**Informe Técnico Final**

Título del Proyecto: ***“*Aplicación de la minería de datos en la citómica y metabolómica para el diagnóstico de las enfermedades metabólicas y sus complicaciones usando R”**

**1. Grupo de Trabajo**

en el grupo de trabajo se incorporaron 2 Investigadores y 2 Estudiantes de Licenciatura

* Dra. Nicté Guadalupe Figueroa Vega (Responsable Técnico)
* Dr. Felipe Pérez Vargas (Colaborador)
* L.M. Carlos Daniel Pérez Sánchez (alumno del Programa de Verano UG 2021)
* QFB Erick Geovanni Galván López (Estudiante del Programa de Verano UG 2021)

**2. Antecedentes**

En los últimos años con el incremento en la incidencia de enfermedades metabólicas (obesidad, diabetes mellitus tipo 2, aterosclerosis, etc.) y sus complicaciones en nuestro entorno y dado el importante papel que juegan diversas moléculas y células inmunes en el desarrollo de éstas, el empleo de la citometría de flujo y espectrometría de masas se ha convertido en un instrumento valioso para la investigación básica aplicada, ya que permite identificar, cuantificar moléculas y células, que han ayudado a un mejor entendimiento y temprano diagnóstico de estas patologías. Ambas son unas poderosas técnicas microanalíticas por su especificidad y sensibilidad que permiten cumplir con los estándares requeridos por la investigación moderna. Los datos obtenidos por estas técnicas permiten utilizar la minería de datos, y a través de un análisis exploratorio se familiarizará con el Software R para cimentar el formalismo matemático. Además permitirá al integrar los datos, buscar patrones clínicos mediante los análisis predictivos y clustering jerárquico que nos sugieran la relación de alguna variable medida con dichas enfermedades para llegar a un diagnóstico temprano clínico de las micro- y macrocomplicaciones de la Diabetes Mellitus tipo 2. Para lo cuál se plantea utilizar K-mean, clustering y clustering jerárquico o árboles de decisión, análisis discriminante, PCA, máquina de soporte vectorial, con el software R.

1. ***Metas científicas:***

Manejo de R y métodos predictivos para establecer diagnósticos temprano del daño renal y alteraciones en el estado de ánimo, a través de la utilización de 3 bases de datos distintas procedentes de pacientes con diabetes mellitus tipo 2 y de mujeres en la menopausia, con diversas variables para los análisis correspondientes.

* Se estableció un método análisis del patrón químico en orina para determinar la huella química de los pacientes con DM2, con y sin enfermedad renal crónica y voluntarios sanos (más abajo se describen los metabolitos).
* Se analizó y correlacionó la expresión de RAGE y sus ligandos por citometría de flujo con los parámetros bioquímicos, metabólicos para establecer su asociación con el fenómeno inflamatorio y su implicación en el desarrollo de daño renal temprano, riesgo cardiovascular y sarcopenia.
* Se determinó la posible asociación de la molécula S100B con los síntomas emocionales de mujeres en la menopausia.
* Se adjunta resumen de los resultados de estos Proyectos de Investigación con los análisis estadísticos correspondientes.
* Con el presente Proyecto se fortaleció el CA: Metabolismo y Reproducción ya que se estableció la línea de investigación en metabolómica y citómica.

1. ***Colaboraciones externas:***

Se estableció colaboración con el Dr. Felipe Pérez Vargas del División de Ciencias de la Vida, Universidad de Guanajuato, campus Irapuato-Salamanca para llevar a cabo los análisis y establecimiento de la huella química.

1. ***Formación de Recursos Humanos de alta calidad***

* Se formaron a los estudiantes L.M. Carlos Daniel Pérez Sánchez (alumno del Programa de Verano UG 2021) ) y QFB Erick Geovanni Galván López (Estudiante del Programa de Verano UG 2021) quienes participaron en el manejo del software R para llevar a cabo los análisis correspondientes.

1. ***Entregables***
   * Presentación en Artículo *en extenso* del Trabajo: ***“Abordaje Metabolómico para la Determinación de Biomarcadores Tempranos de Nefropatía Diabética”***.

Presentación en Cartel y diapositivas del trabajo de Investigación: “**RAGE…..”** para el Congreso

* + Realización de un vídeo de cómo se utiliza R en la minería de datos

**Identificación de marcadores metabolómicos en el desarrollo de nefropatía en pacientes con diabetes mellitus tipo 2**

Carlos Daniel Pérez Sánchez1, Erick Geovanni Galván López1, Felipe Pérez Vargas2, Nicté Figueroa-Vega1

**1**Departamento de Ciencias Médicas, Universidad de Guanajuato, campus León. 20 de Enero No. 929, Col. Obregón, C.P. 37320. León, Gto., México.

2 División de Ciencias de la Vida, Universidad de Guanajuato, campus Irapuato-Salamanca.

ExHacienda del Copal, C.P. 36500*.* Guanajuato, Gto.

*Palabras clave: Diabetes mellitus, Nefropatía diabética, microalbuminuria, metabolómica*

**Introducción:** La diabetes mellitus tipo 2 (DM2), es una enfermedad crónico-degenerativa causada por un defecto en la captación de la insulina por las células (resistencia a la insulina). Aunado a esto, para el 2014 la prevalencia mundial de la DM2 aumentó hasta un 8.5% con respecto al año 1980 (4.7%)1. Además, en nuestro país es causante del 9.25% del total de mortalidad registrada2. Por otra parte, la DM2 ha sido asociada al incremento en el riesgo de desarrollo de complicaciones3. Las complicaciones micro- y macrovasculares tienen alta mortalidad y morbilidad en nuestra población, lo que hacen urgentes la búsqueda e investigación de perspectivas novedosas para identificar precozmente el deterioro micro- y macrovascular. A este respecto, de las personas que cursan DM2 cerca del 40% desarrollarán enfermedad renal (Nefropatía Diabética)4. Uno de los mecanismos para el desarrollo de la nefropatía diabética (ND) deriva principalmente de la hiperglucemia crónica que es característica de pacientes con DM2 generando cambios funcionales y estructurales en el riñón, tales como: engrosamiento de la membrana basal glomerular, incremento en la formación de la matriz extracelular y pérdida de podocitos con alteración de la tasa de filtración glomerular (TFG), etc4,5. La creatinina sérica y urinaria, la excreción urinaria de albúmina, la depuración de creatinina y la estimación de la filtración glomerular son frecuentemente empleadas en el laboratorio clínico de rutina para evaluar la función renal; sin embargo, es conocido que estas pruebas presentan desventajas que limitan su sensibilidad clínica en estadios tempranos, indicando la importancia de acciones preventivas y de diagnóstico oportuno para ofrecer mejor calidad de vida a los pacientes, por lo que la ND se detecta hasta etapas avanzadas cuando el beneficio del tratamiento es limitado. Debido a lo anterior, son necesarios marcadores oportunos de tamizaje de la DM para evitar la progresión a ND a estadios avanzados, así como la implementación de un tratamiento adecuado. Con respecto a lo anterior, se han establecido nuevas estrategias analíticas como la metabolómica, la cual ha surgido como una herramienta que puede tener un gran impacto clínico en la salud, ya que a partir de un simple análisis y un costo relativamente bajo permite la identificación de metabolitos que pueden ser característicos del fenotipo de una enfermedad. En este contexto, con el desarrollo de la metabolómica, se ha comenzado a estudiar la huella química generada en diferentes enfermedades para la identificación de nuevos biomarcadores de daño temprano6–8, por lo que se ha propuesto el uso de esta herramienta con el fin de identificar biomarcadores que aporten un valor predictivo desde las primeras etapas y de esta manera lograr un diagnóstico oportuno y no invasivo de la enfermedad.

**Objetivo:** Identificar biomarcadores metabolómicos en personas diagnosticadas con DM2 que puedan servir como método de tamizaje para evitar la progresión de la ND a estadios avanzados.

**Materiales y métodos:**

**Selección de los pacientes.**

Se incluyeron pacientes de con una edad de 20-50 años, diagnosticados con DM2 sin daño renal (n=19), con DM2 y daño renal (n=9) y sujetos clínicamente sanos (n=21), a los cuales se les registraron datos antropométricos (Talla, Peso, IMC, entre otros) e historia clínica. Se tomó una muestra de orina y de sangre por la mañana con un ayuno de 8 horas. Se obtuvo el plasma mediante centrifugación de las muestras sanguíneas para la determinación de glucosa, hemoglobina glucosilada mediante métodos espectrofotométricos. Por otro lado, en la muestra de orina fue determinada la creatinina, TFG. El análisis del perfil metabolómico en orina se realizó mediante una Nariz Electrónica (Alpha MOS HERACLES Flash Gas Chromatography Electronic Nose) bajo las siguientes condiciones: Volumen de inyección: 500 µl, Temperatura de inyector: 200°C, Flujo constante de hidrógeno de 1 ml/min y temperatura del horno: 50°C, 30 segundos, aumentando 10°C/segundos, hasta alcanzar una temperatura máxima de 280°C. Posteriormente, se determinó la identidad de los metabolitos con base a los índices de Kovats generados por el software del equipo, y con los datos obtenidos: tiempo de retención (TR) y Área Bajo la Curva (AUC) se realizaron análisis estadísticos multivariados: Análisis de Componentes Principales (PCA) y Partial Least Squares-Discriminant Analysis (PLS-DA) en el programa R.

**Resultados:**

Las características de los voluntarios son mostradas en la Tabla 1.

Tabla

Descripción generada automáticamente

Tabla

Descripción generada automáticamente

**Tabla 1.** Diabetes y daño renal. Los valores expresan la medias ± la desviacion estandar. Se compararon los 3 grupos mediante un analisis de ANOVA. Los metabolitos con con una significancia de p<0.05 que se encontraron fueron x2.propanol, butane.2.3.dione, x1.propanol.2.methyl,X2.methylbutane,x2.methylpropanal,butane.2.3.dione,methyl.isoburate.1.

Con una prueba de Tuckey se observo los siguientes metabolitos muestran una diferencia significativa entre los grupos de DM con DR y Dm sin DR: Propenal, x2.propanol, methil.isobutyrato, x2.4 dimethyl.1.3.dioxane

Mediante el análisisPCA se observó una separación entre el grupo de pacientes con DM y sujetos sanos, explicando el 31.6% de la variabilidad de los datos (Figura 1).

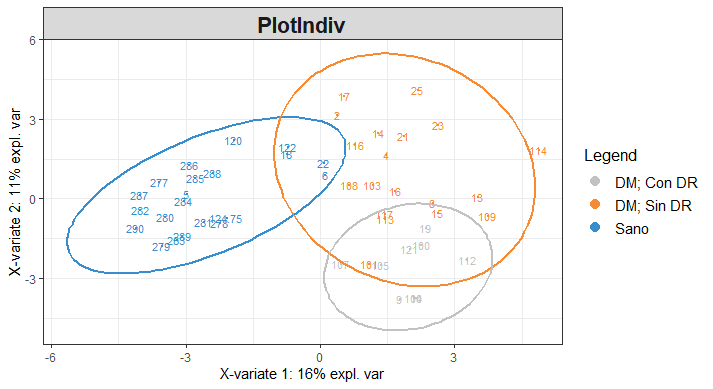
**Figura 1.** PCA daño renal

Gráfico, Gráfico de dispersión

Descripción generada automáticamente

Por otro lado, el PLS-DA mostró que los 3 grupos (Sanos, DM2 con y sin daño renal) pueden distinguirse claramente, explicando el 27% de su variabilidad (Figura 2).

**Figura 2.** PLS-DA daño renal



Además, se logró identificar al propenal, 2-metilpropanal, 2-propanol, alcohol propílico, metil-isobutirato, 2-metilbutano y butan-2,3-diona como los metabolitos más relevantes que distinguía al grupo de los DM2 (fig 3) con respecto al grupo de sujetos sanos, 3 de ellos (2-propanol, alcohol propílico y butan-2,3-diona) se encontraban involucrados en vías metabólicas como: metabolismo de butanoato, metabolismo de glicerolípidos y metabolismo del propanoato.

En relación a aquellos pacientes que padecen DM2 y mostraron daño renal temprano, se observó que los metabolitos

**Discusiones y Conclusiones:**

De acuerdo con los datos obtenidos, el alcohol propílico, 2-propanol y butan-2,3-diona (diacetil) se encuentran asociados a vías metabólicas como: metabolismo de glicerolípidos, metabolismo del propanoato y metabolismo del butanoato, respectivamente. El alcohol propílico se genera a partir del glicerol, que es el principal producto de degradación de los triacilglicéridos (TAGs). A este respecto, la diabetes mellitus se caracteriza por una deposición de lípidos que lleva a una excesiva e incompleta degradación de éstos9, por lo que este proceso puede llevar a la acumulación del alcohol propílico en la orina. Resultados similares fueron mostrados por Van y col. (2012), encontrando metabolitos que pertenecían al grupo de acilcarnitinas y acilglinas que se asocian al metabolismo de glicerolípidos10

Por otro lado, en la DM la glucólisis se encuentra casi inactiva y el organismo recurre a la producción de energía a partir de la activación de otras vías. Por ejemplo, a partir del metabolismo del propanoato y el metabolismo del butanoato, generando la acumulación de cetoácidos con la subsecuente generación de 2-propanol derivado de acetona en el metabolismo del propanoato y la producción de diacetil de manera indirecta durante el metabolismo del butanoato. Algunos estudios han mostrado que en pacientes con ND se encuentran metabolitos como: 3-hidroxipropionato y el ácido 3-metiladipico que se participan en el metabolismo del propanoato y metabolismo de ácidos grasos, respectivamente11

Además, en nuestro estudio se encontraron otros metabolitos como: propenal, también llamado acroleína, producto de peroxidación lipídica, y asociado con la DM y ND13,14. De la misma manera, se detectó el 2-metil propenal (isobutaraldehído) el cual se encuentra incrementado en pacientes diabéticos15. Asimismo, el metil isobutirato estuvo incrementado en el grupo de DM2, el cual es un metabolito que participa en la producción de glucosa a partir de aminoácidos no glucosídicos (Gluconeogénesis)16. Finalmente, el metil-butano también se detectó y puede explicarse como un posible producto de peroxidación lipídica, sin embargo, no hay suficientes estudios que sustenten esta hipótesis.

Finalmente. algunos de estos metabolitos se han reportado presentes en la enfermedad renal; por ejemplo; el propenal se ha encontrado presente en pacientes con enfermedad renal, por lo tanto, este estudio abre una ventana importante para poder detectar y determinar este tipo de metabolitos en esta enfermedad y realizar una metabolómica dirigida.

Este estudio logró identificar metabolitos que se encuentran asociados con DM con enfermedad renal, y de manera más importante la identificación de vías metabólicas que están siendo alteradas.

**Bibliografía:**

1. Mathers, C. D. & Loncar, D. Projections of Global Mortality and Burden of Disease from 2002 to 2030. *PLoS Med.* **3,** e442 (2006).

2. GBD Compare | IHME Viz Hub. Available at: https://vizhub.healthdata.org/gbd-compare/. (Accessed: 18th March 2018)

3. OMS | Diabetes. *WHO* (2017).

4. General, H. & Gonz, M. G. Nefropatía diabética. **5,** (2002).

5. Fundación Mexicana del Riñón A.C. Available at: http://www.fundrenal.org.mx/erc.html. (Accessed: 18th March 2018)

6. Hocher, B. & Adamski, J. Metabolomics for clinical use and research in chronic kidney disease. *Nat. Rev. Nephrol.* **13,** 269–284 (2017).

7. Bouatra, S. *et al.* The Human Urine Metabolome. *PLoS One* **8,** (2013).

8. Su, L. J. *et al.* The Use of Metabolomics in Population-Based Research. *Adv. Nutr. An Int. Rev. J.* **5,** 785–788 (2014).

9. Mihalik, S. J. Increased Levels of Plasma Acylcarnitines in Obesity and Type 2 Diabetes and Identification of a Marker of Glucolipotoxicity. **18,** 1695–1700 (2010).

10. van der Kloet, F. M. *et al.* Discovery of early-stage biomarkers for diabetic kidney disease using ms-based metabolomics (FinnDiane study). *Metabolomics* **8,** 109–119 (2012).

11. Sharma, K. *et al.* Metabolomics Reveals Signature of Mitochondrial Dysfunction in Diabetic Kidney Disease. *J. Am. Soc. Nephrol.* **24,** 1901–1912 (2013).

12. Pena, M. J. *et al.* Urine and plasma metabolites predict the development of diabetic nephropathy in individuals with Type 2 diabetes mellitus. *Diabet. Med.* **31,** 1138–1147 (2014).

13. Zhang, J. *et al.* Metabonomics research of diabetic nephropathy and type 2 diabetes mellitus based on UPLC-oaTOF-MS system. *Anal. Chim. Acta* **650,** 16–22 (2009).

14. Zhu, C., Liang, Q. L., Hu, P., Wang, Y. M. & Luo, G. A. Phospholipidomic identification of potential plasma biomarkers associated with type 2 diabetes mellitus and diabetic nephropathy. *Talanta* **85,** 1711–1720 (2011).

15. Shokry, E., de Oliveira, A. E., Avelino, M. A. G., de Deus, M. M. & Filho, N. R. A. Earwax: A neglected body secretion or a step ahead in clinical diagnosis? A pilot study. *J. Proteomics* **159,** 92–101 (2017).

16. Willems, E. *et al.* Embryonic Protein Undernutrition by Albumen Removal Programs the Hepatic Amino Acid and Glucose Metabolism during the Perinatal Period in an Avian Model. *PLoS One* **9,** (2014).

**S100A12, S100B y HMGB-1) en monocitos y neutrófilos y su aplicación en la evaluación de las complicaciones en la diabetes mellitus tipo 2**

Hay un gran interés en el papel de la inflamación en las enfermedades cardiovasculares, y los marcadores de inflamación que servirían como predictores. Recientemente se ha encontrado que los RAGEs y varios de sus ligandos pudieran actuar como perspectivas interesantes para evaluar el riesgo cardiovascular y diseñar estrategias de tratamiento preventivas (8).

Existen evidencias abundantes de la participación de los RAGE y sus formas solubles (sRAGE y esRAGE) en el daño cardiovascular, pero no hay información completa de su posible valor predictivo (8). Por otra parte, los RAGES tienen efecto en los transportadores tipo cassette de enlace a ATP (ABC), que cuando se alteran experimentalmente aceleran el daño cardiovascular, pero hay información limitada sobre este proceso en el humano (8)(9)(10).

El receptor de AGEs, RAGE ha surgido recientemente como una proteína importante en el proceso inflamatorio sistémico, pues se une además de AGEs,a otras moléculas: HMGB1, S100A y S100B que favorecen la inflamación y daño endotelial (8), e interactúan con transportadores claves en la regulación del transporte inverso del colesterol: ABCA1 y ABCG1 disminuyendo su expresión en los macrófagos. Los conocimientos sobre inflamación y deterioro vascular en la diabetes son incompletos. Un tema prioritario sería esclarecer el papel de la inflamación de bajo grado en estos procesos y en la predicción de daño cardiovascular. A este respecto, se han propuesto nuevos mecanismos de daño vascular durante la DM2 que no han sido abordados en profundidad, como la interacción RAGE-HMGB1-S100B (8) que podrían ofrecer nuevas perspectivas con propósitos predictivos. Por lo tanto, en este trabajo, se evaluaron los cambios de las moléculas RAGE-ligandos en distintas etapas de la historia natural de la DM2, para proponer nuevos procedimientos de diagnóstico temprano, tanto del daño cardiovascular como de la progresión de la enfermedad renal.

**Resultados**

Se seleccionaron adultos de ambos sexos, de 20 a 55 años, de centros de salud, centros de atención primaria y plazas de la ciudadanía de la ciudad de León, Guanajuato.

Se aplicaron las pruebas normalidad de Kolmogorov-Smirnov para identificar las variables paramétricas y no paramétricas. Las variables que sigan una distribución normal se presentaron como media y error estandar (EE), las variables que resulten no paramétricas se presentaron como mediana y rangos intercuartílicos. Se realizaron xxxxx con el paquetee R y se utilizaró un nivel de significancia de 0.05.

La **Tabla 1** muestra las características clínico-patológicas de los participantes evaluados. Se observaron diferencias significativas en el IMC, cifras de tensión arterial media, concentraciones de glucosa y triglicéridos, siendo mayores en los sujetos con DM2 que en los del grupo control.

Tabla

Descripción generada automáticamente

Tabla

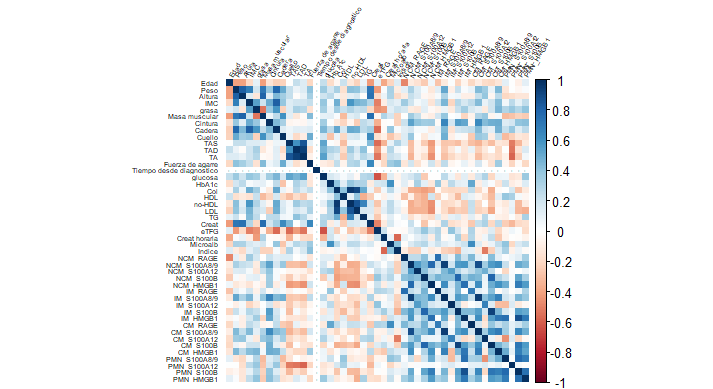
Descripción generada automáticamente

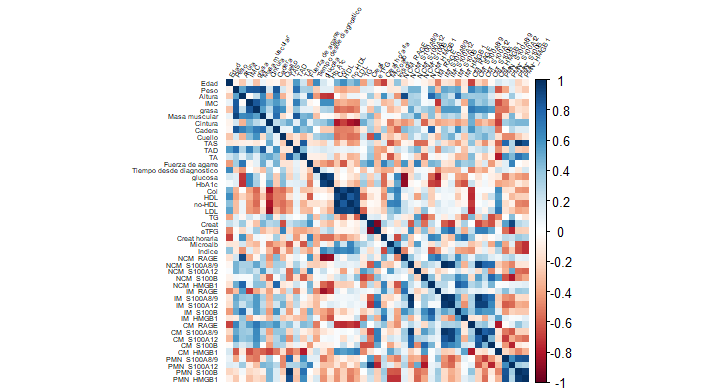
**Tabla 1 Caracteristicas clínicas de** los 3 grupos estudiados (sujetos sano, pacientes con DM sin DR y pacientes con DM y DR) . Mediante un analisis de ANOVA se encontró un nivel de significancia p<0.05 en;Edad,Peso,(%)Grasa,Cintura,Cadera,Glucosa,HbA1c,Trigliceridos, creatinica,index,

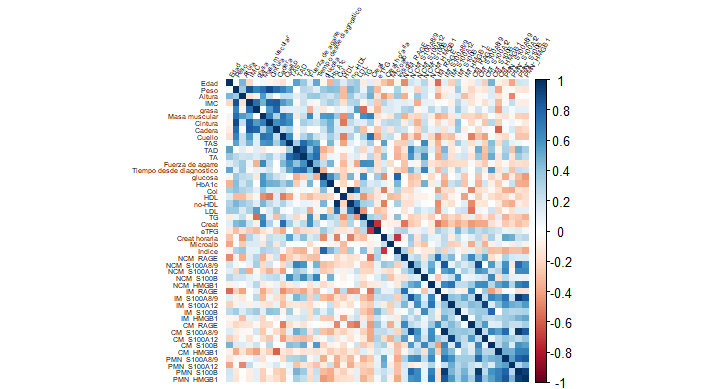
Tabla

Descripción generada automáticamente

**Tabla 2.** Se sometió al mismo procedimiento los grupos de sarcopenia encontrándose que las variables con un nivel de significancia p<0.05 son; Fuerza de agarre, creatinina

a)

b)

c)

**Figura 1.** Heatmaps obtenenidos a través de los coeficientes de correlación de Pearson, de manera grafica es posible observar el grado de correlación entre la variables siendo blancas aquellas donde no hay correlación, en azul las correlaciones positivas y en rojo las correlaciones negativas. a) Sujetos sanos, b)Pacientes con DM sin DR c)Pacientes con DM y DR

Tabla

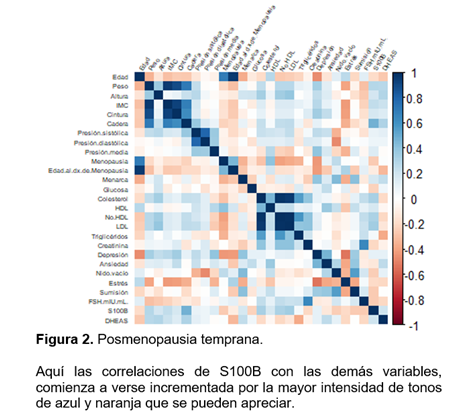
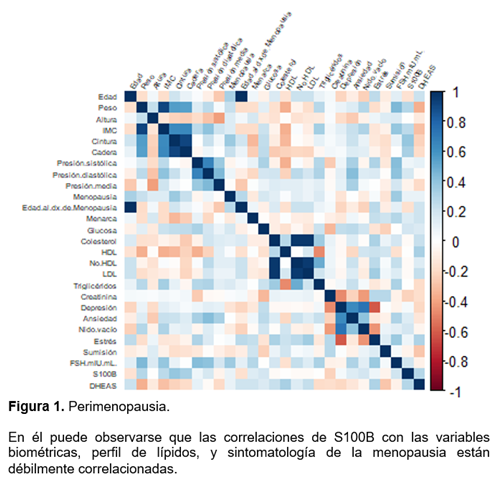
Descripción generada automáticamente

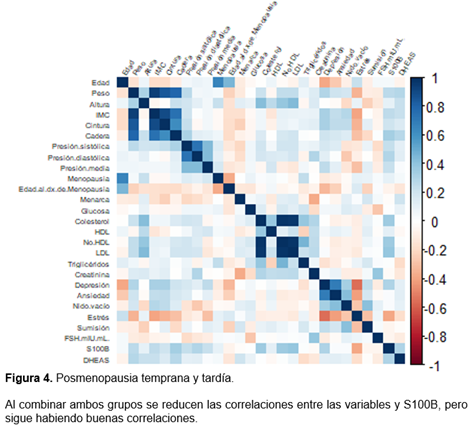
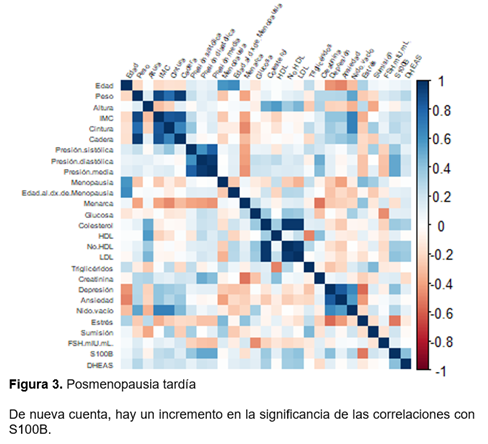
**Tabla 2: Mediante el método de Pearson se encontró las correlaciones existentes entre las variables de los grupos trabajados, se enlistaron las variables con una coeficiente de correlación mayor a 0.5 o menor a 0.5 con una equivalencia a una significancia menor a 0.05.**

* **Correlogramas:** perimenopausia, posmenopausia temprana, posmenopausia tardía, posmenopausia temprana y tardía

Para el estudio de S100B y su relación con diversos parámetros en las mujeres menopaúsicas se llevó a cabo el siguiente análisis.

Primero se obtuvieron las correlaciones de 4 conjuntos: perimenopausia, posmenopausia temprana, posmenopausia tardía y posmenopausia temprana y tardía, según la clasificación Straw.

****



Al analizar estos correlogramas, se decidió tomar como punto de corte para valorar las correlaciones significativas en aquellas mayores o iguales a 0.29 y menores o iguales que -0.29, y una vez filtradas se establecieron modelos de regresión lineal simple, y las correlaciones que obtuvieron una *p* por debajo de 0.05 fueron las siguientes:

Tabla

Descripción generada automáticamente Tabla

Descripción generada automáticamente

Durante la perimenopausia, S100B se correlaciona con la circunferencia de la cintura de la mujer, en la posmenopausia temprana su lugar lo toma la circunferencia de la cadera.

Tabla

Descripción generada automáticamente Tabla

Descripción generada automáticamente

Al ingresar a la menopausia tardía, aparecen las correlaciones con presión arterial sistólica, la presión arterial media y los niveles de estrés.

Finalmente, se decidió agrupar la posmenopausia temprana y tardía para verificar si se incrementaban las correlaciones entre S100B y las 27 variables a comparar, resultando que están asociados, de manera significativa a S100B en este grupo: circunferencia de cadera, presión sistólica, colesterol (principalmente el LDL, seguido del No HDL), depresión, DHEAS y estrés.