017, diciembre 14). *¿Qué es el cáncer de estómago?* American Cancer Society.
<https://www.cancer.org/es/cancer/tipos/cancer-de-estomago/acerca/que-es-el-cancer-de-estomago.html>
- Arenaza, D. P. de. (2022). Análisis multivariado. *Acta Gastroenterológica Latinoamericana*, 52(2), Article 2. <https://doi.org/10.52787/agl.v52i2.206>
- Barboza Besada, E. (1996). Cáncer de estómago. *Revista Medica Herediana*, 7(2), 87-100.
- Board, W. C. of T. E. (2019). *Digestive System Tumours* (5.^a ed., Vol. 1).
<https://publications.iarc.fr/Book-And-Report-Series/Who-Classification-Of-Tumours/Digestive-System-Tumours-2019>
- Caspers, H. (1981). LOUIS LEGENDRE et PAUL LEGENDRE: *Écologie numérique*. Tom 1: Le traitement multiple des données écologiques; Tom 2: La structure des données écologiques; = Collection d'écologie 12 et 13. – 38 fig., 29 tab., XIV/197 pp. +58 fig., 24 tab., XIII/247 pp., 326 ref. Paris, New York, Barcelone, Milan: Masson et Le presses de l'université du Québec 1979. *Internationale Revue Der Gesamten Hydrobiologie Und Hydrographie*, 66(5), 775-776.
<https://doi.org/10.1002/iroh.19810660515>
- Cedeño, R. G. M., Ballesteros, P. E. M., Sani, V. P., & Chica, G. N. R. (2021). Caracterización clínica, histológica y endoscópica del cáncer gástrico en el Hospital de Especialidades Dr. Abel Gilbert Pontón, Ecuador. *Revista colombiana de Gastroenterología*, 36(2), Article 2. <https://doi.org/10.22516/25007440.558>
- Chao, J. (2022, agosto 30). *Inmunoterapia en cáncer de estómago* [Health]. Cancer Research Institute. <https://www.cancerresearch.org/es/cancer-types/stomach-cancer>
- Cisneros, J. T. C., Babici, V. R., Guerrero, C. A. R., Luis, J., & Villardón, V. (2020). *Análisis multivariado HJ-Biplot de la ocurrencia de Helicobacter pylori como*

riesgo para cáncer gástrico, en la ciudadela el Cristo de Consuelo, Milagro Ecuador. 60(2), 116-123.

- Cubilla-Montilla, M., Nieto-Librero, A. B., Galindo-Villardón, M. P., & Torres-Cubilla, C. A. (2021). Sparse HJ Biplot: A New Methodology via Elastic Net. *Mathematics*, 9(11), Article 11. <https://doi.org/10.3390/math9111298>
- Espejo Romero, H., & Navarrete Siancas, J. (2003). Clasificación de los Adenocarcinomas de estómago. *Revista de Gastroenterología del Perú*, 23(3), 199-212.
- Ferlay, J., Ervik, M., Lam, F., Laversanne, M., Colombet, M., Mery, L., Piñeros, M., Znaor, A., Soerjomataram, I., & Bray, F. (s. f.). *Global Cancer Observatory: Cancer Today* [Health]. Global Cancer Observatory. Recuperado 9 de abril de 2024, de <https://gco.iarc.who.int/today/>
- Fritz, A., Percy, C., Jack, A., Shanmugaratnam, K., Sobin, L., Parkin, D. M., & Whelan, S. (2003). Clasificación Internacional de enfermedades para oncología. *Revista do Instituto de Medicina Tropical de São Paulo*, 45(4), 224-224. <https://doi.org/10.1590/S0036-46652003000400018>
- Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58(3), 453-467. <https://doi.org/10.2307/2334381>
- Galindo-Villardón, P. (1986). Una Alternativa de Representación Simultánea: HJ-BIPLLOT. *Qüestiió*, 10(1), 13-23. <https://diarium.usal.es/pgalindo/files/2012/07/0article-HJ-1986.pdf>. <https://orcid.org/0000-0001-6977-7545>
- Hernández, E. J. M. (2015). *Análisis Multivariante del Uso del Tiempo de la Población Colombiana Mediante el HJ Biplot y el Análisis Triádico Parcial* [Trabajo Fin de Máster, Universidad de Salamanca]. https://gredos.usal.es/bitstream/handle/10366/128218/TFM_MAADM_Medina_Hernandez_EdithJohana.pdf?sequence=4

- Hotelling, H. (1936). Simplified calculation of principal components. *Psychometrika*, 1(1), 27-35. <https://doi.org/10.1007/BF02287921>
- Iwu, C. D., & Iwu-Jaja, C. J. (2023, julio 6). Gastric Cancer Epidemiology: Current Trend and Future Direction. *Hygiene*, 3(3), 256-268.
- Jaffrenou, P.-A. (1978). *Sur l'analyse des familles finies de variables vectorielles: Bases algébriques et application à la description statistique*. éditeur non identifié.
- Lordick, F., Carneiro, F., Cascinu, S., Fleitas, T., Haustermans, K., Piessen, G., Vogel, A., & Smyth, E. C. (2022). Gastric cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up☆. *Annals of Oncology*, 33(10), 1005-1020. <https://doi.org/10.1016/j.annonc.2022.07.004>
- Narváez, M. G., Barzola, O. R., & Libroero, A. N. (2020). Análisis multivariante: Un recorrido por las técnicas de reducción de dimensiones. *Matemática*, 18(2), Article 2. <http://www.revistas.espol.edu.ec/index.php/matematica/article/view/747>
- Patino-Alonso, M. C., Recio-Rodríguez, J. I., Magdalena-Belio, J. F., Giné-Garriga, M., Martínez-Vizcaino, V., Fernández-Alonso, C., Arietaleanizbeaskoa, M. S., Galindo-Villardón, M. P., Gómez-Marcos, M. A., & García-Ortiz, L. (2015). Clustering of lifestyle characteristics and their association with cardio-metabolic health: The Lifestyles and Endothelial Dysfunction (EVIDENT) study. *British Journal of Nutrition*, 114(6), 943-951. <https://doi.org/10.1017/S0007114515002500>
- Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11), 559-572. <https://doi.org/10.1080/14786440109462720>
- Puente, J., & de Velasco, G. (2019, diciembre 16). ¿Qué es el cáncer y cómo se desarrolla? - SEOM: Sociedad Española de Oncología Médica © 2019 [SEOM].

Sociedad Española de Oncología Médica. <https://seom.org/informacion-sobre-el-cancer/que-es-el-cancer-y-como-se-desarrolla>

Real-Cotto, J. J., Elizabeth, J.-F. L., Pablo, T.-C. J., Rafael, P.-P. G., Carolina, P.-V. L., & Andrés, R.-R. R. (2023). Mortalidad del Cáncer de Estómago en Guayaquil, según tendencias históricas. *Health and Medical Sciences*, 8(1), 51-55.

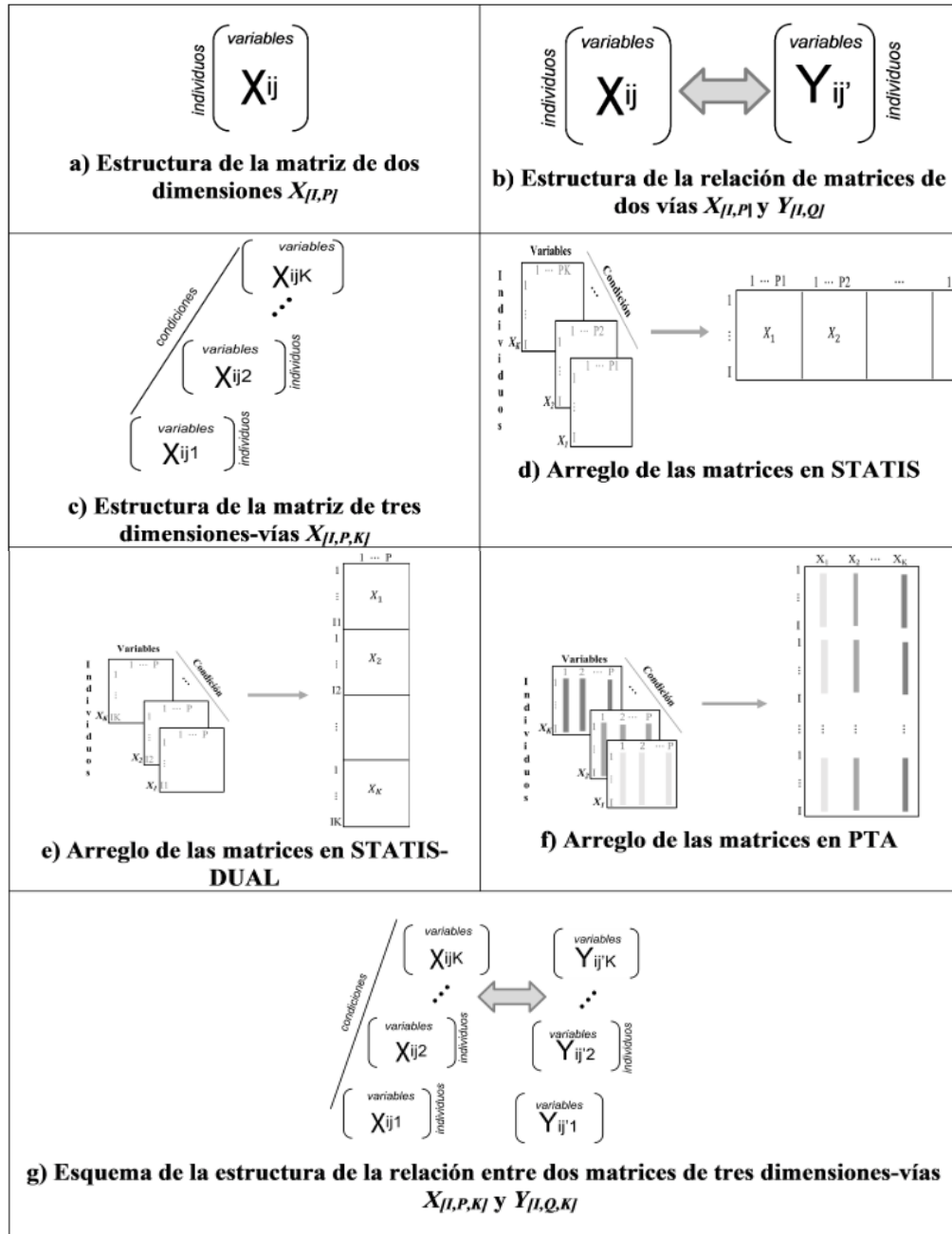
Thioulouse, J., Chessel, D., Dole´dec, S., & Olivier, J.-M. (1997). ADE-4: A multivariate analysis and graphical display software. *Statistics and Computing*, 7(1), 75-83.
<https://doi.org/10.1023/A:1018513530268>

Vicente Villardon, J. L. (2010). *MULTBIPLOT: A package for multivariate analysis using biplots*. <http://biplot.usal.es/ClassicalBiplot/index.html>

7. Apéndices y anexos

7.1. Apéndice A

Figura A1.24 Estructuras y Arreglos Multivariantes



Fuente: (Narvéez et al., 2020) Análisis multivariante: Un recorrido por las técnicas de reducción de dimensiones.

7.2. Apéndice B

Tabla B1.4 Valores propios y varianza explicados

Axis	Eigenvalue	Expl. Var.	Cummulative
Axis 1	8.785	66.536	66.536
Axis 2	1.401	10.611	77.147
Axis 3	1.07	8.103	85.25
Axis 4	0.886	6.713	91.962
Axis 5	0.454	3.439	95.401
Axis 6	0.392	2.969	98.37
Axis 7	0.171	1.293	99.663
Axis 8	0.044	0.337	100

Tabla B2.5 Contribuciones relativas del factor al elemento filas

Row	Axis 1	Axis 2	Axis 3
MC169	872	90	13
MC161	623	23	49
MC168	7	26	528
MC162	900	6	1
HC163	199	661	2
HC165	387	3	46
MC163	302	193	84
HC162	149	281	11
HC169	906	22	7
HC168	192	217	210
MC165	697	8	145
HC161	655	6	45
HC160	6	427	412
MC160	3	93	662
MC164	952	10	4
HC164	835	91	21
HC166	906	3	45
MC166	719	159	40

Column	Axis 1	Axis 2	Axis 3
Adenocarcinoma	838	136	10
Adenoescamoso	197	0	9
Carcinomas	922	3	5
Epidermoide	80	22	648
Indiferenciado	127	17	17
Linfomas	541	227	70
Neuroendócrinos	578	91	2
Sarcoma	558	165	154

7.3. Apéndice C

Tabla C1.6 Valores propios y varianza explicados

Axis	Eigenvalue	Expl. Var.	Cummulative
Axis 1	21.577	84.633	84.633
Axis 2	1.066	4.182	88.815
Axis 3	0.947	3.715	92.529
Axis 4	0.651	2.555	95.085
Axis 5	0.333	1.306	96.391
Axis 6	0.261	1.022	97.413
Axis 7	0.187	0.733	98.146
Axis 8	0.168	0.66	98.807
Axis 9	0.142	0.558	99.365
Axis 10	0.117	0.457	99.822
Axis 11	0.04	0.156	99.978
Axis 12	0.006	0.022	100

Contribuciones relativas del factor al elemento fila

Row	Axis 1	Axis 2	Axis 3
14F	949	5	22
14V	522	405	41
15F	860	1	4
15V	966	4	7
16F	844	40	6
16V	926	42	13
17F	945	1	5
17V	904	1	36
18F	612	83	67
18V	908	3	69
19V	520	60	133
19F	957	8	6
20F	577	32	126
20V	814	15	95
21V	116	84	254
21F	890	1	19
22V	223	42	391
22F	894	0	73
23F	592	180	0
23V	461	89	16

Contribuciones relativas del factor al elemento columna

Column	Axis 1	Axis 2	Axis 3
H[0-19]	3	498	12
H[20-34]	671	8	78
H[35-49]	888	27	34

H[50-64]	863	87	0
H[65-79]	932	6	28
H[80-más]	849	48	17
M[0-19]	13	51	10
M[20-34]	587	51	287
M[35-49]	897	70	0
M[50-64]	815	91	29
M[65-79]	932	5	12
M[80-más]	877	2	3

7.4. Apéndice D

Coeficientes de correlación vectorial RV

	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
2014	1	0,297	0,319	0,221	0,229	0,176	0,274	0,235	0,394	0,280
2015	0,297	1	0,134	0,482	0,296	0,288	0,183	0,291	0,389	0,139
2016	0,319	0,134	1	0,351	0,369	0,224	0,202	0,319	0,197	0,331
2017	0,221	0,482	0,351	1	0,120	0,270	0,368	0,257	0,368	0,225
2018	0,229	0,296	0,369	0,120	1	0,150	0,240	0,304	0,190	0,349
2019	0,176	0,288	0,224	0,270	0,150	1	0,271	0,277	0,201	0,099
2020	0,274	0,183	0,202	0,368	0,240	0,271	1	0,070	0,463	0,198
2021	0,235	0,291	0,319	0,257	0,304	0,277	0,070	1	0,163	0,266
2022	0,394	0,389	0,197	0,368	0,190	0,201	0,463	0,163	1	0,341
2023	0,280	0,139	0,331	0,225	0,349	0,099	0,198	0,266	0,341	1

Coseno 2

	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
	0,5931	0,6560	0,5538	0,6439	0,5456	0,5059	0,5500	0,5478	0,6594	0,5237

Pesos para las k-tablas

	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
	0,324	0,334	0,321	0,353	0,298	0,265	0,308	0,290	0,358	0,300

Valores propios para la inter-estructura

Eigenvalues	3,381	1,207	1,055	0,849	0,766	0,751	0,683	0,568	0,450	0,290	10
%Varianza	33,8%	12,1%	10,6%	8,5%	7,7%	7,5%	6,8%	5,7%	4,5%	2,9%	
%Varianza Acumulada	34%	46%	56%	65%	73%	80%	87%	93%	97%	100%	

Valores propios para el compromiso

Eigenvalues	16,001	1,8084	1,1671	0,538	0,3212	0,0725	19,908
%Varianza	80,4%	9,1%	5,9%	2,7%	1,6%	0,4%	
%Varianza Acumulada	80%	89%	95%	98%	100%		

7.5. Apéndice E

Figura E1.25 Proyección de los ejes principales año 2019

