



Universidad de Valladolid

Facultad de Medicina

Trabajo Fin de Grado

Grado en Ingeniería Biomédica

Control metabólico de pacientes con diabetes mellitus tipo I mediante el uso de glucometrías y parámetros de red

Autor:

D. Juan Encinas Muñoz

Tutores:

Dr. D. Carlos Gómez Peña

Dr. D. Gonzalo Díaz Soto

Valladolid, julio de 2025

TÍTULO: **Control metabólico de pacientes con diabetes mellitus tipo I mediante el uso de glucometrías y parámetros de red**

AUTOR: **D. Juan Encinas Muñoz**

TUTOR/ES: **Dr. D. Carlos Gómez Peña**

Dr. D. Gonzalo Díaz Soto

DEPARTAMENTO: **Teoría de la Señal y las Comunicaciones e Ingeniería Telemática**

**Medicina, Dermatología y Toxicología.
Servicio de Endocrinología y Nutrición
del Hospital Clínico Universitario de
Valladolid**

Tribunal

PRESIDENTE: **Dr. D. Jesús Poza Crespo**

SECRETARIO: **Dr. D. Javier Gómez Pilar**

VOCAL: **Dr. D. Carlos Gómez Peña**

SUPLENTE 1: **Dr. D. Gonzalo C. Gutiérrez Tobal**

SUPLENTE 2: **Dra. D. María García Gadañón**

FECHA: **Julio de 2025**

CALIFICACIÓN:

Agradecimientos

Antes de nada, quiero agradecer a mis tutores del TFG, Carlos Gómez Peña, Jesús Poza Crespo y Gonzalo Díaz Soto, que me han ayudado siempre que lo he necesitado y me han acompañado durante este último año. Además, me han permitido vivir la experiencia de irme de Erasmus y finalizar este trabajo de manera telemática.

A mis amigos de la carrera, que han estado a mi lado durante cuatro años y que me han visto en mis mejores y peores momentos. Sin ellos, esta experiencia no habría sido la misma. A las amigas que conocí en Portugal porque, a pesar de que nos separa la distancia, me apoyan en cada decisión que tomo y son una parte muy importante de mi vida. A toda la gente que he conocido en el Erasmus, que en tan poco tiempo se han convertido en imprescindibles y con los que he vivido momentos maravillosos. A mis amigas del pueblo, que siempre me animan y muchas veces confían más en mí de lo que yo mismo lo hago. A todas aquellas personas que he ido conociendo, por un motivo o por otro, y con las que tomar un café después de meses se siente como si la última vez hubiese sido el día anterior. A Sergio, al que considero ya parte de mi familia después de tantos años a mi lado.

Por supuesto, a toda mi familia, de la que me duele en el alma tener que despedirme. A mis padres, que me han ayudado a ser quién soy y a llegar hasta aquí. A ellos les debo todo y espero ser algún día capaz de devolverles todo el amor que me han dado. A mi hermana, de la que estoy tremendamente orgulloso; te mereces lo mejor en esta vida. A mi tía, que siempre me ha escuchado con entusiasmo y que ha leído fragmentos de este trabajo para darme consejo, ayudándome a sentir que estaba yendo por el buen camino.

Pero este trabajo se lo quiero dedicar a mi abuela, que es una de las personas a las que más quiero y admiro en este mundo. Estoy muy agradecido de poder seguir compartiendo momentos contigo y verte sonreír.

Gracias a todos por haber formado parte de este camino, porque este trabajo no solo es fruto de mi esfuerzo, sino que ha podido ser gracias a todas las personas que me rodean.

Resumen

La diabetes mellitus (DM) es un conjunto de patologías de carácter crónico en las que el paciente presenta una desregulación del nivel de glucosa en sangre, caracterizada por la hiperglucemia. Concretamente, la diabetes mellitus tipo 1 (DM1) consiste en una destrucción autoinmune de las células β de los islotes de Langerhans del páncreas, lo que impide una correcta secreción de la insulina. Para evitar las complicaciones asociadas a la DM1, es esencial que los pacientes presenten un buen control glucémico. Se ha observado que este puede verse favorecido por un nivel de péptido C detectable (> 0.1 ng/mL), lo que indica que el paciente presenta cierta funcionalidad de las células β . Para mantener un buen control, es necesario monitorizar los niveles de glucemia y, en este contexto, un concepto importante es la variabilidad glucémica (VG). A lo largo de las últimas décadas se han planteado numerosos parámetros de VG, pero ninguno de ellos se ha visto que pueda servir como *gold standard*. Ante ello, este trabajo pretende realizar un análisis exploratorio de los diferentes parámetros de VG existentes, de tal forma que se determine cuáles son los que aportan una mayor información con relación al control glucémico expresado por el nivel de péptido C. Para llevar a cabo este análisis, se ha recogido una base de datos conformada por 212 pacientes con DM1 del Hospital Clínico Universitario de Valladolid (HCUV), que se dividen en 70 sujetos con péptido C detectable y 142 con péptido C indetectable. De cada uno de ellos, se extrajeron 28 parámetros de VG a partir de sus curvas de glucemia. Además, se ha realizado el cálculo de 4 variables relativas a la variación del péptido C. Como herramienta imprescindible para determinar las correlaciones entre los diferentes parámetros de VG y de estos con el péptido C, se han empleado las redes de asociación. A partir de este planteamiento inicial, se ha realizado un estudio de clasificación con técnicas de *Machine Learning* en los grupos de péptido C detectable vs. indetectable (nuestra variable *target*). Con él, se observa que una combinación de los parámetros de VG no nos permite mejorar el rendimiento de clasificación. Esto se puede deber a diversos factores, como la imprecisión de la medida de péptido C, la poca correlación entre los parámetros de VG y la variable *target* o la falta de consideración de otros factores, como las comidas, el ejercicio o las dosis de insulina. Además, todos los parámetros de VG comparten mucha información entre ellos, de tal forma que muchas veces son redundantes. Por todo esto, hay aún mucho

margen de mejora para establecer una métrica o conjunto de métricas que nos permitan determinar claramente el estado de control glucémico de un paciente con DM1.

Palabras Clave

Células β , control glucémico, diabetes mellitus tipo 1, estudio de clasificación, *Machine learning*, péptido C detectable, péptido C indetectable, redes de asociación, variabilidad glucémica.

Abstract

Diabetes mellitus (DM) is a variety of chronic disorders in which a patient has a dysregulation in the blood glucose levels, characterized by hyperglycemia. Specifically, diabetes mellitus type 1 (DM1) arises from the β -cells destruction in the pancreatic Langerhans islets due to autoantibodies, which results in an incorrect secretion of insulin. To avoid the complications associated to DM1, it is essential that diabetic patients have a good glycemic control. This can be modulated by a detectable C peptide level (> 0.1 ng/mL), which implies that the patient has a remaining β -cell functionality. Glycemic control can be achieved by monitoring glycemia levels, and, on this context, glycemic variability (GV) is of utmost importance. On the past decades, there have been proposed many GV parameters, but none of them has prevailed as gold standard. Thus, this study aims for an exploratory analysis of the current GV parameters, which can help to establish which ones give us more information about the glycemic control, expressed in terms of C peptide. To perform the analysis, a database with 212 patients with DM1 from the Hospital Clínico Universitario de Valladolid (HCUV) has been constructed. In this database, 70 patients have detectable C peptide level, while 142 have undetectable C peptide. From each of them, 28 GV parameters have been extracted from the glycemic curves, and 4 variables related to the evolution of C peptide have been calculated. Association networks are presented as an indispensable tool to perform the correlation analysis between GV parameters, and among them and the C peptide. From this initial approach, a classification study with Machine learning techniques has been proposed to separate patients with detectable vs. undetectable C peptide – our target variable. In it, it is observed that a combination of variables doesn't improve the performance of classification. The reasons are multiple, such as an imprecision in C peptide measurement, the lower correlation between VG parameters and the target variable, or the absence of external factors, like meals, exercise or insulin bolus. Moreover, all the VG parameters have a lot of information in common, so they can be redundant. All in all, a lot of improvement can be done to establish a measurement or set of measurements that allows us to completely determine the state of a patient with DM1.

Keywords

Association networks, β -cells, classification study, detectable C peptide, diabetes mellitus type 1, glycemic control, glycemic variability, undetectable C peptide, Machine learning.

Índice general

Capítulo 1. Introducción.....	1
1.1 Diabetes mellitus.....	2
1.2 Sistemas de monitorización de la diabetes.....	4
1.3 Concepto clínico de glucemia.....	7
1.4 Redes de asociación.....	10
1.5 Hipótesis y objetivos.....	12
1.6 Descripción del documento.....	13
Capítulo 2. Diabetes mellitus tipo 1.....	17
2.1 Introducción.....	18
2.2 Etiología y factores de riesgo.....	20
2.3 Presentación clínica.....	23
2.4 Diagnóstico.....	24
2.5 Prevención y tratamiento.....	26
2.6 Enfermedades asociadas y complicaciones.....	29
Capítulo 3. Control glucémico.....	33
3.1 Principales sistemas de monitorización de la glucemia.....	34
3.2 Ventajas y limitaciones de los sistemas existentes.....	36
3.3 Papel de la variabilidad glucémica en el control glucémico.....	38
3.4 Diferentes descriptores del control glucémico.....	40
Capítulo 4. Materiales y métodos.....	45
4.1 Introducción.....	46
4.2 Base de datos.....	46
4.2.1 Sujetos.....	46
4.2.2 Variables clínicas.....	48

4.2.3	Registros de glucemia.....	50
4.3	Extracción de características.....	51
4.3.1	Glucometrías clínicas.....	52
4.3.1.1	Intervalos de glucemia e índices derivados	52
4.3.1.2	Medidas de dispersión.....	55
4.3.1.3	Índices de riesgo glucémico	59
4.3.1.4	Medidas frecuenciales	60
4.3.2	Parámetros no lineales	61
4.3.3	Medidas de variación del péptido C	64
4.4	Imputación de datos	65
4.5	Análisis de correlaciones: redes de asociación	66
4.6	Selección de características.....	67
4.7	Clasificadores péptido C detectable vs. indetectable	69
4.8	Análisis estadístico.....	70
	Capítulo 5. Resultados	73
5.1	Introducción	74
5.2	Imputación de datos	74
5.3	Análisis poblacional.....	75
5.4	Análisis de correlaciones.....	86
5.5	Estudio de clasificación.....	90
	Capítulo 6. Discusión y limitaciones	96
6.1	Introducción	97
6.2	Análisis de correlaciones.....	97
6.3	Estudio de clasificación.....	100
6.4	Limitaciones.....	101
	Capítulo 7. Conclusiones y líneas futuras	104

7.1	Introducción	105
7.2	Cumplimiento de los objetivos del trabajo de fin de grado	105
7.3	Conclusiones y aportaciones originales	106
7.4	Líneas futuras	108
	Bibliografía.....	111
	Glosario de siglas y acrónimos	125

Índice de figuras

Figura 1. Comparación entre las bombas de insulina con y sin tubos, donde las áreas sombreadas representan los sitios donde la bomba puede ser llevada por el paciente (Marks et al., 2022).....	7
Figura 2. Estadistado de la DM1 en función de la alteración del perfil glucémico y la presencia de sintomatología. *Normalmente, la presencia de autoanticuerpos se manifiesta meses o incluso años antes de que empiece a haber una pérdida de células β (Katsarou et al., 2017).....	19
Figura 3. Patogenia de la DM1 mediada por la respuesta celular y humoral del sistema inmunológico. DC, células dendríticas; TCR, receptor de los linfocitos T; BCR, receptor de los linfocitos B; MHC, complejo mayor de histocompatibilidad (Katsarou et al., 2017).....	22
Figura 4. Interfaz observada por el médico en LibreView de los pacientes con DM1. Se observa un resumen de las 2 semanas anteriores que incluye parámetros de variabilidad y control glucémico (Battelino et al., 2019).....	36
Figura 5. Registros de glucemia durante dos semanas de cuatro pacientes, en los que se observan en rojo los tramos en los que no hay datos registrados debido a una desconexión del sensor. El nivel de glucemia viene expresado en mg/dL.....	51
Figura 6. Registros de glucemia durante dos semanas en los que están marcados los diferentes intervalos de glucemia. A la izquierda, se observa el TIR en verde, el TAR1 y el TBR1 en amarillo por encima y por debajo, respectivamente, y el TAR2 y el TBR2 en naranja, por encima del TAR1 y por debajo del TBR1, respectivamente. A la derecha, se observa el TITR en verde claro, el TOTR en verde oscuro por encima del TITR, el TAR1 y el TBR1 en amarillo por encima del TOTR y por debajo del TITR, respectivamente, y el TAR2 y el TBR2 en naranja, por encima del TAR1 y por debajo del TBR1, respectivamente.....	54
Figura 7. Gráfica de los niveles glucosa a lo largo del tiempo (en azul), sobre la que se observan las excursiones detectadas por el algoritmo de MAGE (en rojo). Los máximos y mínimos están marcados como círculos en color rojo.....	57

Figura 8.	Visualización de la distancia entre puntos consecutivos de glucosa, L , así como su relación trigonométrica con la variación en tiempo, Δt , y la variación de glucosa, ΔG , ambas marcadas en rojo (Peyser et al., 2018).....	61
Figura 9.	Valores de CONGA en función de la diferencia de tiempo entre muestras consecutivas para todos los pacientes de la base de datos (en azul). En naranja se observa el valor promedio de toda la base de datos.....	62
Figura 10.	P -valor del test U de Mann-Whitney entre el grupo de péptido C detectable e indetectable en función del valor del hiperparámetro r (ρ). A la izquierda se observa el grupo de sujetos pediátricos y a la derecha el de adultos. El radio óptimo en cada caso está marcado con un asterisco.....	64
Figura 11.	Distribución de los histogramas de las variables con datos perdidos antes de la imputación (en gris) y una vez imputadas (en naranja). Arriba a la derecha se observan los histogramas de HbA1c, arriba a la izquierda, los de MODD y abajo, los de Var1 (Caída 6-18 meses).....	75
Figura 12.	Distribución del parámetro HbA1c, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p -valores obtenidos con el test U de Mann-Whitney.....	79
Figura 13.	Distribución del parámetro TBR2 (TBR54), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p -valores obtenidos con el test U de Mann-Whitney.....	79
Figura 14.	Distribución del parámetro TBR1 (TBR54-70), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p -valores obtenidos con el test U de Mann-Whitney.....	79

Figura 15. Distribución del parámetro TIR, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....79

Figura 16. Distribución del parámetro TAR1 (TAR180-250), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....80

Figura 17. Distribución del parámetro TAR2 (TAR250), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....80

Figura 18. Distribución del parámetro TITR, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....80

Figura 19. Distribución del parámetro TOTR (TIR140-180), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....80

Figura 20. Distribución del parámetro CHypo, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....81

Figura 21. Distribución del parámetro CHyper, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....81

Figura 22. Distribución del parámetro GRI, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....81

Figura 23. Distribución del parámetro SD, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....81

Figura 24. Distribución del parámetro CV, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....82

Figura 25. Distribución del parámetro MAGE, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....82

Figura 26. Distribución del parámetro MODD, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....82

Figura 27. Distribución del parámetro IQR (IQR50), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....82

Figura 28. Distribución del parámetro LBGI, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....83

Figura 29. Distribución del parámetro HBGI, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....83

Figura 30. Distribución del parámetro ADRR, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....83

Figura 31. Distribución del parámetro MinGly, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....83

Figura 32. Distribución del parámetro MaxGly, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....84

Figura 33. Distribución del parámetro J, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....84

Figura 34. Distribución del parámetro LAGE, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....84

Figura 35. Distribución del parámetro IGC, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....84

Figura 36. Distribución del parámetro GVP, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....85

Figura 37. Distribución del parámetro MAG, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....85

Figura 38. Distribución del parámetro CONGA, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....85

Figura 39. Distribución del parámetro L _{ZC} , a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....	85
Figura 40. Distribución del parámetro CTM, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.....	86
Figura 41. A la izquierda, se muestran los gráficos de estabilidad y, a la derecha, las redes de asociación con un umbral de 0.5. Arriba del todo se observa la red global de toda la base de datos, en medio la red que solo incluye los pacientes con péptido C detectable y abajo la red con los pacientes que tienen péptido C indetectable.....	87
Figura 42. A la izquierda, la gráfica de estabilidad de la red de asociación y, a la derecha, la propia red de asociación con todos los parámetros de VG y la variable dicotómica ‘isReserve’ (isR).....	89
Figura 43. A la izquierda, la gráfica de estabilidad de la red de asociación y, a la derecha, la propia red de asociación con todos los parámetros de VG, el valor basal de péptido C (C _{pp}) y las diferentes variables de variación del péptido C (V _{rC} , V _{rM} , V _{r1} , V _{r2}).....	89
Figura 44. Curvas ROC de los modelos de clasificación generados a partir de las variables seleccionadas de los clústeres de las redes de asociación. A la izquierda se observa el modelo de LDA, en el medio el de LR y a la derecha el de QDA.....	90
Figura 45. Curvas ROC de los modelos de clasificación generados a partir de las variables seleccionadas con el algoritmo Sequential Forward Selection (SFS). A la izquierda se observa el modelo de LDA, en el medio el de LR y a la derecha el de QDA.....	91
Figura 46. Curvas ROC de los modelos de clasificación generados a partir de las variables seleccionadas con el algoritmo FCBF. En la primera fila, se emplean todas las	

variables seleccionadas en más de 50 iteraciones; en la segunda, en más de 100, y en la tercera, en más de 125. A la izquierda se observa el modelo de LDA, en el medio el de LR y a la derecha el de QDA.....91

Figura 47. Curvas ROC de los modelos de clasificación generados a partir de las variables seleccionadas jugando con el umbral de las redes de asociación. En la primera fila, se emplean las variables más relevantes que no presentan correlaciones más fuertes de 0.7 con ningún otro parámetro y, en la segunda fila, aquellos sin correlaciones más fuertes de 0.8. A la izquierda se observa el modelo de LDA, en el medio el de LR y a la derecha el de QDA.....92

Índice de tablas

Tabla 1.	Datos sociodemográficos y clínicos de la base de datos en función del nivel de péptido C y el grupo de edad. m: media; SD: desviación estándar (standard deviation); H: hombres; M: mujeres.....	48
Tabla 2.	Clasificación de las variables de la base de datos en función de su escala de medición.....	49
Tabla 3.	Resumen de los intervalos de glucemia y los parámetros clínicos de resumen del control glucémico. TBR2: tiempo por debajo de rango nivel 2; TBR1: tiempo por debajo de rango nivel 1; TIR: tiempo en rango; TAR1: tiempo por encima de rango nivel 1; TAR2: tiempo por encima de rango nivel 2; TITR: tiempo en rango estrecho; TOTR: tiempo fuera del rango estrecho; CHypo: carga de hipoglucemia; CHyper: carga de hiperglucemia; GRI: índice de riesgo glucémico.....	54
Tabla 4.	Principales parámetros clínicos de VG.....	56
Tabla 5.	Análisis poblacional de los parámetros de VG para el grupo de sujetos adultos. En la primera columna se observan las variables bajo estudio, en la segunda las medias $\pm$ SD de cada parámetro para los pacientes con péptido C detectable (> 0.1 ng/mL), en la tercera las medias $\pm$ SD de cada parámetro para los pacientes con péptido C indetectable (< 0.1 ng/mL) y en la última el p -valor obtenido con el test U de Mann-Whitney.....	76
Tabla 6.	Análisis poblacional de los parámetros de VG para el grupo de sujetos pediátricos. En la primera columna se observan las variables bajo estudio, en la segunda las medias $\pm$ SD de cada parámetro para los pacientes con péptido C detectable (> 0.1 ng/mL), en la tercera las medias $\pm$ SD de cada parámetro para los pacientes con péptido C indetectable (< 0.1 ng/mL) y en la última el p -valor obtenido con el test U de Mann-Whitney.....	77
Tabla 7.	Valores del área bajo la curva ROC (AUC) obtenidos con los diversos modelos de clasificación para cada uno de los métodos de selección de variables aplicados. Junto al AUC de los modelos, se muestran los AUC de las variables que lo	

conforman si cada una de ellas fuera la única que se emplea para el modelo, en orden
decreciente de valor de AUC.....94

Capítulo 1. Introducción

1.1 Diabetes mellitus.....	2
1.2 Sistemas de monitorización de la diabetes.....	4
1.3. Concepto clínico de glucemia.....	7
1.3 Redes de asociación.....	10
1.4 Hipótesis y objetivos.....	12
1.5 Descripción del documento.....	13

1.1 Diabetes mellitus

La diabetes mellitus (DM) es un conjunto de trastornos del metabolismo de la glucosa que se caracterizan por la presencia de hiperglucemia, generada por un desbalance entre la gluconeogénesis y la glucogenólisis, la secreción de insulina por la célula β pancreática, así como el uso inapropiado de la glucosa como fuente de energía (ElSayed et al., 2025; WHO, 2019). El diagnóstico de este tipo de patologías se basa en la medición de glucosa en plasma en ayunas o tras sobrecarga oral de glucosa, o bien en la determinación de la hemoglobina glicosilada (HbA1c). Dentro de la DM, se distinguen dos grandes grupos: diabetes mellitus tipo 1 (DM1) y tipo 2 (DM2), si bien existen otros tipos como la diabetes mellitus gestacional (DMG), que es primeramente detectada durante el embarazo, y otra serie de tipos específicos, como la diabetes monogénica o pancreopriva. Aparte, la Organización Mundial de la Salud (OMS) incluye en su clasificación de 2019 la llamada diabetes híbrida, que tiene un fenotipo intermedio entre la DM1 y la DM2 (WHO, 2019), pero esta no está contemplada por la Asociación Americana de Diabetes (*American Diabetes Association*, ADA) en su clasificación más reciente en 2025.

La DM1 es una patología que en un 70-90% de los casos presenta una causa autoinmune, de modo que hay una destrucción de las células β de los islotes de Langerhans del páncreas que acaba desembocando en una deficiencia absoluta de insulina en el organismo (ElSayed et al., 2025; WHO, 2019). En la minoría restante, se habla de DM1 idiopática, en la que no se detectan autoanticuerpos y la causa de la enfermedad es desconocida. En global, la DM1 inmunomediada supone un 5-10% del total de casos de diabetes (ElSayed et al., 2025). Dentro de este grupo existe el concepto de diabetes latente autoinmune en adultos (*Latent autoimmune diabetes in adults*, LADA), en la que la destrucción de las células β ocurre lentamente en adultos, de modo que hay una producción marginal de insulina durante un período prolongado de tiempo. Aunque este término es clínicamente correcto y la OMS ha diferenciado este subgrupo dentro de la diabetes híbrida, realmente se sigue considerando un subtipo de DM1. De hecho, estos pacientes no suelen necesitar terapia con insulina al inicio de la enfermedad, de modo que muchas veces son mal clasificados como DM2. Aunque inicialmente presentan un buen control simplemente con cambios en el estilo de vida y agentes orales, acaban necesitando insulina más rápidamente que el clásico paciente con DM2 (WHO, 2019).

Prácticamente el resto de los casos de diabetes, entre el 90% y el 95%, son de pacientes con DM2 (ElSayed et al., 2025; WHO, 2019). En este caso, la patología cursa con el desarrollo de resistencia a la insulina por parte de los tejidos periféricos y posterior fracaso de la liberación de insulina por parte del páncreas. No obstante, el proceso de destrucción de las células β en la DM2 está menos entendido, aunque se cree que están involucrados factores como la predisposición genética, los cambios epigenéticos, la inflamación y el estrés metabólico (ElSayed et al., 2025). El diagnóstico de la DM2 suele realizarse como prediabetes, en la que el paciente presenta cierta disglucemia, con valores de HbA1c de 5.7-6.4%, un estadio intermedio entre la normoglucemia y la diabetes franca (ElSayed et al., 2025). La DM2 se asocia con obesidad, dislipemia e hipertensión, y es cada vez más habitual debido al cambio en los hábitos de vida de la población general. De hecho, aunque suele ser más habitual en adultos, es creciente su incidencia en niños y adolescentes. En muchas ocasiones, la DM2 permanece sin diagnosticar durante muchos años, porque los pacientes no presentan los síntomas clásicos de la diabetes, pero esto no les exime de sufrir las complicaciones asociadas. Por ello, en estos casos es crítico el diagnóstico para establecer un tratamiento que evite la aparición de las diversas complicaciones. Factores de riesgo para el desarrollo de DM2 son la avanzada edad, la obesidad, los hábitos de vida sedentarios o la presencia previa de DMG (ElSayed et al., 2025; WHO, 2019).

Una complicación aguda de la DM1 al debut es la cetoacidosis diabética (CAD), pero en la DM2 rara vez se da de manera espontánea. Normalmente suele estar asociada a una insulinopenia por un incorrecto tratamiento con insulina, la presencia de DM2 propensa a cetosis, el estrés de otra enfermedad como una infección o un infarto de miocardio, determinantes sociales de salud o el uso de ciertos medicamentos como glucocorticoides, antipsicóticos de segunda generación o inhibidores de SGLT-2 (ElSayed et al., 2025).

En ocasiones, puede aparecer un cuadro de diabetes por reducción estructural y funcional de la insulina ante ciertas patologías del páncreas exocrino. Esta patología, normalmente mal diagnosticada como DM2 en un 87.8% de los casos, es conocida como diabetes pancreopriva, y suele ir asociada a trastornos como pancreatitis, trauma o pancreatectomía, neoplasia, fibrosis quística, hemocromatosis, pancreatopatía fibrocalculosa, trastornos genéticos raros, y otras enfermedades idiopáticas (ElSayed et al., 2025; WHO, 2019). También puede ocurrir que la diabetes surja en el contexto de un trasplante, en cuyo caso hablamos de diabetes mellitus postrasplante (DMPT),

independientemente de que la diabetes haya surgido a raíz de la operación o estuviese de forma subyacente, aunque aún sin diagnosticar. En este caso, los factores de riesgo son los habituales de la diabetes, a los que se suman otros específicos del trasplante, como el uso de inmunosupresores (ElSayed et al., 2025). Menos de un 5% de los casos de diabetes son debidas a síndromes diabéticos monogénicos, que generalmente son hereditarios y pueden aparecer en los primeros 6 meses de vida. En este grupo se encuentran defectos monogénicos de la función de la célula β , como la diabetes neonatal o el MODY (*Maturity-Onset Diabetes of the Young*), y defectos monogénicos de resistencia a la insulina, como lipodistrofias monogénicas. Estos últimos son menos frecuentes que el MODY o la diabetes neonatal (ElSayed et al., 2025; WHO, 2019). Especial mención merece la DMG, que es aquella diabetes que aparece durante el embarazo. No obstante, es muy común en muchas mujeres que aparezcan síntomas debidos a una DM2 subyacente que había permanecido sin diagnosticar hasta ese momento (ElSayed et al., 2025).

La incidencia global de la DM presenta una tendencia creciente en todo el mundo, de tal forma que ha pasado de afectar a 108 millones de personas en 1980 a 476 millones en 2017 (Arokiasamy et al., 2021). De hecho, la ratio de incidencia estandarizada por edad en 2017 se sitúa en 285 por cada 100000 habitantes (Liu et al., 2020). Esta tendencia creciente, marcada por un porcentaje anual de cambio estimado del 0.87%, es sobre todo debida a la carga de la DM2; mientras que la DM1 ha tenido una contribución más pequeña. El preocupante impacto de la DM2 es más acuciante en países en desarrollo, principalmente del este de Europa, parte de Asia y África, en los que la prevalencia es más elevada (Arokiasamy et al., 2021; Liu et al., 2020; Untung Halajur & Irfan Setiawan, 2021). La principal causa es el cambio en el estilo de vida, de modo que la población es más propensa a inactividad física, malos hábitos alimenticios, hipertensión y dislipemia. Además, 1.4 millones de muertes, es decir, un 2.5% del total de muertes en el mundo, fueron debidas a la DM en 2017 (Arokiasamy et al., 2021).

1.2 Sistemas de monitorización de la diabetes

La DM es una enfermedad crónica, de tal forma que los pacientes se enfrentan en su vida diaria a la ardua tarea de regular su nivel de glucemia continuamente. Esto es especialmente complejo, porque los pacientes deben monitorizar sus niveles de glucosa en sangre de manera regular y, además, calcular las dosis adecuadas de insulina que se

deben administrar en cada momento. Para ello, es esencial que exista una adherencia del paciente al proceso de monitorización, o de lo contrario puede haber riesgo de un mal control glucémico, que conlleva la aparición de complicaciones asociadas a la hiperglucemia y periodos de hipoglucemia, peligrosos para la salud del paciente (Kamusheva et al., 2021; Waite et al., 2018). De hecho, debe existir un balance entre estos dos aspectos a la hora de controlar el nivel de glucemia.

Así, los sistemas de monitorización de la diabetes incluyen esencialmente un sistema de monitorización de la glucosa y pueden asociarse a un sistema de infusión de insulina, que es posible integrar posteriormente en dispositivos más complejos como sistemas híbridos de asa cerrada (*hybrid closed loop*, HCL). Tradicionalmente, los sistemas de monitorización de la glucosa se han basado en medidores de glucosa mediante punción capilar o *self-monitoring of blood glucose* (SMBG). En este caso, el propio paciente puede medir la glucemia capilar 3-6 veces al día para un correcto seguimiento, pero se desconocen las fluctuaciones de glucemia o la presencia de hipoglucemia asintomática (Jones et al., 2019; Kamusheva et al., 2021). Por este motivo, surgen los sistemas de monitorización continua de glucosa (*continuous glucose monitoring*, CGM). En estos casos, se obtienen medidas de la glucosa intersticial de manera continuada gracias a un sensor subcutáneo colocado en el brazo o en el abdomen (Jones et al., 2019). Su mayor utilidad es que permite detectar las fluctuaciones de glucosa a tiempo real, evaluar las hipoglucemias, incluidas las nocturnas, o identificar hipoglucemias en pacientes que han perdido la sensibilidad a ellas, de modo que no presentan sintomatología en periodos de hipoglucemia (Waite et al., 2018). Este tipo de sensores han sufrido numerosos avances a nivel tecnológico desde la aparición de la CGM en tiempo real, pero especial interés en nuestro caso tienen los sistemas de monitorización flash de glucosa. Dentro de este campo, destacamos el sistema FreeStyle Libre v2 desarrollado por Abbott Diabetes, que es el sistema de monitorización flash de glucosa mayoritario financiado en la Sanidad Pública de Castilla y León (SACYL) (Díaz-Soto et al., 2024).

Para la administración de insulina, existen dos métodos fundamentalmente: la inyección diaria múltiple de insulina (*multiple daily injection*, MDI) y la infusión continua de insulina subcutánea (*continuous subcutaneous infusion of insulin*, CSII) (Waite et al., 2018). La terapia basada en MDI suele consistir en la inyección de insulina, que inicialmente se hacía por medio de jeringas de metal y agujas reusables, pero se trataba de un método muy rudimentario, poco preciso y no exento de riesgos de complicaciones

e infecciones. A raíz de las desventajas de este sistema se desarrollaron las plumas de insulina, que pueden ser reusables o precargadas, ambas con una dosis máxima de 3 mL. En ellas, el paciente puede marcar una dosis de 0.5, 1 o 2 unidades de insulina, que se administran por medio de una aguja (Masieriek et al., 2022). La diferencia entre ellas es que las plumas reusables presentan varios cartuchos que el paciente tiene que ir colocando; mientras que las plumas precargadas tienen los cartuchos ya instalados. El problema con estos sistemas es que, muchas veces, los pacientes se olvidan de inyectarse las dosis de insulina o lo hacen tarde. Por ello, se han desarrollado plumas inteligentes y tapones para plumas, que son tecnologías que se pueden conectar por Bluetooth al teléfono del paciente para ayudar a calcular las dosis de insulina (Marks et al., 2022). Esto se consigue gracias a que estos sistemas digitalizados tienen memoria, de forma que detectan y almacenan los datos de las inyecciones de insulina para después hacer recomendaciones al paciente y recordarles cuando se deben colocar las dosis. La primera pluma inteligente, InPen™, fue desarrollada por Companion Medical y aprobada por la *Food and Drug Administration* (FDA) de EE. UU. en 2017. Por su parte, el primer tapón para plumas fue desarrollado por Bigfoot Biomedical y aprobado por la FDA en 2021 (Masieriek et al., 2022).

La terapia de CSII, por el contrario, suele realizarse por medio de bombas de insulina, que permiten una administración continua de insulina a nivel subcutáneo. Este sistema presenta un perfil mucho más fisiológico, aunque es necesario llevar el dispositivo todo el tiempo. Además, evitamos tener que estar inyectando dosis de manera física cada vez que el paciente necesite insulina y podemos regular los niveles de insulina de manera más personalizada (Marks et al., 2022; Waite et al., 2018). Las bombas de insulina pueden ser de dos tipos: sistemas tradicionales o con tubos, y sistemas sin tubos o parches. Las bombas con tubos presentan una pantalla interactiva donde se pueden introducir los datos de las dosis de insulina, diversos cartuchos con la insulina cargada, un sistema de tubos que conecta la bomba al paciente y una cánula subcutánea para administrar la insulina. Al contrario, las bombas parche no necesitan un sistema de tubos, pero abarcan una mayor extensión en el cuerpo del paciente. En este caso, el cartucho y el sitio de infusión se encuentran integrados en un único dispositivo, pero necesitan un controlador externo. La comparación de ambas se observa en la Figura 1. Cualquiera de los dos sistemas

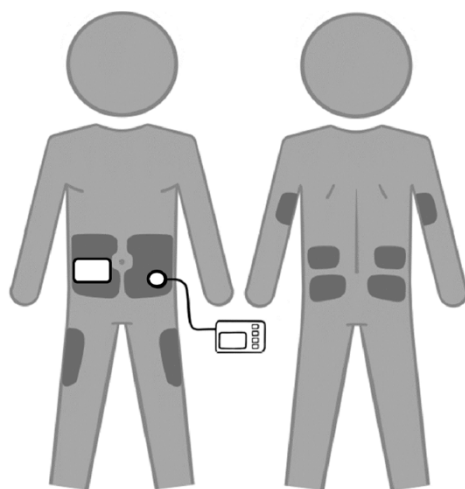


Figura 1. Comparación entre las bombas de insulina con y sin tubos, donde las áreas sombreadas representan los sitios donde la bomba puede ser llevada por el paciente (Marks et al., 2022).

permanece conectado al paciente durante 3 días, período tras el que el punto de infusión debe ser modificado, ya sea el parche o la cánula subcutánea (Marks et al., 2022).

Actualmente, es posible combinar la CGM en tiempo real con las bombas de insulina por medio de sistemas HCL. Estos presentan algoritmos que permiten ajustar las dosis de insulina en base a los datos de CGM del paciente (Marks et al., 2022; Waite et al., 2018). De hecho, el objetivo a futuro es que este sistema cercano a un páncreas artificial también sea capaz de regular la infusión de glucagón y la libración automática de insulina con las ingestas de manera no anunciada. Todo este sistema integrado puede ser telemonitorizado por los profesionales de la salud, de modo que pueden ver en diferido la respuesta del sistema a los diferentes eventos que sufre el paciente, así como el control glucémico que presenta cada paciente de manera personalizada.

1.3 Concepto clínico de glucemia

La variabilidad glucémica (VG) se puede definir como las oscilaciones entre máximos y mínimos que se pueden observar en las curvas de glucemia, tanto a corto plazo como a largo plazo (Breyton et al., 2021; Gómez et al., 2019). A corto plazo es más sencillo de determinar, puesto que la variabilidad de un paciente se estima a partir de las excursiones que se observan a lo largo de un día o en días consecutivos. Aunque parece un concepto muy simple, se puede considerar la magnitud de las excursiones, la duración de eventos críticos (hiperglucemia o hipoglucemia), la cantidad de oscilaciones en un día, la estabilidad de la curva de glucemia en días alternos, etc. Tradicionalmente este aspecto se ha asociado con los eventos de hiperglucemia, que al final son los que producen las

complicaciones microvasculares y macrovasculares a largo plazo; sin embargo, esto tan solo incluye una visión muy reducida dentro de la magnitud del concepto que se intenta describir. Por otra parte, a largo plazo el principal descriptor que existe es la HbA1c, si bien se trata de un aspecto más abstracto. En este caso, podríamos considerar la tendencia del paciente a mantener un patrón de glucemia a largo plazo. En conclusión, podríamos decir que la VG a corto plazo es una variabilidad inter o intradiaria; mientras que la VG a largo plazo es una variabilidad intrapaciente con respecto a un cierto perfil estándar, que es individualizado en cada caso. Como se puede ver, se trata de un concepto con una alta dimensionalidad y que no es tan sencillo de describir con una única medida, como se podría pensar inicialmente. En cualquier caso, cada día se considera más ligado a la homeostasis de la glucosa en pacientes diabéticos.

De hecho, el concepto de las fluctuaciones de glucosa y la VG empezó a evaluarse ya en la década de los 70. No obstante, la aparición de los sistemas de CGM a finales de los 90 supuso el punto de inflexión que permitió por primera vez cuantificar la VG (Mo et al., 2024). Con estos sistemas, que registran los niveles de glucosa intersticial del paciente cada pocos minutos, tenemos una curva de glucemia del paciente a partir de la que se puede extraer diversa información. De hecho, en las últimas décadas han surgido diversas métricas para caracterizar el control glucémico del paciente en términos de variabilidad, desde medidas de dispersión clásicas (desviación estándar, SD, o rango intercuartil, IQR), hasta métricas basadas en la comparación entre días consecutivos (media de diferencias diarias, MODD), en la morfología de las curvas de glucemia (amplitud media de excursiones de glucosa, MAGE, o cambio de glucosa absoluta media, MAG), o en índices de riesgo (índice de riesgo glucémico, GRI; índice de baja glucosa en sangre, LBGI, o índice de alta glucosa en sangre, HBGI) (Khurana et al., 2025; Mo et al., 2024). Incluso se ha llegado a plantear un análisis en frecuencia por medio de la Transformada de Fourier o el planteamiento de métricas no lineales que describen la complejidad de series temporales, como la dimensión fractal o la entropía multiescala (Breyton et al., 2021; Rodbard, 2018). Al final, en la clínica se emplea como métrica de VG el denominado coeficiente de variación (CV), que se calcula dividiendo la SD por el promedio de los niveles de glucemia en el intervalo temporal considerado. No obstante, no existe una medida *gold standard* que se haya impuesto por encima de las demás, a pesar de la gran cantidad de análisis que se han realizado.

Otro concepto que se considera que está muy asociado con la VG es la reserva pancreática. En este contexto, se refiere al número de células β funcionantes que siguen segregando insulina en un paciente con diabetes. Las células β productoras de insulina se localizan en los islotes de Langerhans, que conforman la porción endocrina del páncreas (~2-3%). Este número debe ser suficiente para que se mantenga la homeostasis de la glucosa, donde se suele considerar un umbral del 50% para que la porción de células β funcionante pueda hacerse cargo de la regulación de los niveles de glucemia de todo el organismo. En el caso de los pacientes con DM1, cuando inician la fase sintomática su reserva se suele encontrar por debajo del 25%, de modo que los niveles de secreción de insulina son muy bajos y aparece la sintomatología asociada con la hiperglucemia (Flatt et al., 2021). Normalmente, para medir la función de las células β en pacientes diabéticos se suele emplear el péptido C como medida indirecta. Este péptido es producido junto a la insulina en las células β , donde el conjunto completo se denomina proinsulina. Luego, de esta proteína se eliminan algunas regiones, una de ellas el péptido C. Posteriormente, se puede utilizar como biomarcador, ya que su vida media es de 20-30 minutos, lo que da una ventana más estable que la insulina, cuya vida media es de 3-5 min (Leighton et al., 2017). En una persona sana, los niveles de esta molécula son habitualmente de 0.9-1.8 ng/mL, con una subida de hasta 3-9 ng/mL en período posprandial (Venugopal et al., 2023). Cuando el paciente presenta DM, los niveles están reducidos, normalmente por debajo de 0.6 ng/dL, valor que está estrechamente correlacionado con la presencia de DM1. De hecho, se ha observado que, en el marco de desarrollo de la patología, los niveles de péptido C se van reduciendo con el tiempo (Leighton et al., 2017).

A la hora de tomar las medidas del péptido C, hay que tener en cuenta que es metabolizado por los riñones, excepto un 5-10% que es excretado en la orina. Por ello, las medidas pueden estar alteradas en pacientes con enfermedad renal crónica. Además, los pacientes con altos niveles de autoanticuerpos contra la insulina pueden tener un nivel erróneamente elevado de péptido C, ya que estos autoanticuerpos tienen una elevada afinidad por la proinsulina y el péptido C (Leighton et al., 2017; Venugopal et al., 2023). Los niveles de péptido C se pueden medir en orina o en plasma, si bien se prefiere este último. Aunque es necesario tener en cuenta que el péptido C es lábil dentro del plasma debido a su susceptibilidad a enzimas proteolíticas, se puede mantener estable durante 24 horas en tubos EDTA preparados. Las muestras en sangre se pueden tomar de manera aleatoria, en ayunas o en un estado de estimulación. Mientras que las medidas aleatorias

se pueden hacer en cualquier momento del día sin tener en cuenta las comidas, las medidas de péptido C en ayunas se deben hacer tras 8-10 h de ayuno. Por su parte, las medidas en estado de estimulación incluyen diversos estímulos, como glucagón, glucosa oral/intravenosa, tolbutamida, sulfonilurea o una comida mixta. En el caso del HCUV, normalmente se hace una medida en ayunas, porque facilita el incluirla en un análisis de sangre rutinario.

1.4 Redes de asociación

Como podemos ver, la VG en el contexto de la DM es un concepto muy amplio y que engloba aspectos muy diversos. En relación a este tema, han surgido una gran cantidad de métricas en el ámbito clínico en un intento de caracterizarla (Breyton et al., 2021; Rodbard, 2018). Debido a que ninguna de ellas se ha establecido como *gold standard*, en este estudio se pretende hacer un análisis exploratorio de la gran cantidad de medidas existentes. Y, para ello, surgen como una potente herramienta las redes de asociación. Estas se definen como un tipo de red en las que cada uno de los nodos representa a una variable y los enlaces entre nodos nos indican la correlación que existe entre esas dos variables (Ramos-Fernández et al., 2009). El grosor de los enlaces entre nodos nos indica la fuerza de estas correlaciones, de modo que cuanto más grueso sea el vínculo que une dos variables, mayor correlación existe entre ellas y, en consecuencia, comparten más cantidad de información.

Al igual que cualquier otro tipo de red, los enlaces de las redes de asociación pueden ser binarios o ponderados. En caso de los grafos binarios, solo existen dos valores de asociación posibles: 0 en caso de que se considere que las variables no tengan asociación, y 1 cuando hay una asociación presente entre las variables. Esta es una visión más simplista, en la que se reduce todo a determinar la presencia o ausencia de asociación, como si se tratase de un clasificador. En cambio, los grafos ponderados presentan enlaces con valores continuos entre 0 y 1, de tal forma que la correlación entre variables es más fuerte cuanto más cerca del 1 y más débil cuanto más cerca del 0 (Rubinov & Sporns, 2010). En nuestro caso de estudio, se emplean grafos semi-ponderados, en los que las correlaciones también toman valores continuos, pero solo se muestran en la red cuando superan un cierto umbral. También es importante entender si las correlaciones entre variables son directa o inversamente proporcionales, ya que esto nos aporta más información que solo la fuerza de la correlación. Otro concepto importante en las redes

en general es la direccionalidad de los enlaces. En caso de que los enlaces sean direccionales, se pueden observar relaciones causa-efecto en las redes (Barabási, 2016). No obstante, este no suele ser el caso de las redes de asociación, en las que los enlaces suelen ser no direccionales.

Esta aproximación nos permite visualizar relaciones complejas entre diferentes variables de muy distinta naturaleza, lo que suele ser habitual en el ámbito biomédico. Nos pueden ayudar a entender la fisiopatología de las enfermedades de una manera más integrada, puesto que se pueden incluir en los nodos factores etiológicos, genes alterados, mecanismos de expresión activados/inhibidos, manifestaciones clínicas, etc. De hecho, en enfermedades como el cáncer se puede observar que la transición del estado sano al enfermo se produce ante un aumento de la conectividad entre los genes mutados y los mecanismos moleculares que involucran (Ranea et al., 2022). También en este contexto, se puede estudiar el papel de la microbiota en la aparición de enfermedades. En este caso, el abordaje más adecuado es un estudio de la proteómica, en el que se analizan las correlaciones entre las diferentes enzimas, entre las que se incluyen las enzimas de los microorganismos pertenecientes a la microbiota (Ai et al., 2019). Otro campo muy estudiado es la psicopatología y la psiquiatría, en el que se estudian redes de síntomas, que presentan las principales manifestaciones de diversas alteraciones psiquiátricas. Normalmente, estas redes están prácticamente inactivas en una persona sana; mientras que en una persona que presenta algún trastorno psiquiátrico, se activan ante la presencia de ciertos *triggers*, incrementando la conectividad entre todos los nodos de la red (Borsboom, 2017; Borsboom & Cramer, 2013; Jimeno et al., 2020). En el campo de la apnea del sueño, este tipo de aproximaciones también es muy útil, ya que a estos pacientes se les hace una cantidad enorme de pruebas durante las polisomnografías, de modo que hay una gran cantidad de datos que presentan elevada heterogeneidad (Gutiérrez-Tobal et al., 2021). Por último, en el campo de la bioinformática, las redes de asociación son una herramienta versátil para poder trabajar con la elevada cantidad de datos existentes. En estos casos, los nodos son genes, proteínas y metabolitos, de manera que las relaciones entre ellos nos dan una idea de los mecanismos moleculares en los que está implicado cada uno de ellos (Gao et al., 2019). Asimismo, esto nos permite explorar las interacciones gen-proteína y proteína-proteína que se dan en las células de nuestro organismo de forma mucho más visual (Szklarczyk et al., 2023).

En este estudio, las redes de asociación nos van a permitir establecer una relación entre las diferentes medidas de VG existentes en la clínica, así como algunas otras de análisis no lineal de series temporales complejas, con parámetros de control glucémico. Las variables que se van a manejar en referencia al control glucémico realmente no son infalibles, pero son aquellas que han demostrado a nivel clínico una mayor aplicabilidad para caracterizar el estado del paciente, como son la HbA1c, el TIR o el péptido C (Battelino et al., 2019; Bergenstal, 2015; Leighton et al., 2017).

1.5 Hipótesis y objetivos

Este TFG aborda la elevada dimensionalidad de la VG en la DM1 por medio de un análisis de correlaciones. Esto se debe a que la mayor parte de los estudios existentes analizan uno o varios parámetros de VG, pero se establecen correlaciones individuales, sin llegar a obtener una visión global del problema bajo estudio (Breyton et al., 2021; El-Laboudi et al., 2016; Gómez et al., 2019; Khurana et al., 2025; Lazar et al., 2023; Mo et al., 2024). Por ello, con este trabajo se plantea intentar definir el estado del paciente, no seleccionando una única medida que sea el *gold standard*, sino combinando aquellas que sean más relevantes y que permitan caracterizar todos los aspectos de la VG. Por tanto, la hipótesis de partida de este TFG es que **una combinación de los parámetros de VG permite determinar el estado de control glucémico de un paciente con DM1**, así como identificar qué aspectos del control necesitan ser mejorados.

En base a esta hipótesis, se propone como objetivo principal del TFG **determinar cuáles son los parámetros de VG que aportan más información con respecto al control metabólico de los pacientes diabéticos**. Para poder abordar el problema, se ha llevado a cabo un análisis exploratorio con técnicas estadísticas clásicas y redes de asociación. A la hora de comprobar la eficacia de estos parámetros, se han planteado diversos clasificadores que involucran algunas de las medidas bajo estudio. En caso de identificar las variables más importantes, esto nos puede ayudar a describir de manera numérica el riesgo que tiene el paciente de sufrir eventos de hipoglucemia o hiperglucemia, la estabilidad de sus niveles de glucemia a lo largo del día, o si sus picos posprandiales son demasiado elevados.

Para lograr el propósito principal del TFG, es necesario que se cumplan diversos objetivos específicos, que son expuestos a continuación:

- i. Hacer una búsqueda bibliográfica sobre la DM, centrada en la DM1, para entender mejor los aspectos fisiopatológicos, la etiopatogenia, el diagnóstico, las complicaciones y las principales formas de monitorización.
- ii. Realizar la recogida de datos de pacientes con DM1, así como la descarga de los datos de glucemia en crudo.
- iii. Hacer una revisión de los principales parámetros de VG que se emplean en la clínica, así como de las fórmulas para su implementación, a los que se añadirán parámetros de análisis no lineal que cuantifican variabilidad y/o complejidad.
- iv. Realizar un curado de la base de datos, de tal forma que se eviten los valores perdidos y las variables queden correctamente codificadas.
- v. Llevar a cabo un análisis exploratorio inicial para comprobar si los parámetros bajo estudio presentan distribuciones muy distintas en pacientes con y sin reserva pancreática.
- vi. Realizar un análisis de correlaciones para establecer grupos de variables que describen el mismo aspecto de la VG, así como determinar aquellas que presentan más correlación con los niveles de péptido C.
- vii. Intentar establecer una clasificación entre pacientes con reserva pancreática y sin reserva por medio del empleo de múltiples parámetros de VG.
- viii. Analizar los resultados obtenidos para extraer las conclusiones pertinentes y determinar las limitaciones y líneas futuras.

1.6 Descripción del documento

La memoria del presente TFG está conformada por 7 capítulos, que son presentados y descritos brevemente a continuación:

- **Capítulo 1. Introducción.** Se presenta un breve resumen de la patología bajo estudio – la DM –, los principales mecanismos de monitorización de la diabetes y

las métricas empleadas para determinar el estado del paciente. Además, se incluye una breve descripción de las redes de asociación, así como sus principales aplicaciones en el ámbito biomédico. El capítulo concluye con la presentación de la motivación del trabajo, así como la hipótesis y los objetivos, y un breve resumen de los contenidos del TFG.

- **Capítulo 2. Diabetes mellitus tipo 1.** El capítulo inicia con la descripción del concepto de DM1 y las dos grandes clases existentes: DM1A y DM1B. Se continúa con una explicación de la etiopatogenia de la enfermedad, las principales manifestaciones clínicas y los mecanismos de prevención y tratamiento existentes. Se finaliza con la presentación de las comorbilidades y complicaciones más importantes asociadas a la DM1.
- **Capítulo 3. Control glucémico.** Este capítulo se centra en describir los principales mecanismos de monitorización de la glucemia del paciente, así como sus ventajas e inconvenientes. Está centrado en el sistema FreeStyle Libre 2, del que se expone la interfaz y los parámetros más importantes que se pueden observar en ella. Finalmente, se mencionan los diferentes parámetros de VG que se han estudiado en el ámbito clínico.
- **Capítulo 4. Materiales y métodos.** Se inicia con una descripción de la base de datos, así como los mecanismos para su recogida. Tras ello, se explican las fórmulas y algoritmos empleados para realizar la extracción de parámetros a partir de las curvas de glucemia, además de la imputación de los datos para un correcto curado de la base de datos. Posteriormente, se exponen la generación de redes de asociación, los métodos de selección de variables empleados y los diferentes clasificadores. Por último, este capítulo describe los métodos estadísticos empleados para validar las diferentes técnicas aplicadas en el estudio.
- **Capítulo 5. Resultados.** En este capítulo se presentan las diferentes gráficas y estadísticos que se han obtenido durante el proyecto. Primeramente, se exponen diversos resultados de análisis poblacional de la base de datos; después, se muestran las correlaciones por medio de las redes de asociación y, por último, se incluyen los rendimientos de los diversos modelos de clasificación.

- **Capítulo 6. Discusión y limitaciones.** Este capítulo está dedicado a comentar los principales resultados obtenidos a lo largo del estudio. Además, se presentan las principales limitaciones que han surgido a lo largo de la realización del TFG.
- **Capítulo 7. Conclusiones y líneas futuras.** Este último capítulo expone las conclusiones más importantes que se extraen del TFG llevado a cabo, así como las principales líneas de mejora de cara a futuros estudios relacionados.

Capítulo 2. Diabetes mellitus tipo 1

2.1 Introducción.....	18
2.2 Etiología y factores de riesgo.....	20
2.3 Signos y síntomas.....	23
2.4 Diagnóstico.....	24
2.5 Tratamiento.....	26
2.6 Enfermedades asociadas.....	29

2.1 Introducción

La diabetes mellitus tipo 1 (DM1) es una enfermedad crónica en la que hay una alteración de la homeostasis de la glucosa, que se manifiesta habitualmente en forma de hiperglucemia. Esto ocurre porque hay una destrucción de las células β de los islotes de Langerhans del páncreas, que son las encargadas de producir insulina (Katsarou et al., 2017). En este contexto, la insulina juega un papel esencial, porque es la única hormona hipoglucemiante de nuestro cuerpo. Por tanto, solamente ella es capaz de reducir los niveles de glucosa en sangre e introducirla en los diferentes órganos. De hecho, un déficit de insulina conlleva la imposibilidad por parte de los tejidos para emplear la glucosa como fuente de energía (Rahman et al., 2021).

Tradicionalmente, tanto la ADA como la OMS han clasificado la DM1 en dos clases: diabetes tipo 1A (DM1A) y 1B (DM1B) (Qiu et al., 2024). En 2019, la OMS introdujo modificaciones en su clasificación, de modo que englobaba a todos los subtipos de diabetes tipo 1 (DM1A y DM1B) como una única clase y añadía un tipo de diabetes híbrida que presenta un fenotipo intermedio entre DM1 y DM2 (WHO, 2019). No obstante, hoy en día la ADA sigue manteniendo la diferenciación entre DM1A y DM1B (ElSayed et al., 2025). En la DM1A, también llamada diabetes inmunomediada, existe una estrecha relación con un alelo del antígeno leucocitario humano (*human leukocyte antigen*, HLA) que provoca una elevada autoinmunidad en el paciente, por lo que casi siempre aparece una deficiencia de insulina severa. La activación del sistema inmune da lugar a la producción de autoanticuerpos contra el ácido glutámico descarboxilasa (*glutamic acid decarboxylase*, GAD), el antígeno de los islotes 2 de la tirosina fosfatasa (*islet antigen 2*, IA-2) y el transportador de zinc 8 (*zinc transporter 8*, ZnT8). Esta variante es la más prevalente, puesto que supone el 70-90% de los casos de DM1 (WHO, 2019). Por su parte, la DM1B engloba aquellas formas de diabetes que cursan con una deficiencia de insulina sin que exista evidencia de autoinmunidad en las células β . Estos síndromes se suelen denominar diabetes idiopática porque su etiología es desconocida, pero suponen un porcentaje muy pequeño de los pacientes con DM1 (ElSayed et al., 2025).

La ADA también separa a los pacientes con DM1 en tres estadios de evolución, que dependen de la presencia de alteraciones en el nivel de glucemia y la sintomatología (ElSayed et al., 2025). Estos se pueden observar en la Figura 2. Las dos primeras etapas se consideran DM1 presintomática; mientras que el tercer estadio es la DM1 sintomática.

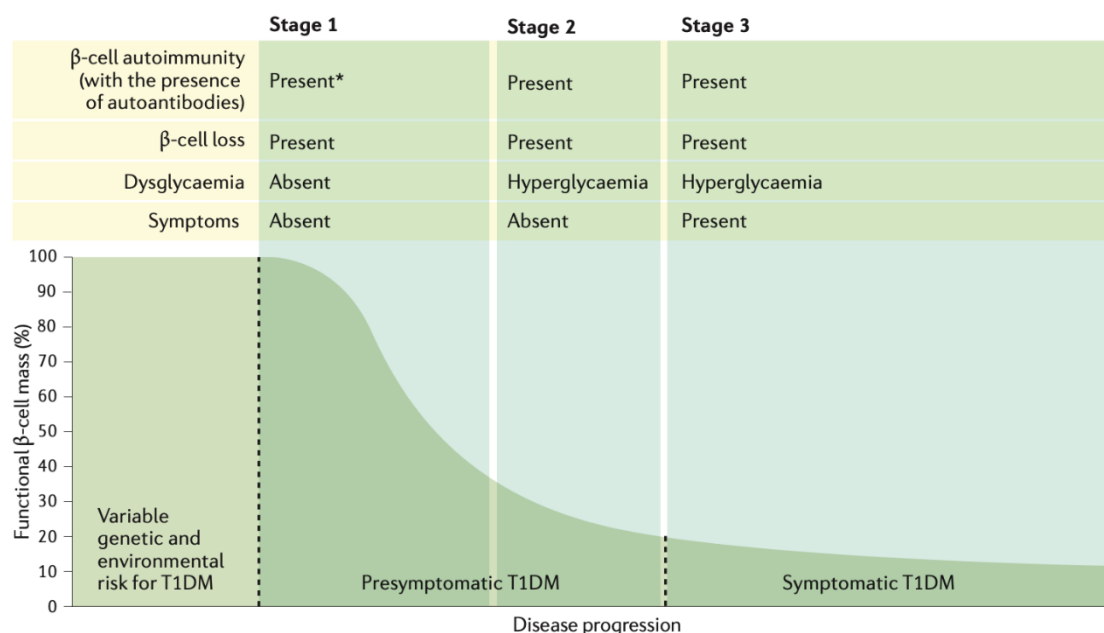


Figura 2. Estadiaje de la DM1 en función de la alteración del perfil glucémico y la presencia de sintomatología.
*Normalmente, la presencia de autoanticuerpos se manifiesta meses o incluso años antes de que empiece a haber una pérdida de células β (Katsarou et al., 2017).

En el primer estadio, se observa la presencia de autoanticuerpos, así como pérdida de células β , pero todavía no existe disglucemia. En el segundo estadio, además de la presencia de autoanticuerpos, hay una alteración del perfil glucémico. Finalmente, en el tercer estadio, a los autoanticuerpos y la hiperglucemia se suma la presencia de sintomatología (Katsarou et al., 2017). Realmente, esta separación en estadios encaja en el contexto de la DM1A, que es la que presenta autoinmunidad.

Del total de pacientes con diabetes, únicamente un 10-15% presentan DM1, de tal forma que la variante más prevalente es la DM2 (Katsarou et al., 2017). No obstante, se trata de una patología con un enorme impacto en niños, adolescentes y adultos jóvenes. De acuerdo con el Estudio de Carga Global de Enfermedades (*Global Burden of Disease*, GBD) de 2019, ha habido un incremento preocupante tanto en la incidencia como en la prevalencia de la DM1 desde 1990 hasta 2019, que es más acusado en el rango de edades de 20 a 24 años. Así, el porcentaje promedio de cambio anual (*average anual percent change*, AAPC) en la década de 2010 ha sido de 1.09 para la prevalencia y de 1.35 para la incidencia (Gong et al., 2025). No obstante, sí que es cierto que, en general, la mortalidad en jóvenes con DM1 ha descendido en los últimos 30 años, con un AAPC en 2010-19 de -0.48. Cabe mencionar que esta se sigue manteniendo más elevada en países con un bajo nivel socioeconómico (Gong et al., 2025). También es importante destacar que, en los últimos años, ha habido un incremento del número de pacientes adultos que

sufren DM1 debido en gran medida al incremento de la esperanza de vida (Tomic et al., 2025).

La región europea es la que ha tenido un mayor incremento de incidencia y prevalencia de DM1 en los últimos 30 años (Gong et al., 2025). Concretamente, en España ha habido una ratio de incidencia de 17.47 en niños y adolescentes en el período de 2013 a 2022 (Ruiz-Grao et al., 2024). De hecho, aunque algunos países nórdicos, como Finlandia, Suecia y Noruega presentan tasas de incidencia muy por encima, España se encuentra dentro de los países europeos con una mayor incidencia de pacientes con DM1. Por tanto, se trata de una enfermedad que, en nuestro país, presenta una gran relevancia tanto a nivel clínico como de investigación.

2.2 Etiología y factores de riesgo

A pesar de que solo un 10-15% de los pacientes con DM1 tienen un familiar de primer o segundo grado con presencia de la patología, el riesgo de padecer la enfermedad se ve incrementado en gran medida cuando un pariente cercano la presenta. El riesgo genético de desarrollar DM1 es debido en un 40-50% de los casos a la presencia de ciertos haplotipos del HLA. Concretamente, los alelos polimórficos HLA-DR4-DQ8 y HLA-DR3-DQ2 son los dos mayores factores de riesgo en el mundo occidental asociados con la DM1, puesto que están presentes en un 90% de los pacientes con esta patología frente a un 40% en los individuos sanos (Paschou et al., 2018). Además, estos haplotipos también están relacionados con los autoanticuerpos dirigidos contra las células β , lo que también puede inducir al desarrollo de DM1. En función de cuál de los dos esté presente, aparecerá primero un autoanticuerpo u otro. Mientras que el haplotipo HLA-DR3-DQ2 se relaciona más con la aparición precoz de autoanticuerpos tipo GAD; el HLA-DR4-DQ8 se asocia con anticuerpos que tienen como diana directa la insulina (Katsarou et al., 2017).

A pesar de que el principal factor genético determinante en el desarrollo de la DM1 es la presencia de alelos complejos del gen que codifica para el HLA, en los últimos años se han descubierto otros que también pueden tener cierta influencia (Paschou et al., 2018). El otro gran factor de riesgo identificado, que supone el 10% de los casos con predisposición a DM1, son los polimorfismos en la región del promotor para el gen que codifica la insulina. Aunque esta región no codifica para la proteína en sí, es esencial para

la formación de la proinsulina. Se trata de una región de repeticiones en tándem variables (*variable number of tandem repeats*, VNTRs) que incluye dos clases comunes: la clase pequeña o tipo I y la clase grande o tipo III. Esta última supone un factor de protección en el desarrollo de DM1, porque contribuye a la destrucción de los linfocitos T implicados en la destrucción de las células β . Otro polimorfismo que predispone a la aparición de DM1 se produce en el gen CTLA-4 (*cytotoxic T lymphocyte antigen-4*). En un individuo sano, este gen codifica para el receptor CTLA-4, que se une a moléculas coestimuladoras de los linfocitos T para evitar la activación excesiva de estas células inmunitarias. En cambio, cuando aparece el polimorfismo, esta inhibición no se produce. Otros factores genéticos estudiados son los polimorfismos del gen PTPN22 (*protein tyrosine phosphatase non-receptor type 22*), mutaciones del gen AIRE (*autoimmune regulator*), alteraciones en el gen del factor de transcripción FoxP3, mutaciones *de novo* en el factor de transcripción STAT3, polimorfismos de nucleótido único (*single nucleotide polymorphisms*, SNPs) del gen IFIH1 (*interferon-induced helicase domain 1*), supresión del gen HIP14 (*Huntingtin-interacting protein 14*) y SNPs del gen ERBB3 (*Erb-B2 receptor tyrosine kinase 3*).

Aparte de la predisposición genética que el individuo pueda tener hacia la DM1, también tienen bastante influencia los factores ambientales. Esto queda demostrado en gemelos monocigóticos, en los que la aparición de la enfermedad en ambos tiene una incidencia entre el 50 y el 100%, sin llegar nunca a este último valor (Paschou et al., 2018). Uno de los principales factores ambientales son los virus, especialmente la rubeola y los enterovirus, cuya presencia durante el período fetal puede activar mecanismos de la inmunidad innata de los niños. Además, también se ha observado que podrían estar implicados mecanismos de imitación molecular, ya que los antígenos de los virus son muy similares a los autoantígenos que son diana del sistema inmune de pacientes con DM1. No obstante, algunos enterovirus se han presentado como un factor de protección en ciertos estudios experimentales. Por otra parte, el papel de la dieta en el desarrollo de la DM1 también se ha estudiado. Así, se observa que una introducción precoz en la dieta de leche de vaca o cereales, la exposición a nitratos ante la ingesta de agua, una toma inadecuada de omega-3 y el déficit de vitamina D pueden predisponer a la aparición de DM1. También se ha observado que los pacientes con DM1 presentan una microbiota intestinal alterada, con una ratio Firmicutes/Bacteroidetes reducida. En contraposición, hay ciertos factores protectores relacionados con el embarazo, como la producción de un parto natural, que supone un contacto con la microbiota vaginal de la madre, o la presencia

de DMG, que expone al feto a elevados niveles de insulina (Katsarou et al., 2017; Paschou et al., 2018). En cualquier caso, como en la mayoría de las enfermedades, el desarrollo de diabetes está inducido por la acumulación de diversos factores más que por la presencia de uno de ellos de forma aislada.

En los estadios iniciales de la patología, se ha observado una reacción inflamatoria mediada por macrófagos, células dendríticas y linfocitos T y B. En concreto, los linfocitos T son CD4⁺ y CD8⁺, específicos de los autoantígenos de las células β . Aunque no está del todo claro cuál es la relación entre la inflamación y la apoptosis de células β , se sabe que hay una implicación de diversas citoquinas proinflamatorias en el proceso, como IL-1 (interleucina 1), TNF- α (factor de necrosis tumoral α) e IF- γ (interferón γ) (Paschou et al., 2018). Este proceso se puede prolongar durante meses o años hasta que empieza a aparecer la sintomatología en el estadio 3, en el que se pueden detectar en sangre todos los tipos celulares implicados en el proceso inflamatorio. Ante ello, se puede constatar que es necesaria una prolongada autoinmunidad contra las células β antes de que surjan los síntomas de la enfermedad (Katsarou et al., 2017; Paschou et al., 2018). Aunque esto suele asociarse a la muerte de células β , también cabe la posibilidad de que exista un malfuncionamiento de estas causado por la presencia de inflamación en los islotes de Langerhans. Es importante remarcar que los linfocitos T CD4⁺ y CD8⁺ son activados gracias a la presentación de antígenos específicos de las células β por parte de los linfocitos B y las células dendríticas.

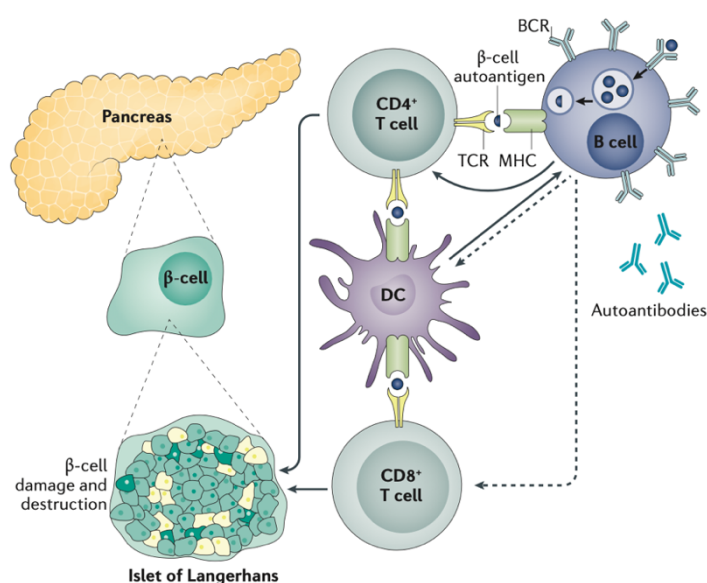


Figura 3. Patogénesis de la DM1 mediada por la respuesta celular y humoral del sistema inmunológico. DC, células dendríticas; TCR, receptor de los linfocitos T; BCR, receptor de los linfocitos B; MHC, complejo mayor de histocompatibilidad (Katsarou et al., 2017).

A nivel humoral, la activación de los linfocitos B debida a los antígenos de las células β da lugar a la producción de autoanticuerpos. En este sentido, los principales autoanticuerpos que se han detectado en pacientes con DM1 están dirigidos contra la insulina, el GAD (70% de los pacientes), el IA-2 (60% de los pacientes) y el ZnT8 (60-80% de los pacientes) (Paschou et al., 2018). Normalmente, los primeros en aparecer son los de la insulina y el GAD; mientras que los dirigidos al IA-2 aparecen posteriormente y están más relacionados con la progresión de la enfermedad. Por otra parte, los autoanticuerpos contra el ZnT8 han sido descubiertos de forma más reciente.

2.3 Presentación clínica

Aunque el signo más precoz de DM1 es la desregulación del nivel de glucemia, el cuadro clínico normalmente no se presenta hasta que la enfermedad está más avanzada, momento en el que empiezan a aflorar los síntomas. La presentación clínica de diabetes está basada normalmente en los denominados síntomas clásicos o la CAD (Gregory et al., 2013).

La tríada de síntomas clásicos que suele aparecer en niños y adolescentes es poliuria, polidipsia y pérdida de peso. Los dos primeros son los más habituales, presentes conjuntamente en un 97% de los casos; mientras que la pérdida de peso afecta a un 46% (Roche et al., 2005). Otro síntoma importante es la enuresis, especialmente en niños menores de 10 años, que puede ir acompañada de nocturia. Además, los pacientes pueden quejarse de fatiga o visión borrosa. La aparición de sintomatología en niños suele ser bastante abrupta y la duración es variable, desde unos pocos días hasta varios meses (Roche et al., 2005).

La CAD tiene una incidencia variable dependiendo del país, pero es uno de los síntomas de presentación más habituales en pacientes con DM1 *de novo*, tanto pediátricos como adultos. Se trata de una condición aguda que surge ante la deficiencia severa de insulina y que se manifiesta en forma de hiperglucemia, cetosis y acidosis metabólica. De hecho, es una de las principales causas de muerte en niños menores de 15 años con DM1 recién diagnosticada (Peng et al., 2021). La CAD suele ir acompañada de la sintomatología habitual, aunque es frecuente que provoque náuseas, vómitos, deshidratación, somnolencia y dolor abdominal. En casos graves pueden aparecer eventos trombóticos, edema cerebral e incluso llegar al coma. Algunos factores de riesgo que predisponen a la aparición de CAD son la temprana edad – por debajo de 2 años –, el bajo

nivel socioeconómico, las minorías étnicas, el acceso limitado a servicios médicos, la falta de seguro médico o la ausencia de parientes de primer grado con DM1 (Gregory et al., 2013; Peng et al., 2021).

Así pues, es una minoría de los casos en los que la hiperglucemia es el primer signo que se detecta de la enfermedad. No obstante, la monitorización de la glucosa en sangre cada vez es más común en los hogares. En consecuencia, en ocasiones algún familiar con diabetes puede realizar un análisis de glucosa en sangre y detectar hiperglucemia aun en ausencia de síntomas. También puede ocurrir que se detecte glucosuria al realizar una prueba en pacientes que llegan con molestias totalmente diferentes, como una candidiasis vulvovaginal (Gregory et al., 2013).

2.4 Diagnóstico

La DM1 se detecta habitualmente durante la infancia y presenta picos entre los 10 y los 14 años. No obstante, el diagnóstico ocurre también en un porcentaje importante en adultos mayores de 30 años (Quattrin et al., 2023). Es habitual que el diagnóstico se retrase hasta la etapa 3 o sintomática, en la que muchos pacientes presentan CAD, que es una condición de riesgo (Katsarou et al., 2017). De hecho, la presentación de síntomas habituales junto con CAD es más habitual en pacientes más jóvenes. En ellos, generalmente el control de la enfermedad es peor.

Aunque realmente una prueba rápida de glucemia elevada o un valor anormal durante una prueba de tolerancia a la glucosa oral (PTGO) servirían para hacer un diagnóstico preciso, la mayor parte de los pacientes son identificados debido a la sintomatología asociada a la hiperglucemia (poliuria, polidipsia y pérdida de peso) (Quattrin et al., 2023). No obstante, para confirmar el diagnóstico inicial derivado de los síntomas es necesario realizar varias PTGO. En este contexto, existen una serie de criterios diagnósticos establecidos por la ADA en 2016, basados principalmente en signos de un metabolismo anormal de la glucosa (Gregory et al., 2013; Katsarou et al., 2017). Así, para un diagnóstico fiable, el paciente debe cumplir al menos uno de ellos. Los criterios son los mostrados a continuación:

- Un nivel de HbA1c mayor o igual a 6.5%, que debe ser registrada con un método certificado por el Programa Nacional de Estandarización de la Glicohemoglobina.

La HbA1c hace referencia al porcentaje de hemoglobina en sangre que está unido a la glucosa, de forma que es una medida indirecta del nivel de glucemia en sangre.

- Una prueba rápida de glucosa en plasma que sea igual o superior a 126 mg/dL repetida en 2 ocasiones. Esta prueba se debe realizar con la persona en un ayuno de al menos 8 horas, sin que haya habido ingesta calórica.
- Una medida de PTGO igual o superior a 200 mg/dL, realizada dos horas después de una carga de glucosa de 1.75 g por kg de peso, siempre que no se supere la dosis de 75 g. En caso de que el resultado no sea concluyente, es necesario realizar repetidas pruebas para confirmar el diagnóstico.
- Una medida aleatoria de glucemia en sangre venosa superior a 200 mg/dL, acompañada de los síntomas típicos de la hiperglucemia o de una crisis hiperglucémica.

Una cuestión que genera preocupación actualmente es la creciente incidencia de obesidad y DM2 en niños y adolescentes, lo que complica el diagnóstico de DM1 en estas etapas. No obstante, es habitual que la obesidad presente en niños con DM2 sea mucho más grave que en aquellos que presentan DM1. Además, en pacientes con DM2 también aparecen otra serie de síntomas como acantosis nigricans, hipertensión o dislipemia, que son poco comunes en la DM1 (Gregory et al., 2013). La DM1 suele aparecer en la infancia, mientras que la DM2 suele aparecer tras la pubertad, aunque existen casos de adolescentes que presentan un cuadro clínico indicativo de DM2. Normalmente, en estos casos suele haber cierta autoinmunidad, por lo que suelen considerarse pacientes con diabetes híbrida (Katsarou et al., 2017). Por tanto, pese a que la presencia de autoanticuerpos se suele asociar con DM1, no es exclusiva de este tipo de diabetes. Otro factor confusor es el nivel de péptido C, que puede estar bajo en pacientes recién diagnosticados de DM2 y tener un nivel normal en el debut de la DM1. Cabe añadir que las etnias hispana, afroamericana y americana nativa suelen tener una mayor prevalencia de DM2 que otras etnias no hispanas.

Como hemos visto, el diagnóstico realizado actualmente es bastante tardío, cuando los síntomas ya están presentes e, incluso, el paciente puede padecer alguna condición aguda como CAD o edema cerebral. Por ello, es necesario establecer programas de

screening que nos permitan detectar la patología de manera precoz. En este contexto, es esencial establecer una estandarización de las pruebas de autoinmunidad, centradas en detectar autoanticuerpos contra GAD65 o IA-2. Esta estandarización, iniciada por la OMS, fue finalmente implementada en 2006 por el Instituto Nacional de Salud (*National Institute of Health*, NIH) de Estados Unidos (Katsarou et al., 2017). No obstante, estas pruebas solo nos permiten identificar la DM1 autoinmune (DM1A); mientras que la DM1 idiopática no es tan sencilla de identificar, porque no existen anticuerpos que detecten los posibles autoantígenos en este caso. Los pacientes con DM1 idiopática suelen presentar una edad más avanzada en el momento del diagnóstico, son habitualmente africanos o asiáticos, y tienen un índice de masa corporal (IMC) más elevado (Katsarou et al., 2017). También cabe considerar el HLA, que suele ser un indicador en estadios precoces de la enfermedad, pero no siempre está presente (Kahn et al., 2004). Por tanto, es necesario hacer un seguimiento continuado de los pacientes.

2.5 Prevención y tratamiento

Un punto crítico en el desarrollo de DM1 consiste en intentar prevenir la destrucción de las células β de los islotes de Langerhans. Esto es efectivo en niños que presentan un riesgo genético de aparición de autoanticuerpos. Para ello, se han probado modificaciones en la dieta, suplementos como ácidos grasos ricos en omega-3, o introducción de insulina oral, intranasal o subcutánea, todos ellos con poco éxito (Gregory et al., 2013; Katsarou et al., 2017). Un abordaje prometedor es el uso de inmunomoduladores que actúan sobre células inmunitarias y citoquinas implicadas en la destrucción de las células β . El mayor problema es que sus efectos suelen disminuir tras discontinuar el tratamiento, pero ahora mismo se valora el uso combinado de varios inmunomoduladores y la aplicación de una terapia intermitente. Al final, el uso de estos fármacos pretende aplanar la curva de pérdida de las células β , de tal forma que se ralentice el establecimiento de la patología (Gregory et al., 2013). Otra opción es realizar una repoblación de células β por medio de productos fabricados en el laboratorio. Estas dos opciones se consideran una prevención secundaria, porque son empleadas cuando el paciente ya presenta autoinmunidad, pero todavía no padece los síntomas típicos.

Sin embargo, los métodos de prevención todavía están en vías de desarrollo, de forma que la mayor parte de pacientes que llegan al hospital presentan sintomatología. El abordaje en estos casos va a depender en gran medida de la gravedad del paciente, de

modo que son mucho más críticos aquellos que presentan una descompensación metabólica. Además, es esencial la participación de diversos especialistas (endocrinólogo, enfermeras, nutricionista, trabajadores sociales, psicólogos y educadores en diabetes), así como de la familia, el propio paciente y diversos círculos de apoyo, como la escuela o el trabajo.

Lo más importante en este tipo de pacientes es monitorizar los niveles de glucemia de manera continua. Por tanto, hay que enseñar tanto a los pacientes como a sus familiares a usar un glucómetro para medir la glucemia en sangre desde casa. Esta medida se debe realizar antes de las comidas, antes de acostarse y previo al ejercicio. Además, también es necesaria cuando hay sospecha de hipoglucemia, después de tratar bajos niveles de glucemia y antes de tareas importantes como conducir. De hecho, la detección de hipoglucemia es esencial, así como una adecuada respuesta frente a ella. Normalmente, cuando se entra en hipoglucemia es suficiente con una pequeña dosis de azúcar (15 g), aunque no siempre, lo que hace necesaria la monitorización posterior del nivel de glucemia. Como referencia, los niveles de glucosa en sangre de una persona sana en ayunas son de 80-130 mg/dL y nunca se elevan más de 180 mg/dL en el pico posprandial (después de las comidas) (Katsarou et al., 2017).

Otra de las tareas relevantes que el paciente y la familia deben aprender es a cargar e inyectar las dosis de insulina, que normalmente se administra de forma subcutánea. De hecho, la terapia con insulina es el principal tratamiento existente hoy en día, que permite a los pacientes tener un buen control glucémico, es decir, mantenerse en niveles de euglucemia. Para hacer una calibración inicial de la ratio de insulina/carbohidratos (I:C) que necesita cada paciente, normalmente se hace una estimación de la dosis diaria total (DDT), que suele situarse en 0.5-1 unidades/kg/día de insulina en pacientes con diabetes de larga evolución (Gregory et al., 2013). Luego, este valor se va titulando de manera empírica para cada uno. En base a ello, se establecen las dosis de insulina, que van dirigidas a reducir la duración y magnitud de los picos de hiperglucemia, así como facilitar la metabolización y almacenaje de los carbohidratos ingeridos. El abordaje más habitual es el régimen de “bolo basal” o “inyección diaria múltiple”, en el que se emplean un par de dosis diarias de insulina de acción lenta para mantener el nivel basal de insulina en ayuno y dosis de insulina de acción rápida para compensar los carbohidratos de las comidas. Este método permite una mayor flexibilidad en cuanto a las cantidades y horarios de las comidas. Aunque más rígido en este aspecto, una alternativa es el régimen

“de insulina mixta/dividida”, en el que se emplea una mezcla de insulinas de acción corta e intermedia que se administra en dos dosis: una antes del desayuno y otra antes de la cena. La principal ventaja es que reducimos el número de dosis diarias.

En caso de que el paciente llegue con una descompensación metabólica, normalmente CAD, hay que trasladarle con urgencia a una unidad de cuidados intensivos. La CAD se detecta habitualmente con niveles de glucemia superiores a 200 mg/dL acompañados de un pH venoso inferior a 7.3 o un nivel de bicarbonato por debajo de 15 mmol/L (Gregory et al., 2013). Lo primero que hay que hacer es detectar posibles complicaciones asociadas, normalmente edema cerebral, y llevar a cabo fluidoterapia para restablecer los niveles de agua y electrolitos. Una vez que el paciente está estable, se inicia una inyección intravenosa de insulina de manera continua para normalizar la glucemia y eliminar la cetogénesis y lipólisis. Para evitar que el paciente caiga en hipoglucemia, se puede combinar la insulina con dextrosa, ya que la hiperglucemia suele corregirse antes que la cetoacidosis.

Una vez el paciente ha sido diagnosticado y el tratamiento iniciado, es necesaria una monitorización por parte del especialista endocrinólogo, ya que se trata de una enfermedad crónica. Para ello, es esencial que el paciente realice 4 o más medidas de glucosa en sangre, puesto que además esto está asociado con un mejor control glucémico (Gregory et al., 2013). Una medida que resulta muy útil es la HbA1c, que nos da una medida promedio de los niveles de glucemia de los tres meses anteriores. Esto es porque la vida media de los glóbulos rojos es de 3 meses. Normalmente, el objetivo es mantener unos niveles de HbA1c por debajo de 7.5% en niños y adolescentes; mientras que en adultos se han propuesto umbrales de 6.5-7%. En el caso de las embarazadas el criterio es más estricto, de modo que la HbA1c debe situarse por debajo del 6% (Katsarou et al., 2017). Otros aspectos que se deben cuidar son la dieta, la hipoglucemia, el ejercicio y las enfermedades. Todos ellos son factores críticos que pueden alterar los niveles de glucemia, de modo que son situaciones que los pacientes con DM1 deben aprender a manejar. Todo esto puede suponer un estrés para el paciente debido a la cantidad de ajustes que tienen que implementar en su vida diaria. De hecho, en el caso concreto de los pacientes pediátricos, hay una tendencia a desarrollar trastornos psicológicos, especialmente ansiedad y depresión (Gregory et al., 2013).

Para una dosis más precisa de insulina, existen actualmente bombas de insulina en las que el paciente introduce los carbohidratos que va a ingerir y, directamente, el bolo de insulina es calculado automáticamente (Gregory et al., 2013). Con estos sistemas, el valor basal de insulina se puede ir ajustando en diferentes momentos del día, de forma que se ajuste a los ritmos circadianos del paciente. Si estos sistemas de inyección automática de insulina se combinan con un sensor de glucosa, obtenemos lo que comúnmente se llama “páncreas artificial”. Estos sistemas mejoran a diario para detectar la insulina que hay que introducir al paciente en base a sistemas de inteligencia artificial que analizan las excursiones de glucosa y la actividad del paciente. Aquí, los principales aspectos a mejorar es conseguir medir la glucosa en sangre en vez de glucosa intersticial, y modificar la inyección de insulina para que no sea subcutánea, cuya acción es más lenta que si fuera intravenosa, ni periférica, lo que reduce su actividad en el hígado. También tenemos el problema añadido de la posibilidad de desarrollar resistencia a la insulina (Powers, 2021).

2.6 Enfermedades asociadas y complicaciones

Debido al carácter fuertemente autoinmune de la DM1A, que es la variante más prevalente de la DM1, existen diversos trastornos autoinmunes que suelen aparecer asociados. Uno de los más importantes es la enfermedad de Addison, donde la incidencia asciende de 1 de cada 10000 en la población general a 1 de cada 200 en pacientes con DM1 (Kahn et al., 2004). Normalmente, estos pacientes presentan autoanticuerpos contra la 21-hidroxilasa, aunque estos pueden permanecer latentes y no dar lugar a la patología. También hay una fuerte asociación con la enfermedad celiaca, que afecta al 5% de los pacientes con DM1A, en comparación con un 0.5% de la población general (Kahn et al., 2004). En este caso, el autoanticuerpo presente es contra la transglutaminasa, que aparece en un 10% de los pacientes con DM1A. Los pacientes que presentan este trastorno no pueden tomar cereales, lo que se añade a las restricciones dietéticas habituales de los pacientes diabéticos. Tanto la enfermedad de Addison como la enfermedad celiaca también están asociadas con haplotipos del gen HLA. Además, un cuarto de los pacientes con DM1A presentan autoinmunidad en el tiroides con presencia de autoanticuerpos (Kahn et al., 2004). No obstante, rara vez progresa a una enfermedad tiroidea (insuficiencia tiroidea o enfermedad de Graves). Por ello, se suele hacer una medición rutinaria de la tirotropina para detectar la aparición de este tipo de patologías.

Aparte de los diversos trastornos autoinmunes, la DM1 está asociada con varios síndromes autoinmunes. Uno de ellos es el síndrome autoinmune poliglandular tipo 1 (*Autoimmune Polyendocrine Syndrome type I*, APS-I), que se desarrolla en la infancia e incluye alguno o varios de los siguientes trastornos: candidiasis mucocutánea, enfermedad de Addison e hipoparatiroidismo. Es un síndrome poco prevalente, pero un 18% de los pacientes que lo padecen acaban desarrollando DM1 (Kahn et al., 2004). Otro ejemplo es el síndrome autoinmune poliglandular tipo 2 (*Autoimmune Polyendocrine Syndrome type II*, APS-II), que cursa con dos o más de los siguientes trastornos autoinmunes: enfermedad de Addison, enfermedad de Graves, DM1A y tiroiditis. De nuevo, está asociado con alelos del gen HLA. Por último, tenemos el síndrome de inmunodesregulación, poliendocrinopatía y enteropatía ligada al cromosoma X (*X-linked syndrome of polyendocrinopathy, immune dysfunction and diarrhea*, XPID). El síndrome XPID va ligado a DM1, enteropatía con atrofia vellositaria, dermatitis crónica, inmunodeficiencia variable y muerte infantil. En él, existe una mutación en el gen *Foxp3* que es imprescindible para la producción de linfocitos T reguladores CD4⁺ y CD25⁺ (Kahn et al., 2004).

Una vez que el paciente ha desarrollado DM1, la presencia continuada de hiperglucemia puede dar lugar a diversos tipos de complicaciones, que podemos separar en microvasculares y macrovasculares. Realmente, estas complicaciones resultan de una combinación multifactorial de la susceptibilidad genética a la exposición de glucosa, los cambios epigenéticos inducidos por la hiperglucemia, la dislipemia asociada, y diversas rutas celulares (Powers, 2021). Además, son la principal causa de la morbilidad en pacientes con DM1.

Las principales complicaciones microvasculares son la nefropatía, retinopatía y neuropatía. La nefropatía inicia con un cuadro de microalbuminuria, que es una ratio de albumina/creatinina de 30-299 mg/g en una muestra de orina (Gregory et al., 2013). Si no es correctamente tratado, este estadio inicial puede avanzar hacia una proteinuria y, finalmente, desembocar en una enfermedad renal. Por ello, la ADA recomienda realizar un *screening* para detectar microalbuminuria en pacientes a partir de los 10 años que llevan al menos 5 años de desarrollo de la enfermedad (Gregory et al., 2013). Para evitar el deterioro del filtrado glomerular, son útiles los inhibidores de la enzima convertidora de angiotensina (ECA) o los bloqueantes del receptor de angiotensina (Powers, 2021). En la retinopatía diabética hay una alteración del flujo sanguíneo en los vasos retinianos, de

modo que se promueve la generación de neovasos para compensar este defecto. El problema es que los nuevos vasos son frágiles e hiperpermeables, de modo que son frecuentes las hemorragias y la liberación de proteínas hacia la retina. Finalmente, la disminución en la perfusión de la retina puede causar ceguera (Katsarou et al., 2017). No obstante, es difícil diagnosticar esta comorbilidad cuando el paciente ha sido recién diagnosticado de DM1, de tal forma que la ADA recomienda realizar *screenings* de manera regular a partir de los 10 años y cuando el paciente ya lleva unos 3-5 años de evolución (Gregory et al., 2013). Las terapias más empleadas para remitir las retinopatías están basadas en anti-VEGF, que son fármacos que bloquean la secreción del factor de crecimiento endotelial vascular (*Vascular endothelial growth factor*, VEGF) (Powers, 2021). Por último, también puede aparecer neuropatía diabética, aunque es menos común en niños y adolescentes. De todas formas, la ADA recomienda realizar revisiones anuales a partir de la pubertad (Gregory et al., 2013). Lo que tienen en común todas las células afectadas por estas comorbilidades es que no son capaces de reducir la introducción de glucosa en situaciones de hiperglucemia, de modo que hay una sobreproducción de superóxido en las mitocondrias, lo que deriva en una situación de estrés oxidativo (Katsarou et al., 2017).

Al contrario que las complicaciones microvasculares, las complicaciones macrovasculares no son específicas de la diabetes. Principalmente incluyen la enfermedad coronaria, enfermedades cerebrales y enfermedades de las arterias periféricas, todas ellas resultado de una arteriosclerosis. Aunque la arteriosclerosis no es observable habitualmente hasta la edad adulta, sí que se ha detectado un engrosamiento de las paredes de la carótida media íntima ya desde la infancia y la adolescencia (Gregory et al., 2013). Aparte de esto, más recientemente se ha asociado un mal control en la DM1 con un incremento del riesgo de insuficiencia cardíaca, así como una alteración de las funciones cognitivas (Katsarou et al., 2017). En el proceso de la insuficiencia cardíaca también se cree que pueden estar implicadas alteraciones hemodinámica y del gasto cardíaco. Para tratar estas complicaciones, se han propuesto inhibidores del cotransportador de sodio-glucosa tipo 2 (*Sodium glucose cotransporter-2*, SGLT-2) y agonistas del péptido similar al glucagón tipo 1 (*Glucagon-like peptide-1*, GLP-1) (Powers, 2021). Los inhibidores del SGLT-2 alteran el manejo del sodio y la glucosa a nivel renal, de tal forma que tienen un efecto diurético que mejora la función arteriolar renal (Katsarou et al., 2017).

Todas las complicaciones derivadas de la DM1 se pueden evitar si el paciente presenta un buen control glucémico, pero, una vez establecidas, los daños producidos no suelen ser reversibles. Por este motivo, es esencial que los pacientes se mantengan en un rango de euglucemia y eviten picos de hiperglucemia elevados y prolongados en el tiempo. Ante ello, surgen diversos sistemas de monitorización de glucemia, esenciales para que el médico pueda tener un control sobre los pacientes y que los propios pacientes puedan reaccionar más rápido ante eventos críticos.

Capítulo 3. Control glucémico

3.1 Principales sistemas de monitorización de la glucemia.....	34
3.2 Ventajas y limitaciones de los sistemas existentes.....	36
3.3 Papel de la variabilidad glucémica en el control glucémico.....	38
3.4 Diferentes descriptores del control glucémico.....	40

3.1 Principales sistemas de monitorización de la glucemia

El sistema de monitorización de la glucemia más tradicional en pacientes diabéticos es la denominada monitorización de glucosa en sangre mediante punción digital (SMBG). Este sistema está basado en el uso de glucómetros, de tal forma que es el propio paciente el que tiene que extraer manualmente muestras de sangre capilar de los dedos. A partir de estas muestras, va registrando sus niveles de glucemia día a día, donde se recomienda un mínimo de 3-4 medidas cada día (Czupryniak et al., 2014; Jones et al., 2019).

La otra gran familia de métodos para la monitorización de la glucemia son los sistemas de monitorización continua de glucosa (CGM), en los que se mide la glucosa intersticial cada 1-5 minutos por medio de un sensor subcutáneo, por lo que son sistemas mínimamente invasivos (Marks et al., 2022). Estos sistemas constan del propio sensor, que es el que detecta los niveles de glucosa intersticial; un transmisor, que envía la señal de glucosa desde el sensor hasta un receptor, y finalmente un receptor, normalmente un dispositivo electrónico en el que se muestran al paciente los valores de glucosa. Estos sistemas generalmente duran unos 7-14 días, tiempo tras el que el paciente debe reponer el sensor. Muchos de estos sensores vienen ya calibrados, de forma que no es necesario realizar una punción digital para calibrar el dispositivo. Normalmente, no incluyen ninguna recomendación sobre las dosis o bolos de insulina, de tal forma que todas las decisiones recaen sobre el paciente en función de los valores observados. No obstante, no solo se ven los valores de glucosa, sino que también se pueden observar las tendencias de las curvas de glucosa intersticial. Esto permite predecir episodios de hiperglucemia e hipoglucemia antes de que sucedan, ante los que el paciente puede responder con un margen. Dentro de este grupo, distinguimos la CGM en tiempo real, en la que los datos de glucosa son recibidos por el paciente según se van midiendo. Esto es lo que ocurre en sistemas como Dexcom G6 o Medtronic Guardian 3. Por otro lado, tenemos los sistemas CGM flash, en los que hay que escanear el transmisor con el dispositivo móvil para poder observar los valores de glucosa. Un gran ejemplo de esta tecnología es FreeStyle Libre 2, que es el sistema de monitorización de la glucemia que presentan los pacientes incluidos en este estudio.

La tecnología de FreeStyle Libre, desarrollada por Abbott, fue aprobada por la FDA para uso clínico en septiembre de 2017 (Blum, 2018). El sistema consta de un sensor subcutáneo que se coloca en la parte superior del brazo, encargado de almacenar los

valores de glucosa intersticial del paciente cada 15 minutos y actualizar la lectura cada minuto (Bidonde et al., 2017). Para que el paciente pueda observar los datos que se están leyendo, se debe escanear el sensor con un lector, normalmente el teléfono móvil. Así, en el lector se puede observar la concentración actual de glucosa, la tendencia de los niveles de glucosa y los registros de las 8 horas anteriores en forma de tendencia. Para realizar el escaneado del sensor, el lector debe situarse a una distancia de 1-4 cm de este. No necesita ser calibrado, ya que viene calibrado de fábrica, y dura unos 14 días, momento en el que se debe sustituir y los datos deben ser cargados en la plataforma Libreview. Además, es resistente al agua, de forma que puede permanecer 30 minutos bajo una columna de 1 metro de agua.

Los datos cargados en la plataforma Libreview son observados por el profesional sanitario, que puede encontrar una gran diversidad de información (Bidonde et al., 2017), mostrada en la Figura 4. Está disponible el LogBook, que incluye las lecturas de glucosa, así como las notas del propio paciente. En estas notas se pueden incluir, por ejemplo, las causas de los eventos de hipoglucemia, el número de carbohidratos de la ingesta, medidas de glucemia capilar que se hayan realizado, etc. Al final, esto es algo muy personal de cada paciente y depende mucho de la responsabilidad individual de cada uno de ellos. También se puede observar un gráfico diario con las lecturas de glucosa, en el que se marca si estas lecturas están dentro del rango diana. Dentro de él, se observan los promedios de glucemia cada 6 horas, de forma que hay 4 medidas diarias. Además, se proporciona un patrón diario en el que se indican la mediana de glucosa en cada momento del día, así como su variabilidad por medio del primer y tercer cuartil, así como los percentiles 5% y 95%. Esto es lo que se denomina perfil ambulatorio de glucosa (*ambulatory glucose profile*, AGP), que nos permite resumir los rangos en los que se mueve la glucosa, así como su constancia, en los 14 días anteriores (Battelino et al., 2019). Como valores numéricos, se pueden ver el tiempo de diana, los eventos de baja glucemia y el porcentaje de uso del sensor. El tiempo de diana nos indica el porcentaje de tiempo que el paciente se encuentra tanto en el rango de glucosa deseado como por encima y por debajo. Normalmente, el rango diana es el denominado tiempo en rango (TIR), que abarca desde los 70 mg/dL a los 180 mg/dL. Su valor ideal debe ser superior al 70% (Battelino et al., 2019). El porcentaje del uso del sensor se refiere al tiempo que el sensor subcutáneo ha estado activo y recogiendo datos de glucosa, un valor que debe ser superior al 70%

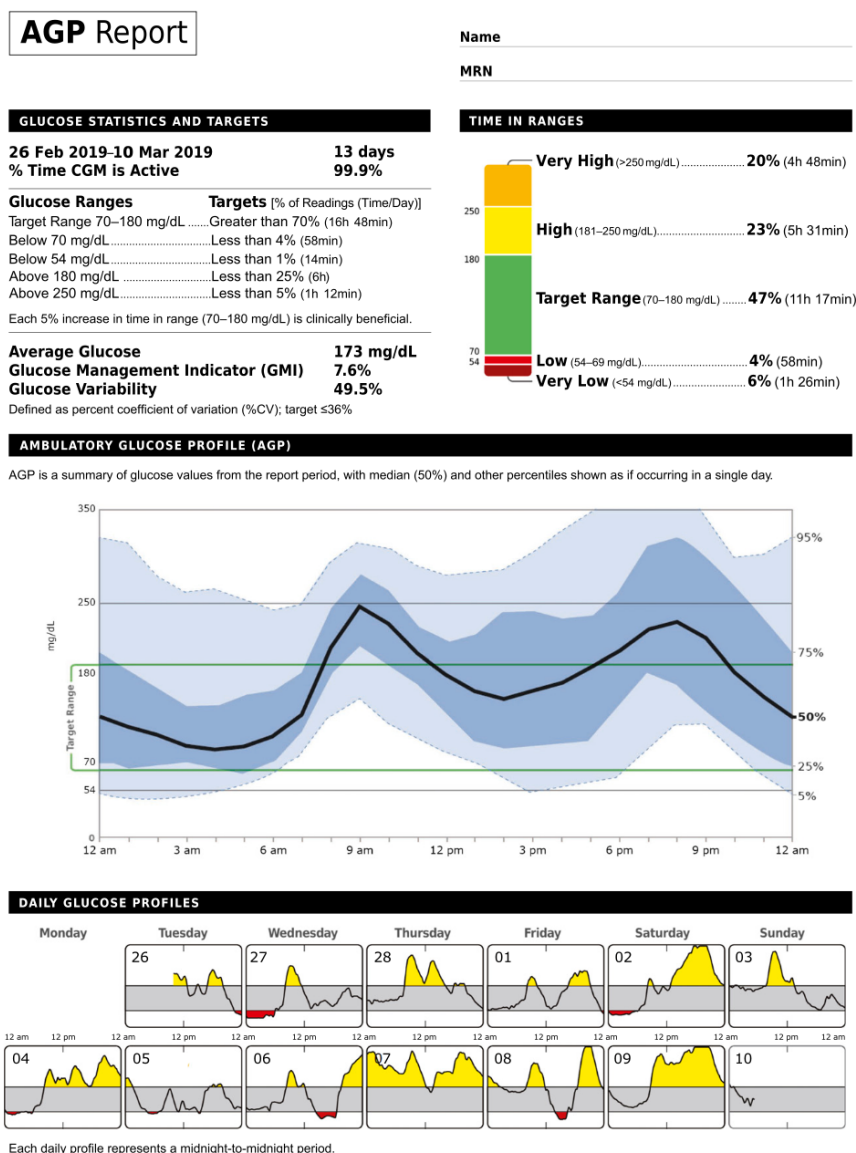


Figura 4. Interfaz observada por el médico en Libreview de los pacientes con diabetes mellitus tipo 1. Se observa un resumen de las 2 semanas anteriores que incluye parámetros de variabilidad y control glucémico (Battelino et al., 2019).

para que un resumen de 14 días pueda ser representativo de los últimos 3 meses (Battelino et al., 2019). Esto es esencial, porque permite la correlación con la HbA1c, que da un resumen de los niveles de glucosa en los últimos 3 meses.

3.2 Ventajas y limitaciones de los sistemas existentes

Actualmente, los sistemas SMBG nos permiten hacer medidas de manera rápida y eficaz, en tan solo 5 segundos, y con una muestra de sangre inferior a 2 μ L. Además, permiten calcular medias, distinguir entre glucemia pre y posprandial y almacenar hasta 500 resultados en su memoria (Czupryniak et al., 2014). En general, este tipo de

dispositivos están financiados, aunque hay variaciones según el país, incluso dentro de la Unión Europea. Sin embargo, el principal inconveniente de estos aparatos es que el paciente debe hacerse una punción digital al menos 3-4 veces al día, lo que puede acabar generando dolor en los dedos o hematomas.

Por otra parte, los sistemas CGM incluyen mucha más información, de modo que el paciente tiene un contexto más amplio para responder correctamente frente a períodos de hipoglucemia e hiperglucemia. Esto es porque aparecen unas flechas de tendencia de la glucosa que nos indican si el nivel de glucemia va a aumentar o disminuir y con qué rapidez, de tal forma que se pueden establecer unas guías de actuación que son específicas de cada sistema (Marks et al., 2022). En consecuencia, con ello se consigue corregir la dosis de insulina para evitar un infra o sobretratamiento (Karakuş et al., 2021). Además, esta tecnología incluye alarmas que suenan o vibran cuando se entra en hiperglucemia o hipoglucemia, cuyos umbrales pueden ser ajustados por el paciente. En caso de que persista el evento, también se puede establecer un período de repetición de la alarma. Todas ellas son opcionales, aunque hay algunos sistemas que presentan una alarma por defecto que no se puede eliminar para períodos de hipoglucemia severa (<54 mg/dL) (Marks et al., 2022). Este sistema de alarmas es especialmente útil en niños diabéticos (Karakuş et al., 2021). Otra gran ventaja es que los datos del sensor pueden ser subidos por el paciente a la nube o almacenados en un ordenador, de modo que luego el médico tenga acceso a ellos sin necesidad de un gran esfuerzo por parte del paciente. También permite a los usuarios que tienen el lector en el teléfono móvil compartir sus datos con familiares y amigos (Marks et al., 2022). Así, con este sistema se consigue mejorar la calidad de vida del paciente y reducir en gran medida los niveles de HbA1c (Roze et al., 2021).

Uno de los principales problemas de CGM es que mide la glucosa intersticial y no la glucosa en sangre como SMBG, por lo que hay un cierto retraso en la detección de valores fuera de rango. Esto se debe a que hay un gradiente de glucosa desde los vasos hacia el líquido intersticial (Marks et al., 2022). Cuando los cambios de glucemia son lentos, esto no supone un problema. Sin embargo, puede ocurrir que no detecte bajadas bruscas de forma eficiente o que siga marcando una hipoglucemia tras la administración de insulina, lo que puede llevar a un sobretratamiento. Este retraso se ha visto que en la práctica puede ser de 5-40 min (Marks et al., 2022). También ocurre que hay una cierta discrepancia entre las medidas observadas con CGM y SMBG, de tal forma que muchas veces la

precisión de los sistemas de CGM no es perfecta (Karakuş et al., 2021). No obstante, estas diferencias son siempre inferiores al 10% (Marks et al., 2022). Aunque en un inicio se creía que igualmente era necesario usar sistemas SMBG para la calibración de los dispositivos, hoy en día la mayoría de ellos vienen ya calibrados de fábrica (Marks et al., 2022). Otra cuestión problemática es el coste, ya que el precio de los sistemas CGM es más elevado y no siempre están financiados (Karakuş et al., 2021; Roze et al., 2021). Por último, al iniciar el uso de estos sistemas, muchos niños tienen problemas para la inserción del sensor. A ello se suman problemas con la adhesión en casos de excesiva sudoración o al ir a nadar en verano (Karakuş et al., 2021).

En particular, nos centramos de nuevo en el sistema FreeStyle Libre, que es aquel que emplean los pacientes de nuestro estudio. Al igual que el resto de los sistemas de CGM, nos da información de la glucosa intersticial de manera continua, lo que permite al paciente ajustar la dosis de insulina en base a los valores de glucosa y las tendencias. Tampoco es necesario realizar una punción capilar diariamente para el calibrado, ya que el sistema viene calibrado de fábrica. Además, el sensor es muy pequeño, de forma que las molestias se reducen al mínimo, se puede esconder debajo de la ropa, y es resistente al agua. Las principales desventajas son similares a las de otros sistemas de CGM, puesto que presenta cierto retraso en las medidas con respecto a la glucemia capilar y puede generar irritación en la piel. Su coste también es una barrera para el uso si bien el sistema se encuentra financiado en el Sistema Nacional de Salud para todo paciente con DM1, ya que hay que sustituirlo cada 14 días. Aunque no debería haber problemas en personas con sobrepeso, puede haber limitaciones en la medida en personas con un peso anormalmente bajo debido a que presentan un espacio subcutáneo demasiado estrecho (Bidonde et al., 2017).

3.3 Papel de la variabilidad glucémica en el control glucémico

La interfaz de FreeStyle Libre 2 incluye varias medidas de VG que ayudan en la valoración del control glucémico del paciente. En primer lugar, tenemos un descriptor más gráfico, que es el perfil ambulatorio de glucosa (AGP), en el que se observa un resumen de los niveles de glucosa intersticial de los últimos 14 días (Bidonde et al., 2017; Rodbard, 2018). Lo que se observa es una especie de diagrama de dispersión de un día de duración, lo que puede resultar útil para detectar en qué momentos del día el paciente sufre eventos de hipoglucemia e hiperglucemia. Esto es gracias a que el TIR está marcado

en el gráfico, de modo que se puede observar cuándo el paciente está fuera del rango diana. Para ver la tendencia de glucosa con mayor precisión, aparece el valor de la mediana cada 30-60 minutos; valores que luego son conectados por medio de una curva suavizada. De esta curva se puede obtener información de la variabilidad intradiaria, normalmente caracterizada por la SD. Aparte de la mediana, aparecen el primer y el tercer cuartil (o equivalentemente, los percentiles 25% y 75%), de manera que se puede observar la dispersión de los datos de la mediana. A partir de ellas, se puede calcular el IQR. Esto sigue siendo válido incluso para la distribución marcadamente asimétrica de datos de glucosa, que suele estar más concentrada en valores fisiológicos (70-180 mg/dL), pero con una larga cola a la derecha debida a los eventos de hiperglucemia. Por último, el AGP también nos muestra el intervalo de confianza al 90%. Con esto, podemos ver los momentos del día en los que el paciente se acerca a valores extremos.

No obstante, este resumen de los niveles de glucosa no es perfecto (Rodbard, 2021). Aunque nos muestra la distribución de los valores de la glucosa, no incluye información de los principales factores que influyen los episodios de hiperglucemia o hipoglucemia. Los más relevantes son el momento, la cantidad y la duración de las comidas; el momento de inyección de las dosis de insulina y el bolo introducido, y la actividad física, que se podría mostrar con una escala gracias a un acelerómetro. Además, la información sobre los picos postprandiales queda enmascarada debido a que la hora de las comidas no es siempre la misma y estas pueden ser más o menos copiosas. Por tanto, sería recomendable realizar una normalización de las comidas, de tal forma que se sincronicen los inicios de todos los picos posprandiales y se elimine el efecto de la glucosa basal. También quedan enmascaradas las diferencias según el día de la semana, ya que muchas veces el horario y tipo de comidas no es igual entre semana que el fin de semana. Otro problema con la glucemia es que la distribución de glucosa es altamente asimétrica, de modo que es más difícil distinguir la gravedad de la hipoglucemia, cuyos valores están normalmente más comprimidos, que la de la hiperglucemia. Por último, en una versión promediada, por mucho que incluyas diversos percentiles, se pierde mucha información sobre los *outliers* o valores atípicos, que al final son los que interesan porque son los que se refieren a los episodios de hiperglucemia o hipoglucemia graves.

Como segundo gran descriptor de la VG en el sistema de FreeStyle Libre tenemos el CV, que en este caso es de tipo puramente numérico. Normalmente, junto con la HbA1c, este es uno de los principales parámetros que se observa para determinar si el paciente

presenta un buen control glucémico. El valor que se establece como corte es un 36%, ya que el CV se expresa normalmente en porcentaje (Gómez et al., 2019). Las principales ventajas de este parámetro es que es independiente de la media de glucosa, ya que está normalizado por este valor, y que es sencillo de calcular (Mo et al., 2024). También se ha observado que está altamente correlacionado con la HbA1c y que, cuanto mayor sea, mayor es la frecuencia con la que ocurren los eventos de hipoglucemia en el paciente (Breyton et al., 2021; Gómez et al., 2019). De hecho, algunos estudios establecen el punto de corte del CV por debajo del 36% para asegurar que se reduce el riesgo de episodios de hipoglucemia (Lazar et al., 2023). Por tanto, surge como uno de los parámetros más prometedores para resumir el control glucémico del paciente. No obstante, este parámetro por sí solo falla a la hora de resumir toda la información subyacente sobre la condición del paciente, precisamente debido a su simplicidad y a que el concepto de VG es mucho más amplio (Lazar et al., 2023).

En FreeStyle Libre también se presenta como índice de riesgo glucémico el indicador de manejo de la glucosa (*glucose management indicator*, GMI). Este parámetro fue propuesto en 2017 como una forma de convertir la CGM en una estimación de la HbA1c (Selvin, 2024). Para calcularlo, se emplea una regresión lineal cuyos coeficientes fueron determinados en base a un estudio realizado en una población determinada. No obstante, una vez que nos salimos de esta población objetivo con la que se determinaron los coeficientes, se pueden observar grandes discordancias entre GMI y HbA1c. Por ello, esta conversión es un poco limitada para resumir el control del paciente y no es equivalente a lo que se observa con la HbA1c. Ante ello, otros parámetros de riesgo o control glucémico pueden ser útiles, y muchos de ellos caen dentro de la categorización de la VG.

3.4 Diferentes descriptores del control glucémico

Aunque el parámetro por excelencia que se emplea a nivel clínico es el CV, han surgido infinidad de otras métricas que intentan describir diferentes aspectos de la VG. Una de las más estudiadas es la amplitud media de excursiones de glucosa (*mean amplitude of glucose excursions*, MAGE), que es un parámetro que resume las fluctuaciones entre mínimos y máximos (Breyton et al., 2021). Este parámetro requiere de un algoritmo de cálculo complejo y, realmente, no hay un consenso, por lo que muchas veces los algoritmos varían entre sí (Rodbard, 2018). Además, las oscilaciones tienen que ser superiores a la SD, por lo que excursiones que sean de poca magnitud no quedan

registradas en el parámetro, que igual pueden tener relevancia para mantener la glucosa en el rango diana (Mo et al., 2024). Algunos autores proponen que, para un buen control glucémico, el valor de MAGE debe ser inferior a 40 mg/dL, pero se ha observado que los pacientes diabéticos exceden enormemente este límite (Khurana et al., 2025). Realmente, lo que nos podría estar indicando este parámetro es una presencia de hiperglucemia o hipoglucemia en caso de estar elevado, aunque realmente no se ha observado ninguna correlación con el tiempo de hipoglucemia (Gómez et al., 2019; Lazar et al., 2023).

Otro parámetro también muy estudiado es la clásica SD (Breyton et al., 2021). Al contrario que el CV, esta variable se correlaciona estrechamente con la media de glucosa; no obstante, no existe una correlación tan fuerte con el tiempo de hipoglucemia, al igual que ocurre con MAGE (Gómez et al., 2019; Rodbard, 2018). Otro gran problema de este parámetro es que asume que la distribución de glucosa es similar a una normal, lo que no siempre se alinea con la realidad (Mo et al., 2024). También ha sido bastante analizada la media de diferencias diarias (*mean of daily differences*, MODD), que calcula la diferencia entre dos muestras de glucosa tomadas en el mismo momento de días consecutivos (Breyton et al., 2021; Lazar et al., 2023). Un valor bajo de este parámetro nos indicaría que el paciente presenta una repetición de patrones, de modo que la aplicación de tratamiento podría ser más efectiva. En este caso, se asume que los patrones de comidas son siempre los mismos, lo que no es habitual en un entorno no controlado (Rodbard, 2018). Este parámetro, en conjunto con MAGE y SD, nos da un resumen de la dispersión de los valores de glucosa (Lazar et al., 2023). Un parámetro más alineado con el análisis no lineal de series temporales es la red de solapamiento continua de acción glucémica (*continuous overlapping net glycemic action*, CONGA), en el que se calcula la SD para series temporales en las que las muestras están espaciadas un tiempo específico, n , por lo que depende de un hiperparámetro (Breyton et al., 2021; Khurana et al., 2025). El período de tiempo n se expresa en horas y, en caso de que sea $n = 24$, entonces estamos trabajando en la misma escala que MODD (Rodbard, 2018). El problema que presentan tanto MODD como CONGA es que no fueron diseñadas para ser comprensibles desde el punto de vista de la VG a lo largo del tiempo, de tal forma que son poco utilizadas en la clínica (Mo et al., 2024).

Dos índices muy empleados son el índice de baja glucosa en sangre (*low blood glucose index*, LBGI) y el índice de alta glucosa en sangre (*high blood glucose index*, HBGI). En ellos, se hace una transformación logarítmica para incrementar la fuerza de

los valores extremos de hipoglucemia e hiperglucemia, respectivamente (Mo et al., 2024). Realmente no son medidas de variabilidad como tal, sino que se trata de índices de riesgo: LBGI de hipoglucemia y HBGI de hiperglucemia. Por ello, LBGI es uno de los parámetros que mejor correlación tiene con los eventos de hipoglucemia (Gómez et al., 2019). Similar a estos dos índices, tenemos el promedio de ratios de riesgo diarios (*average daily risk ratio*, ADRR) y el índice de control glucémico (*index of glycemic control*, IGC). En el caso de ADRR, se toman los valores máximos y mínimos de glucemia de cada día para hacer un promedio; mientras que IGC es simplemente la suma de LBGI y HBGI (Breyton et al., 2021; El-Laboudi et al., 2016; Lazar et al., 2023). Ambos se consideran también índices de riesgo, pero en este caso no diferencian entre riesgo de hipoglucemia o hiperglucemia, sino que es global. También hay parámetros más sencillos, como el mínimo (MinGly) o el máximo de glucemia (MaxGly).

De parámetros de dispersión clásica, también se ha empleado el IQR y un parámetro similar a MAGE, denominado amplitud máxima de excursiones de glucosa (*large amplitude of glucose excursions*, LAGE). La diferencia entre MAGE y LAGE es que en esta última no se da un resumen de todas las excursiones de glucosa – en forma de un promedio –, sino que directamente se toma la excursión de mayor magnitud. Además, se pueden obtener índices como combinación de medidas clásicas, como el índice J (J), que se calcula a partir de la media de glucosa y la SD, de modo que es sensible tanto a cambio en la VG como a eventos de hiperglucemia (Breyton et al., 2021; Lazar et al., 2023). Por último, hay dos medidas que nos permiten caracterizar la VG frente al tiempo, que son el cambio de glucosa absoluta media por unidad de tiempo (*mean absolute glucose change per unit of time*, MAG) y el porcentaje de variabilidad glucémica (*glycemic variability percentage*, GVP) (Rodbard, 2018). MAG calcula la diferencia entre muestras consecutivas de glucosa y lo normaliza por el intervalo temporal, de modo que nos permite obtener un valor promedio del cambio de glucosa por unidad tiempo. No obstante, es muy dependiente de las unidades usadas tanto para medir la glucosa como el tiempo, así como de la frecuencia con la que se registran los valores de glucosa. De hecho, no suele estar muy correlacionada con medidas más tradicionales de VG, como MAGE o SD. Por su parte, GVP estima también la magnitud de las excursiones por medio de una comparación de distancias, donde determina cuánta distancia de más hay que recorrer entre dos medidas consecutivas de glucosa en comparación con la distancia en tiempo, es decir, de una línea horizontal. Tanto en MAG como en GVP se intenta eliminar la

dependencia del tiempo, para que no sea tan relevante la cadencia con la que el sensor mide muestras de glucosa.

A pesar de la gran cantidad de parámetros de VG existentes, así como la cantidad de estudios que han analizado su capacidad para determinar el control glucémico y eventos adversos, no hay una conclusión clara (Breyton et al., 2021). De hecho, ninguno se ha establecido como *gold standard* y, a falta de ello, el CV se mantiene como el parámetro clínico por excelencia. No obstante, lo que está claro es que la VG no se puede definir solo con uno de estos parámetros, ya que involucra más aspectos que solo la magnitud de las excursiones, la carga de hipoglucemia e hiperglucemia o la presentación de un patrón estable de comidas a lo largo del día. Por tanto, se ha considerado que una mejor aproximación para intentar resumir el estado general del paciente es no limitarse a uno de los parámetros, sino intentar utilizar varios para tener una percepción más completa.

Capítulo 4. Materiales y métodos

4.1 Introducción.....	46
4.2 Base de datos.....	46
4.2.1 Sujetos.....	46
4.2.2 Variables clínicas.....	48
4.2.3 Registros de glucemia.....	50
4.3 Extracción de características.....	51
4.3.1 Glucometrías clínicas.....	52
4.3.1.1 Intervalos de glucemia e índices derivados.....	52
4.3.1.2 Medidas de dispersión.....	55
4.3.1.3 Índices de riesgo glucémico.....	59
4.3.1.4 Medidas frecuenciales.....	60
4.3.2 Parámetros no lineales.....	61
4.3.3 Medidas de variación del péptido C.....	64
4.4 Imputación de datos.....	65
4.4 Análisis de correlaciones: redes de asociación.....	66
4.5 Selección de características.....	67
4.6 Clasificadores péptido C detectable vs. indetectable.....	69
4.8 Análisis estadístico.....	70

4.1 Introducción

En la primera parte de este capítulo, se describe de forma detallada la base de datos empleada para realizar el estudio de variabilidad y control glucémico en pacientes con DM1. Esto incluye el proceso de descarga de los datos de glucemia y de adquisición de los datos clínicos de los pacientes. También se explican los principales parámetros que se han extraído a partir de las curvas de glucemia, tanto aquellos de naturaleza más clínica como los que son propios del análisis no lineal de series temporales.

En la segunda parte de este capítulo, se abordan las diferentes herramientas que se han empleado para relacionar los parámetros extraídos con el control glucémico de los pacientes, principalmente representado por el nivel de péptido C en ayunas en sangre. Este análisis se basa en las redes de asociación como herramienta exploratoria, a partir de las que se realiza una selección de características con diversas técnicas. Finalmente, se exponen los principales clasificadores que se han empleado, así como los análisis estadísticos utilizados para estimar la eficacia de los métodos implementados.

4.2 Base de datos

4.2.1 Sujetos

La base de datos empleada fue recogida en el edificio Rondilla del HCUV y estaba compuesta inicialmente por 227 sujetos con DM1 asociados a este hospital, es decir, del Área Este Sanitaria de Valladolid y sus alrededores. Todos los pacientes incluidos en la base de datos presentan, al menos, dos años de evolución de la patología. Para generarla, se ha empleado el servicio Jimena del HCUV, en el que se guarda la historia clínica de cada paciente asociada a un número identificativo, el número de historia clínica (NHC). Aparte de la recogida de datos clínicos, también ha sido necesario realizar la descarga de los datos en crudo de glucemia de todos los pacientes. Esto es posible gracias a la utilización de Libreview, que es una plataforma online desarrollada por Abbott en la que se guardan los valores de glucemia de los pacientes diabéticos que presentan el sistema FreeStyle Libre 2 de monitorización de la glucemia. De la base de datos inicial se tuvieron que eliminar 12 sujetos porque no incluían la medida de péptido C en las analíticas de sangre, que es el principal indicador de la reserva pancreática y el referente que se ha usado para establecer el control glucémico. Otro sujeto más se retiró de la base de datos

porque tenía un repunte de péptido C con el tiempo de manera acusada, lo que no tiene sentido fisiológico, porque el nivel de este péptido solo se reduce con el tiempo en pacientes diabéticos, nunca aumenta. Por ello, la razón más probable de esta anomalía es un error en la analítica, por lo que se ha considerado conveniente no incluirlo en el estudio. Por último, se excluyeron otros dos sujetos debido a que no presentaban registros de glucemia en la plataforma de Libreview, por lo que es imposible extraer ninguna característica de sus curvas, ya que no había datos suficientes. Finalmente, la base de datos que se ha empleado en el estudio consta de 212 sujetos.

La base de datos está compuesta de dos grandes grupos de pacientes clasificados en función de su nivel de péptido C, expresado en ng/mL. Por un lado, están los sujetos con péptido C detectable, que tienen un nivel de péptido C de al menos 0.1 ng/mL, ya que las medidas por debajo de este valor ya no son fiables con las técnicas de medida existentes (Rickels et al., 2020). Por otro lado, tenemos los pacientes con péptido C indetectable, que tienen niveles de péptido C en plasma inferiores a 0.1 ng/mL. Así, en la base de datos hay 70 sujetos con péptido C detectable y 142 sujetos con péptido C indetectable. Dentro de los sujetos con péptido C detectable, podemos separar en otros dos grupos por medio de un umbral en 0.6 ng/mL, por encima del cual consideramos que la función de las células β residuales es considerable (Rickels et al., 2020). De esta forma, hablamos de pacientes con péptido C elevado por encima de este umbral y con péptido C detectable/moderado en el rango 0.1-0.6 ng/mL. En nuestra base de datos hay un total de 34 pacientes con péptido C elevado y 36 pacientes con péptido C moderado. También es importante mencionar que la base de datos contiene tanto sujetos adultos como pediátricos (menores de 18 años). En total, hay 143 sujetos adultos y 69 sujetos pediátricos.

En nuestra base de datos también hay pacientes que incluyen varias medidas del péptido C, que fueron tomadas en distintos momentos en las diversas analíticas de sangre. De toda la base de datos, 56 pacientes tienen al menos dos medidas; de estos 56, solo 34 tienen tres medidas; y, de los 34, únicamente 13 tienen una cuarta medida. De todos los pacientes que presentan un evolutivo de péptido C, 40 son pediátricos y 16 son adultos.

A continuación, en la Tabla 1, se presenta la información general sobre la base de datos, tanto estratificada en base al nivel de péptido C como las dos categorías de edad presentes.

SUJETOS						
DATOS	Indetectable		Moderado		Elevado	
	Adultos	Pediátricos	Adultos	Pediátricos	Adultos	Pediátricos
Número de sujetos	106	36	18	18	19	15
Edad (m $\pm$ SD) (años)	47.44 $\pm$ 14.62	15.86 $\pm$ 2.28	44.50 $\pm$ 15.10	13.50 $\pm$ 4.44	47.42 $\pm$ 10.86	15.60 $\pm$ 2.10
Género (H:M)	43:63	19:17	11:7	13:5	10:9	9:6
Péptido C (m $\pm$ SD) (ng/mL)	0.02 $\pm$ 0.01	0.02 $\pm$ 0.01	0.31 $\pm$ 0.18	0.41 $\pm$ 0.15	1.30 $\pm$ 0.90	1.41 $\pm$ 0.56

Tabla 1. Datos sociodemográficos y clínicos de la base de datos en función del nivel de péptido C y el grupo de edad. m: media; SD: desviación estándar (*standard deviation*); H: hombres; M: mujeres.

4.2.2 Variables clínicas

Como ya se ha comentado, todas las variables sociodemográficas y clínicas incluidas dentro de la base de datos han sido obtenidas a partir del programa Jimena IV de SACYL. En esta plataforma, cada paciente presenta un identificador único al que se asocian su historia clínica, diversas pruebas de imagen realizadas, analíticas y tratamientos que presenta actualmente o que ha tenido en algún momento. Este número es el denominado NHC, de modo que a través de él se puede localizar a un paciente de forma inequívoca. Además, para realizar un pseudoanonimizado de los datos, a cada paciente incluido en la base de datos inicial se le ha asociado un identificador del tipo “PXXX”, donde “XXX” es el número asignado al paciente dentro de la base de datos. De esta forma, no se compromete la información de los pacientes, ya que este número es aleatorio y no se correlaciona con el NHC.

Las diversas variables sociodemográficas y clínicas se pueden clasificar en diferentes categorías de acuerdo con lo que se muestra en la Tabla 2.

TIPOS DE VARIABLES

Cualitativa dicotómica nominal	Cualitativa ordinal nominal	Cuantitativa continua
Adulto/pediátrico Género Reserva pancreática	Nivel de reserva	Nivel de péptido C HbA1c Edad Fecha de péptido C

Tabla 2. Clasificación de las variables de la base de datos en función de su escala de medición.

Es importante que las variables cualitativas sean codificadas, de tal manera que sea más fácil trabajar con ellas por medio de las diferentes herramientas que se van a emplear, especialmente las redes de asociación. Esto es porque, como la mayoría de los métodos de *Machine Learning*, las redes de asociación trabajan con variables cuantitativas, de modo que no son capaces de procesar datos cualitativos. Así, las codificaciones son las siguientes:

- **Adulto/pediátrico.** En este caso, se ha creado un vector de clasificación denominado ‘isAdult’, que vale 0 cuando estamos ante un niño o adolescente y 1 cuando se trata de un adulto. En nuestro caso, se ha considerado pediátricos a todos los pacientes hasta 18 años, ya que se trata de pacientes que llevan al menos 2 años diagnosticados.
- **Género.** Se codifica como 1 a los hombres y como 0 a las mujeres.
- **Reserva pancreática.** Esta variable se ha denominado ‘isReserve’ y se codifica como 0 si el nivel de péptido C es indetectable (<0.1 ng/mL) y como 1 si es detectable (≥ 0.1 ng/dL).
- **Nivel de reserva.** Esta variable se ha denominado ‘Pancreatic reserve’ y se codifica como 0 si el nivel de péptido C es indetectable (<0.1 ng/mL), como 1 si es detectable/moderado ($0.1-0.6$ ng/mL) y como 2 si es elevado (>0.6 ng/mL).

Respecto a las variables cuantitativas, la edad se expresa en años y nos indica la edad de los pacientes actual, a fecha de enero de 2025. El nivel de péptido C está expresado en ng/mL y es una medida que se ha hecho en sangre, extraída directamente de las analíticas de los pacientes. Del mismo modo se ha obtenido la HbA1c, que se expresa en porcentaje, ya que es el porcentaje de hemoglobina que aparece unida a la glucosa. Mención aparte recibe la fecha de toma del péptido C, que es una variable de tipo ‘timetable’ en MATLAB® con el formato DD-MMM-YYYY (D: día; M: mes, Y: año). Esta variable es importante porque nos indica la fecha de la analítica de la que se han tomado el nivel de péptido C y la HbA1c, lo que es esencial para extraer los parámetros de las curvas de glucemia en ese mismo momento.

4.2.3 Registros de glucemia

Los registros de glucemia de los pacientes, que van a ser la base para extraer todas las características, tanto clínicas como de análisis de series temporales, se han descargado en crudo de la plataforma Libreview. Esta plataforma online desarrollada por Abbott está conectada con la aplicación móvil FreeStyle Libre 2, que presenta un sistema Flash de monitorización de la glucosa. En ella, se realiza una lectura del nivel de glucosa intersticial cada 15 minutos por medio de un sensor subcutáneo colocado en el brazo del paciente y la información es almacenada en la aplicación móvil (Blum, 2018). Además, el paciente puede añadir anotaciones de lo que ha comido u otros comentarios. Luego, el paciente puede cargar toda la información en la plataforma Libreview, lo que se hace normalmente cada dos semanas.

Aunque había una gran cantidad de información, nos quedamos inicialmente con la más relevante: la fecha y hora de la medida de glucemia, el valor de glucemia del paciente, las dosis de insulina de acción rápida y lenta (en caso de que los pacientes tengan plumas de insulina inteligentes), los carbohidratos ingeridos en las comidas en gramos, y los valores de glucemia capilar. No obstante, finalmente la mayoría de los datos son escasos y dependen mucho de la responsabilidad de cada paciente. Por tanto, finalmente lo único que se ha empleado en el estudio es el valor de la glucemia intersticial a lo largo del tiempo, expresada en mg/dL. Además, hay que tener en cuenta que los pacientes deben escanear el sensor con su móvil cada 8 horas para que se mantenga la conexión entre la aplicación y el sensor subcutáneo. Por esta razón, muchas veces aparecen huecos dentro de los registros de glucemia, que son imposibles de interpolar debido a que abarcan varias

horas o incluso días, y que la frecuencia de muestreo al final es muy baja – 1 muestra cada 15 minutos, lo que equivale a 1.1 mHz. Todo el acondicionamiento a partir de los datos en crudo se ha llevado a cabo con MATLAB®. Un ejemplo de cómo se verían algunos registros de glucemia se observa en la Figura 5.

4.3 Extracción de características

A partir de los datos de la glucemia de los pacientes, se han extraído diversas características que nos permiten cuantificar diversos aspectos de la VG. Para ello, se ha seleccionado un intervalo temporal de 2 semanas que sea lo más cercano posible al momento en el que está registrado el péptido C de la analítica de sangre. Se emplea esta ventana temporal porque está demostrado que la monitorización de las curvas de glucemia del paciente cada dos semanas permite resumir el promedio de los 3 meses anteriores (Xing et al., 2011). Esto es útil porque se puede comparar esta información con la de la HbA1c, que resume la glucosa en sangre promedio de los últimos 3 meses. Para el cálculo de todos los parámetros, se ha empleado el lenguaje de programación MATLAB®. Dentro

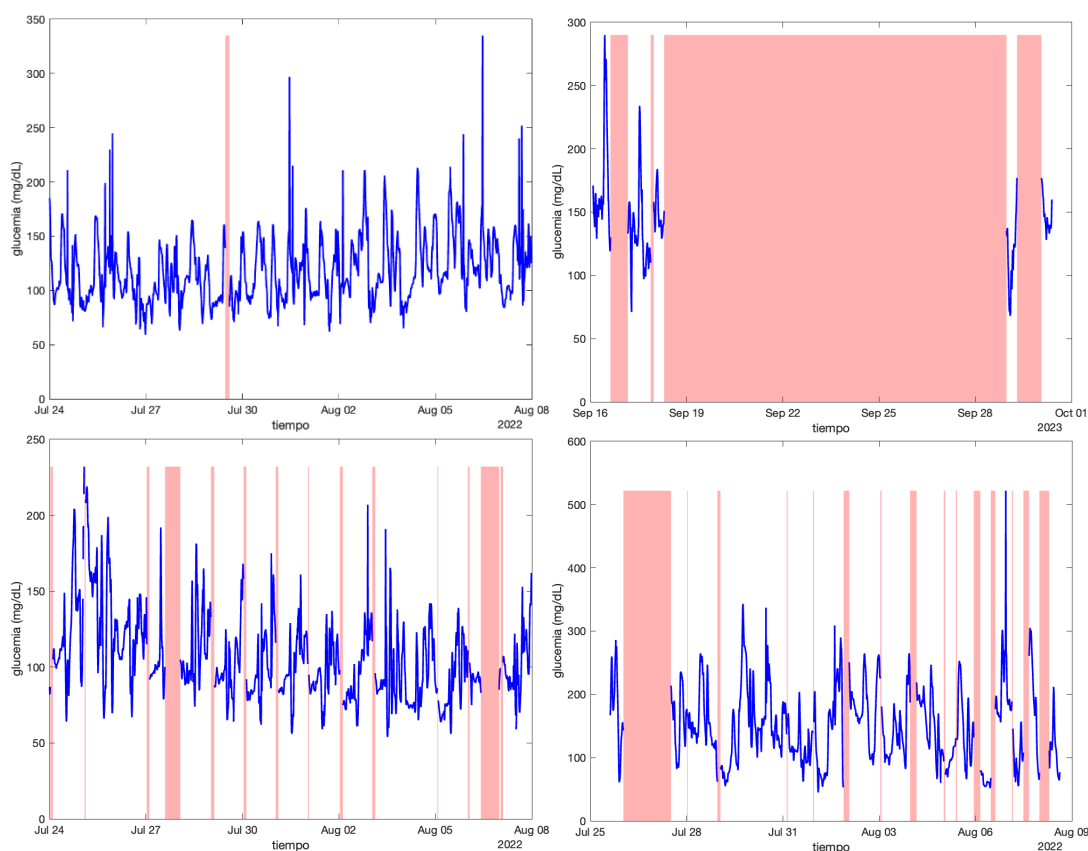


Figura 5. Registros de glucemia durante dos semanas de cuatro pacientes, en los que se observan en rojo los tramos en los que no hay datos registrados debido a una desconexión del sensor. El nivel de glucemia viene expresado en mg/dL.

de la gran cantidad de características extraídas, se pueden distinguir dos grandes grupos: glucometrías clínicas, que tienen cierto significado fisiológico, y parámetros no lineales, que son normalmente empleados para el análisis de series temporales complejas.

4.3.1 Glucometrías clínicas

4.3.1.1 Intervalos de glucemia e índices derivados

Las glucometrías clínicas más simples consisten en establecer una serie de rangos o intervalos de glucemia, de modo que se calcula el porcentaje de tiempo que el paciente permanece en cada uno de ellos. Este análisis nos permite diferenciar entre el rango de normoglucemia o euglucemia, que tradicionalmente se ha situado entre 70 mg/dL y 180 mg/dL y los rangos de hipoglucemia – por debajo de 70 mg/dL – e hiperglucemia – por encima de 180 mg/dL. En este contexto, podemos entender la hipoglucemia como un nivel bajo de glucosa en sangre y la hiperglucemia como un nivel alto de glucosa en sangre, ambos considerados perjudiciales para la salud del paciente. Así, podemos distinguir los siguientes intervalos:

- **Tiempo por debajo de rango (*time below range*, TBR).** Normalmente se suele diferenciar entre una hipoglucemia más crítica (*time below range level 2*, TBR2), que incluye valores por debajo de 54 mg/dL, y una hipoglucemia más leve (*time below range level 1*, TBR1), que incluye valores entre 54 mg/dL y 70 mg/dL. El valor de estos dos parámetros debe ser inferior al 4% para el TBR1 y al 1% para el TBR2 (Battelino et al., 2019). Esto es esencial para conseguir un buen control glucémico y evitar la presencia de hipoglucemia, lo que incrementa el riesgo de la disminución de la percepción de la hipoglucemia. Esto conlleva que los pacientes puedan sufrir neuroglucopenia e hipoglucemia severa de forma más habitual, incrementado el riesgo cardiovascular (Christou et al., 2023).
- **Tiempo en rango (*Time in range*, TIR).** Se trata del rango de glucemia en el que el paciente debe permanecer la mayor parte del tiempo para que tenga un buen control glucémico. Tradicionalmente, se ha establecido entre 70 mg/dL y 180 mg/dL, e idealmente debe ser superior al 70% (Battelino et al., 2019).
- **Tiempo por encima de rango (*Time above range*, TAR).** Al igual que con el TBR, se suele distinguir entre una hiperglucemia más acusada (*time above range*

level 2, TAR2) – que incluye valores por encima de 250 mg/dL, y una hiperglucemia más leve (*time above range level 1*, TAR1), que abarca valores entre 180 mg/dL y 250 mg/dL. Para un buen control glucémico, estos dos parámetros deben ser inferiores al 5% y el 25%, respectivamente (Battelino et al., 2019). Con ello, evitamos la presencia de hiperglucemia continuada, que puede provocar diversas complicaciones microvasculares y macrovasculares (Gregory et al., 2013; Katsarou et al., 2017; Powers, 2021).

Todos estos parámetros se obtienen en la plataforma Libreview cuando el paciente carga sus datos de las últimas dos semanas desde la aplicación FreeStyle Libre. Sin embargo, dan un valor sin decimales, de modo que se pierde parte de la precisión. Por ello, se ha realizado el cálculo de estos parámetros desde MATLAB®, donde se puede obtener una precisión con varios decimales, lo que nos permite detectar valores inferiores al 1% del TVBR, que son muy habituales en la clínica. Otro intervalo de glucemia que no está incluido en Libreview y que es cada vez más utilizado es el tiempo en rango estrecho (*time in tight range*, TITR) (Dunn et al., 2024). Este parámetro es un poco más estricto que el TIR, de modo que abarca desde 70 mg/dL hasta 140 mg/dL de glucemia. Su valor ideal para un buen control debería ser superior al 50% (Bahillo-Currieses et al., 2024). Además, se ha calculado la diferencia entre el TIR y el TITR, que se ha denominado el tiempo fuera del rango estrecho (*time out of tight range*, TOTR). La suma de todos estos intervalos debería suponer el 100%, que se puede obtener de varias maneras:

$$100\% = TBR2 + TBR1 + TIR + TAR1 + TAR2 = TBR2 + TBR1 + TITR + TOTR + TAR1 + TAR2$$

Derivado de estos parámetros, se puede calcular el denominado índice de riesgo glucémico (*glycemia risk index*, GRI) (Díaz-Soto et al., 2024; Pérez-López et al., 2023). Este índice se calcula a partir de dos parámetros, que son la carga de hipoglucemia (CHypo) y de hiperglucemia (CHyper), en los que se da más peso a los valores extremos por debajo de 54 mg/dL y por encima de 250 mg/dL que a los valores más cercanos al TIR (Klonoff et al., 2023). Este parámetro pretende sintetizar toda la información del control glucémico del paciente en un único parámetro, en el que se tienen más en cuenta los valores más extremos, que representan situaciones fisiopatológicas más críticas. Las fórmulas para calcular CHypo, CHyper y GRI, junto con un resumen de los intervalos de glucemia, se pueden observar en la Tabla 3. Además, en la Figura 6 se observa una representación gráfica de los diferentes intervalos, tanto con la aproximación clásica del TIR como con la más novedosa del TITR.

Parámetro	Cálculo	Valor ideal
TBR2	% tiempo <54 mg/dL	<1%
TBR1	% tiempo ≥54 mg/dL y <70 mg/dL	<4%
TIR	% tiempo ≥70 mg/dL y ≤180 mg/dL	>70%
TAR1	% tiempo >180 mg/dL y ≤250 mg/dL	<25%
TAR2	% tiempo >250 mg/dL	<5%
TITR	% tiempo ≥70 mg/dL y ≤140 mg/dL	>50%
TOTR	% tiempo >140 mg/dL y ≤180 mg/dL	—
CHypo	$TBR2 + 0.8 \cdot TBR1$	—
CHyper	$TAR2 + 0.5 \cdot TAR1$	—
GRI	$3.0 \cdot CHypo + 1.6 \cdot CHyper$	—

Tabla 3. Resumen de los intervalos de glucemia y los parámetros clínicos de resumen del control glucémico. TBR2: tiempo por debajo de rango nivel 2; TBR1: tiempo por debajo de rango nivel 1; TIR: tiempo en rango; TAR1: tiempo por encima de rango nivel 1; TAR2: tiempo por encima de rango nivel 2; TITR: tiempo en rango estrecho; TOTR: tiempo fuera del rango estrecho; CHypo: carga de hipoglucemia; CHyper: carga de hiperglucemia; GRI: índice de riesgo glucémico.

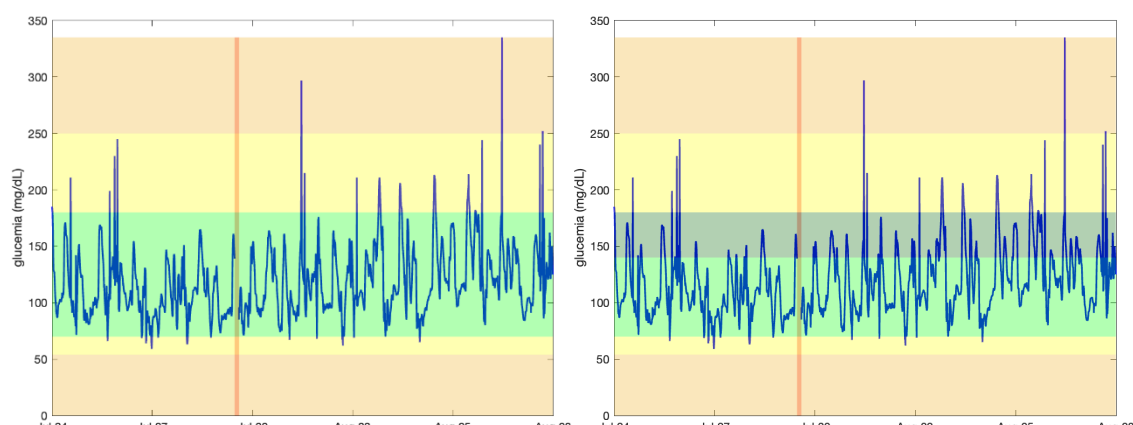


Figura 6. Registros de glucemia durante dos semanas en los que están marcados los diferentes intervalos de glucemia. A la izquierda, se observa el TIR en verde, el TAR1 y el TBR1 en amarillo por encima y por debajo, respectivamente, y el TAR2 y el TBR2 en naranja, por encima del TAR1 y por debajo del TBR1, respectivamente. A la derecha, se observa el TITR en verde claro, el TOTR en verde oscuro por encima del TITR, el TAR1 y el TBR1 en amarillo por encima del TOTR y por debajo del TITR, respectivamente, y el TAR2 y el TBR2 en naranja, por encima del TAR1 y por debajo del TBR1, respectivamente.

4.3.1.2 Medidas de dispersión

El resto de las glucometrías definidas nos permiten describir diferentes aspectos de la VG (Breyton et al., 2021; Rodbard, 2018). Todas ellas se han empleado a nivel de investigación, pero ninguna es aplicada actualmente en la clínica (Khurana et al., 2025; Mo et al., 2024). La única que aparece incluida en la interfaz de la plataforma Libreview es el CV. En cualquier caso, el resumen de cómo calcular cada una de ellas se puede observar en la Tabla 4.

Como medidas clásicas de dispersión, se encuentran la desviación estándar (*standard deviation*, SD) y el rango intercuartil (*interquartile range*, IQR). La SD mide la dispersión de los valores de glucosa (G_t) con respecto al valor de glucosa media (MG) en un cierto período temporal (Lazar et al., 2023). Por su parte, el IQR mide la variabilidad intradiaria de los niveles de glucosa, de tal forma que cuanto más amplio sea su valor, más cambios hay en los niveles de glucosa del paciente en diferentes días (Mo et al., 2024). De hecho, se ha observado que valores de IQR por debajo de 5.55 mg/dL se correlacionan con valores de TIR por encima del 70% (Tokutsu et al., 2022). En este caso, estas dos medidas están directamente implementadas en funciones propias de MATLAB®, por lo que no ha sido necesario programar código para calcularlas. En cambio, el resto de las medidas se han tenido que implementar desde cero, aunque la mayoría de ellas presentan fórmulas sencillas que no son complejas de estructurar con código.

Como ya hemos visto, el CV se calcula de manera sencilla dividiendo la SD por la MG. Normalmente se expresa en porcentaje, por lo que el resultado se multiplica por cien. Al igual que SD e IQR, la media es un parámetro que ya viene implementado dentro de las funciones de MATLAB®.

Muy distinto es el cálculo de MAGE, pues se trata de uno de los parámetros más complejos. Esto se debe a que hay que identificar los picos máximos y mínimos de la curva de glucosa para calcular la diferencia entre ellos. Estas diferencias son las que se denominan excursiones (λ), pero no todas ellas son válidas, ya que es necesario que sean superiores a la SD. Por tanto, solo nos quedamos con los extremos cuyas diferencias superan este umbral, de tal forma que todas las excursiones calculadas presentan una continuidad entre ellas. De esta manera, si unimos los extremos que se detectan,

Parámetro	Fórmula	
SD (mg/dL)	$\sqrt{\frac{1}{n-1} \sum_{t=0}^{t_n} (G_t - MG)^2}$	n: número de valores de glucosa G _t : nivel de glucosa para el tiempo t MG: nivel de glucosa medio
CV (%)	$(SD/\bar{G}) \times 100$	
MAGE (mg/dL)	$\sum \frac{\lambda}{x}, \text{ if } \lambda > SD$	λ: diferencia absoluta entre máximos y mínimos x: número de observaciones válidas
MODD (mg/dL)	$\frac{1}{k} \sum_{t=0}^{t_k} G_t - G_{t-1440} $	k: número de observaciones con otra a las 24h t: tiempo expresado en horas
IQR (mg/dL)	$Q_3 - Q_1$	Q ₁ : primer cuartil Q ₃ : tercer cuartil
LBGi	$\frac{1}{n} \sum_{i=1}^n r_l(G_i)$ $r_l(G_i) = \begin{cases} 22.77 \cdot f(G_i)^2 & \text{si } f(G_i) < 0 \\ 0 & \text{si } f(G_i) > 0 \end{cases}$	$f(G_i) = \ln(G_i)^{1.084} - 5.381$
HBGI	$\frac{1}{n} \sum_{i=1}^n r_h(G_i)$ $r_h(G_i) = \begin{cases} 0 & \text{si } f(G_i) < 0 \\ 22.77 \cdot f(G_i)^2 & \text{si } f(G_i) > 0 \end{cases}$	
ADRR	$\frac{1}{M} \sum_{j=1}^M (\max\{r_l(G_1), \dots, r_l(G_k)\} + \max\{r_h(G_1), \dots, r_h(G_k)\})$	M: número de días
MinGly (mg/dL)	$\min\{G_1, \dots, G_n\}$	
MaxGly (mg/dL)	$\max\{G_1, \dots, G_n\}$	
CONGA (mg/dL)	$SD(D), \quad D_t = G_t - G_{t-m}$	m: versión de grano grueso
Índice J ((mg/dL) ²)	$0.001 \cdot (\bar{G} + SD)^2$	
LAGE	$\max\{\lambda_1, \dots, \lambda_x\}$	
IGC	$LBGI + HBGI$	
GVP (%)	$100 \times \left(\frac{L}{L_0} - 1\right)$	L: distancia entre muestras consecutivas de glucosa L ₀ : lapso de tiempo entre muestras de glucosa
MAG (mg/dL/h)	$\frac{1}{T} \sum_{t=0}^{t_{n-1}} G_{t+1} - G_t $	T: tiempo total de registro

Tabla 4. Principales parámetros clínicos de VG.

obtenemos una versión de la señal que recoge solo la información de las oscilaciones de gran magnitud de la señal de glucemia, que se puede observar en la Figura 7. Una vez que se tienen todas las excursiones válidas, se realiza una media ponderada por la duración de cada una de ellas. Realizar esto de forma manual es muy tedioso, por lo que hay autores que proponen algoritmos automáticos (Baghurst, 2011). No obstante, en este estudio se ha optado por emplear el código de la función MAGE que hay disponible en el paquete ‘*iglu*’ del lenguaje de programación R (Broll et al., 2021). En este caso, la identificación de extremos de la función se hace por medio de la determinación de los cruces entre dos medias móviles, una corta y otra larga, de tal forma que se obvia el ruido de las tendencias prolongadas de glucosa y las fluctuaciones locales (Fernandes et al., 2022). Este código se ha traducido al lenguaje de MATLAB® para poder manejar todos los parámetros con el mismo lenguaje de programación.

Muy similar a MAGE es LAGE. En este caso, en vez de realizar una media ponderada de todas las excursiones, solo se toma aquella que tiene mayor magnitud, por lo que el algoritmo empleado es el mismo. Se trata de una aproximación un poco más refinada que la implementación tradicional, que simplemente calcula la diferencia entre el máximo y el mínimo de glucemia en un cierto período temporal (Zhang et al., 2021). Mientras que MAGE queda suavizada en el tiempo y nos indica si el paciente tiene períodos extremos de hiperglucemia o hipoglucemia, LAGE es una medida que es capaz de detectar eventos

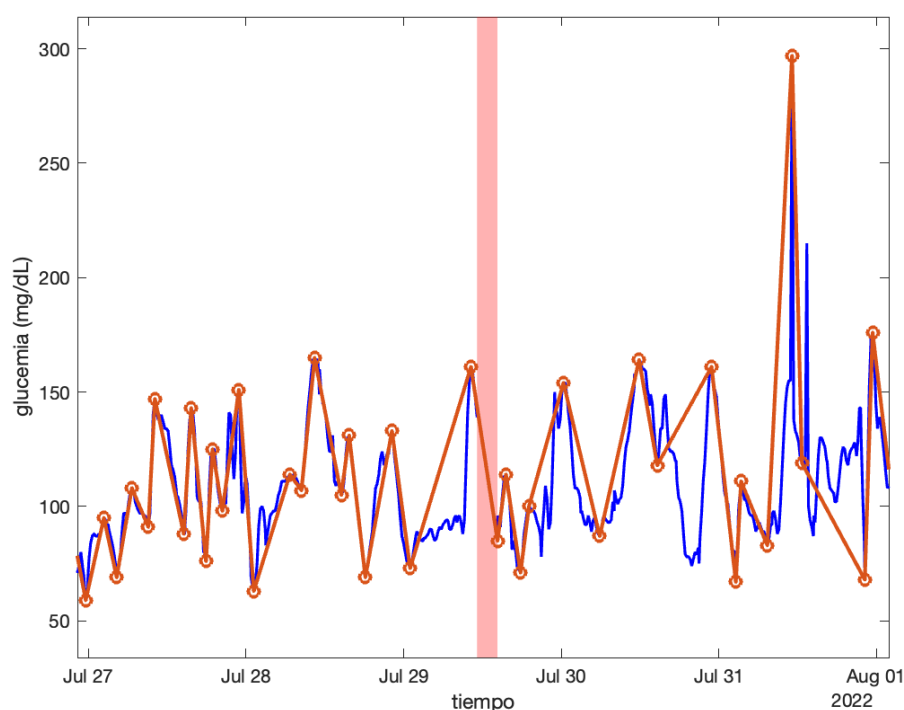


Figura 7. Gráfica de los niveles glucosa a lo largo del tiempo (en azul), sobre la que se observan las excursiones detectadas por el algoritmo de MAGE (en rojo). Los máximos y mínimos están marcados como círculos en color rojo.

más extremos. De hecho, se ha demostrado que puede ayudar en la detección de hipoglucemia nocturna asintomática (Wang et al., 2022).

MODD calcula un promedio de la diferencia en los valores de glucosa que se tomaron en el mismo momento en días consecutivos. Estas diferencias se miden siempre en valor absoluto. Al contrario de lo que ocurre con MAGE y SD, que miden la variabilidad intradiaria, MODD nos permite detectar la variabilidad interdiaria. Esto quiere decir que, cuanto más elevado sea el valor de MODD, el paciente presenta unos patrones irregulares de estilo de vida (Bowler et al., 2024). Esto implica que, en estos pacientes, es más difícil predecir la efectividad de un tratamiento con insulina, ya que cuando se intentan reducir los niveles de glucosa y, con ellos, el nivel de HbA1c, hay un riesgo elevado de provocar episodios de hipoglucemia (Lazar et al., 2023).

CONGA se calcula como una SD, lo que pasa que no directamente sobre la serie temporal de glucosa. En este caso, se generan diversas versiones de grano grueso de la señal de glucosa a lo largo del tiempo, en las que la diferencia temporal entre muestras cada vez es mayor (Breyton et al., 2021). Después, sobre cada una de esas versiones se calcula la SD, de tal forma que no tenemos un solo valor del parámetro, sino una curva de valores que nos muestra la dimensionalidad de la señal de glucosa. En este caso existe un parámetro adicional, n , que nos indica la diferencia de tiempo entre muestras consecutivas en cada versión de la señal. Normalmente las medidas empleadas en la literatura suelen ser con una diferencia de pocas horas, y generalmente nunca superan las 12 horas (Lazar et al., 2023; Sparks et al., 2022). De hecho, hay algunos intervalos más usados, como el intervalo de 1 h, que nos indica la VG global diaria; el intervalo de 2 h, que mide la VG relacionada con aperitivos (o con las comidas en pacientes que son tratados con insulina de acción rápida); el intervalo de 4 h, que refleja la VG referente a las comidas, y el intervalo de 12 h, empleado para medir la VG asociada a la insulina basal. No obstante, en este estudio se ha optado por realizar numerosas escalas – en concreto 96 – con una diferencia mínima entre muestras de 15 min (que es la frecuencia de registro del sensor FreeStyle Libre) y una diferencia máxima de 24 h. De esta forma, obtenemos una visión global de la evolución del parámetro en función de la escala empleada y, a partir de las curvas generadas, se pueden extraer parámetros no lineales como la dimensión fractal.

4.3.1.3 Índices de riesgo glucémico

Muy similares en su manera de cálculo son LBGi y HBGI. En este caso, lo primero que se hace es una transformación logarítmica de los niveles de glucosa, $f(G)$, que hace que los valores de normoglucemia queden comprimidos en torno al cero – que se corresponde con un nivel de glucemia de 112.52 mg/dL –, mientras que los valores de hipoglucemia e hiperglucemia se ven potenciados. Después, cada uno de los dos índices nos indica el riesgo de un tipo de eventos: LBGi informa sobre el riesgo de hipoglucemia; mientras que HBGI representa el riesgo de hiperglucemia (Lazar et al., 2023). De esta forma, LBGi descarta todos los valores por encima de 0, correspondientes a períodos de hiperglucemia; mientras que HBGI hace lo contrario, es decir, descarta valores menores que 0 asociados con la hipoglucemia. Los valores no descartados se elevan al cuadrado para, de nuevo, potenciar la detección de eventos adversos – ya sea hipoglucemia o hiperglucemia –, y son multiplicados por un factor. Finalmente, son promediados (Breyton et al., 2021). Para estos índices de riesgo, se han establecido una serie rangos. En el caso de LBGi, un valor por debajo de 2.5 nos indica un riesgo leve de hipoglucemia, entre 2.5 y 5 un riesgo moderado, y un valor superior a 5 es indicativo de un elevado riesgo de hipoglucemia. Por otro lado, valores de HBGI menores de 4.5 indican un riesgo leve de hiperglucemia, entre 4.5 y 9 hay un riesgo moderado de hiperglucemia y valores por encima de 9 presentan un riesgo severo de hiperglucemia (Köhlmoos & Dittmar, 2025).

También usan esta misma transformación ADRR e IGC. El parámetro ADRR consiste en tomar el mayor valor de la función de baja glucemia empleada en LBGi, $r_l(G)$, que solo considera los valores por debajo de 0, y también el mayor valor de la función de alta glucemia de HBGI, $r_h(G)$, que considera los valores por encima de 0. Estos valores se seleccionan para cada día bajo estudio, se suman y se hace un promedio dividiendo por el total de días, M . En cambio, IGC es más sencillo y simplemente se calcula como la suma de LBGi y HBGI (Breyton et al., 2021). Ambos dos reflejan el riesgo de eventos críticos, sin distinguir si son eventos de hipoglucemia o hiperglucemia. Se observa que, en el caso de ADRR, un valor elevado se correlaciona con una mayor sensibilidad a la insulina y, en este caso, es un mejor indicador de elevados niveles de glucosa (Fabris et al., 2016; Lazar et al., 2023).

Dos índices mucho más sencillos son MinGly y MaxGly, en los que simplemente se extrae el valor mínimo y máximo de glucosa, respectivamente, en un cierto intervalo de

tiempo. Al igual que LBG y HBGI, estos dos parámetros nos indican la presencia de eventos de hipoglucemia (MinGly) e hiperglucemia (MaxGly). Aunque pueden parecer parámetros muy simples, muchas veces una aproximación menos compleja nos proporciona resultados más interpretables desde el punto de vista clínico.

También basado en parámetros sencillos es el índice J, que se calcula como una combinación de SD y MG, elevadas al cuadrado y multiplicadas por un factor concreto (Breyton et al., 2021). Este parámetro, además de ser sensible a diversos aspectos de la VG – la dispersión de los valores de glucosa y las subidas de la glucosa promedio –, también es una medida de la calidad del control glucémico (Lazar et al., 2023; Leksic et al., 2022). De hecho, se ha observado que su monitorización puede ser útil en embarazos con DM1, donde un elevado valor del índice J puede ser indicativo de un neonato grande para su edad gestacional (Leksic et al., 2022).

4.3.1.4 Medidas frecuenciales

La mayoría de los parámetros descritos, aunque nos indican el riesgo de eventos adversos en las curvas de glucosa, no tienen en cuenta la frecuencia con la que estos ocurren. Para este tipo de abordaje, se han descrito variables como GVP y MAG. GVP compara la diferencia entre la distancia entre dos puntos de nivel de glucosa consecutivos, L , y la distancia en línea recta, L_0 , que básicamente es el intervalo temporal entre muestras (Rodbard, 2018). La distancia L es fácil de calcular empleando trigonometría, ya que es la hipotenusa de un triángulo formado por la variación en tiempo (Δt), la variación de glucosa (ΔG) y la propia curva de glucemia que une dos puntos consecutivos, lo que se puede observar en la Figura 8. Su cálculo se detalla a continuación:

$$L = \sqrt{\Delta t^2 + \Delta G^2}$$

Esta medida es bastante intuitiva en el ámbito médico, porque viene expresada en porcentaje. Además, se ha comprobado que la información que nos aporta es ligeramente distinta de métricas tradicionales como MAGE y CV, probablemente porque tiene en cuenta la frecuencia de las fluctuaciones de glucemia (Peyser et al., 2018). Se ha establecido un corte en 20% para diferenciar entre pacientes diabéticos y no diabéticos.

Por el contrario, el MAG no se expresa en porcentaje, lo que dificulta su interpretabilidad. En este caso, se calcula la suma de las diferencias entre muestras

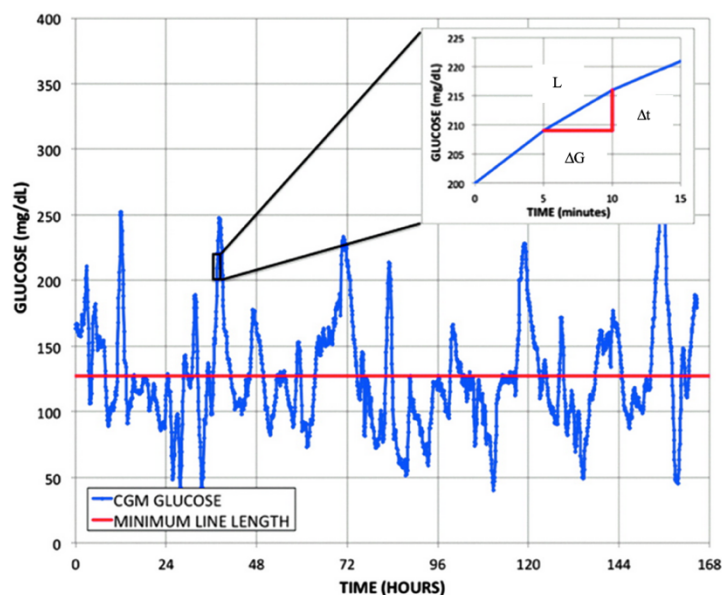


Figura 8. Visualización de la distancia entre puntos consecutivos de glucosa, L , así como su relación trigonométrica con la variación en tiempo, Δt , y la variación de glucosa, ΔG , ambas marcadas en rojo (Peyser et al., 2018).

consecutivas de glucosa, todas en valor absoluto, para después normalizar el valor con respecto al tiempo dividiendo entre el intervalo temporal bajo estudio, T (Rodbard, 2018). En esta medida es esencial la frecuencia de muestreo de la glucosa intersticial, así como las unidades escogidas para el nivel de glucosa y el tiempo, porque pueden cambiar los resultados obtenidos.

4.3.2 Parámetros no lineales

Aparte de la gran cantidad de parámetros que existen en la clínica, en este trabajo se han querido incluir diversos parámetros no lineales empleados para el análisis de series temporales complejas. Dentro de este grupo, podemos distinguir medidas de variabilidad y medidas de complejidad. Aunque estamos intentando medir la variabilidad, una estimación de la complejidad de la señal también puede ser útil para determinar el control glucémico del paciente, ya que añadimos una mayor dimensionalidad a nuestro conjunto de características.

Uno de los parámetros de complejidad por excelencia es la dimensión fractal, que es un concepto que se usa para describir objetos en la naturaleza denominados fractales. Este tipo de objetos presentan una geometría irregular que hace que su dimensionalidad sea superior al número de dimensiones topográficas que presentan, normalmente un número decimal. Un ejemplo muy habitual de fractalidad son los diferentes sistemas presentes en el cuerpo humano y, por ende, las señales biomédicas derivadas de ellos (Azizi, 2022).

Una aproximación común en el caso de los sistemas fisiológicos es la dimensión fractal de Higuchi, en la que se considera una relación entre dos magnitudes: un factor de escala, n , y un parámetro calculado para diferentes versiones de grano de la señal. Para poder determinar la dimensión fractal, lo más habitual es trabajar en una escala logarítmica, de tal forma que la pendiente que relaciona las dos variables es una estimación de la dimensión fractal (Armonaité et al., 2024). En nuestro caso, tenemos una cosa similar con el parámetro CONGA, que puede ser calculado para diferentes escalas de tiempo, n . Si representamos ambas variables en escala logarítmica, podemos observar una curva que, en el primer tramo presenta una pendiente bastante marcada, para luego estabilizarse a partir de cierta escala, lo que se puede observar en la Figura 9. Por tanto, podemos hacer una estimación de la pendiente en esta primera etapa por medio de una interpolación polinómica. De hecho, a partir de aquí, cada vez que se mencione CONGA, nos referiremos a esta estimación de la dimensión fractal.

Como segunda medida de complejidad se ha escogido la complejidad de Lempel-Ziv (*Lempel-Ziv complexity*, LZC). Para su cálculo, hay que binarizar la secuencia temporal de la señal bajo estudio, en este caso la curva de glucosa. Sobre esta secuencia, se determina el número de patrones distintos existentes, que será más elevado cuanto más compleja sea la señal (Borowska, 2021). Normalmente, el número de secuencias distintas se normaliza por el número de muestras que contiene la señal, de tal forma que LZC toma un valor entre 0 y 1.

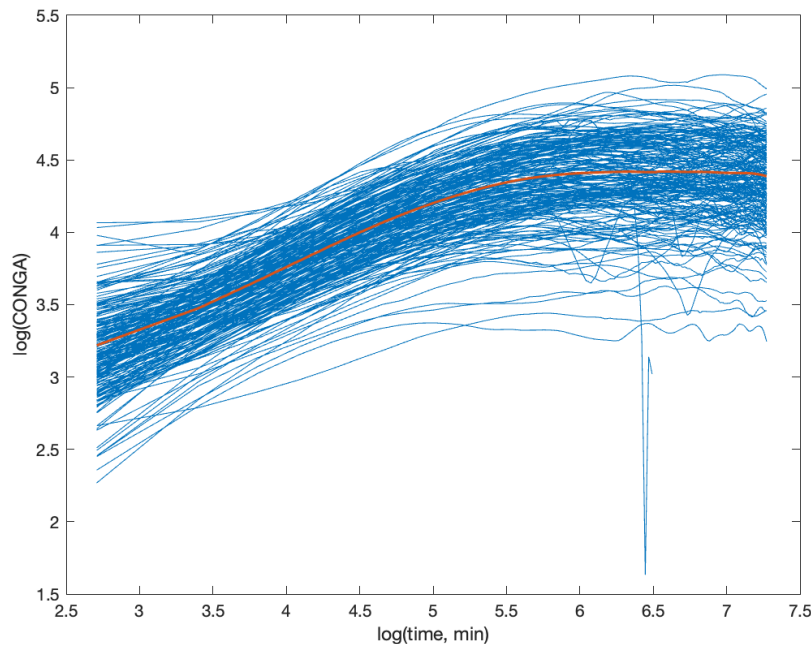


Figura 9. Valores de CONGA en función de la diferencia de tiempo entre muestras consecutivas para todos los pacientes de la base de datos (en azul). En naranja se observa el valor promedio de toda la base de datos.

Finalmente, como medida de variabilidad pura se ha incluido la medida de tendencia central (*central tendency measure*, CTM). En este caso, se cuenta el número total de puntos de un gráfico de diferencias de segundo orden que caen dentro de una circunferencia de radio r (Salankar et al., 2022). En estos gráficos, se representan en los ejes x e y las siguientes diferencias: $G(i+1) - G(i)$ y $G(i+2) - G(i+1)$, donde $G(i)$ es el nivel de glucosa en el instante de tiempo i . Para normalizar su valor, de tal forma que se mueva en un rango entre 0 y 1, se divide el número de puntos que quedan dentro de la circunferencia por el número total de puntos menos dos – esto se debe a que tenemos diferencias de segundo orden. Podemos ver la fórmula de cálculo a continuación:

$$CTM(r) = \frac{1}{n-2} \sum_{i=1}^{n-2} \delta(d(i)), \delta(d(i)) = \begin{cases} 1 & \text{si } d(i) = ([G(i+2) - G(i+1)]^2 + [G(i+1) - G(i)]^2)^{0.5} < r \\ 0 & \text{en otro caso} \end{cases}$$

Una cosa que cabe destacar es que, cuanto más pequeño sea el valor de CTM, mayor es la variabilidad de la señal, ya que hay más puntos que caen fuera de la circunferencia definida por r , lo que quiere decir que las diferencias entre muestras son mayores. La dificultad añadida de este parámetro es que presenta el hiperparámetro r , cuyo valor hay que definir. A priori, no hay ningún valor de r estándar, porque depende mucho de la aplicación concreta. Por ello, se ha hecho un análisis exploratorio para un gran número de valores de r hasta encontrar el radio óptimo. En este caso, hemos considerado que el radio óptimo es aquel que diferencie mejor entre sujetos con péptido C detectable y sujetos con péptido C indetectable, medido a través del p -valor obtenido con el test U de Mann-Whitney, aunque realmente partiremos de unas medidas de CTM que, por defecto, van a clasificar bien entre grupos. Esto se realiza porque no tenemos ninguna referencia de estudios previos. Normalmente este tipo de medidas no lineales requieren de una normalización de los valores de la señal, pero teniendo en cuenta que lo que queremos es detectar la variabilidad de las curvas de glucemia y que estas no tienen ningún tipo de ruido que altere su morfología – ya que vienen de datos analíticos y no son señales eléctricas –, en este caso no se va a aplicar ninguna normalización. El problema de no normalizar es que los valores en el gráfico de diferencias de segundo orden no quedan restringidos en torno al valor de 1, por lo que hay que realizar un análisis más extenso. De hecho, se inició probando radios en el rango de 0 a 200, para luego ir acotando en las zonas donde las diferencias entre los dos grupos (péptido C detectable vs. indetectable) eran menores. Este análisis se hizo por separado en el grupo de sujetos pediátricos y adultos y, finalmente, se obtuvieron como radios óptimos $r = 5.85$ para sujetos pediátricos

y $r = 10.45$ para adultos, como se observa en la Figura 10. No obstante, para simplificar los análisis se va a tomar el mismo radio para ambos grupos, $r = 6.5$.

4.3.3 Medidas de variación del péptido C

Para todos los análisis que se van a realizar, la variable *target* es el nivel de péptido C de cada paciente. Pero, al final, el objetivo de este trabajo es comprobar si existe alguna correlación entre las características de VG y la evolución del péptido C. Por este motivo, se han incluido varias variables relativas a la variación de péptido C en los pacientes que presentan más de una medida en sus analíticas de sangre:

- **Variación global de péptido C (VarC).** Esta es la diferencia entre el valor de la primera medida de péptido C (valor basal) y la última medida. Para que todos estos valores sean comparables, se divide la diferencia por el valor basal de cada paciente, de modo que vemos cuánto ha disminuido el nivel de péptido C en proporción al valor del que se partía.
- **Variación promedio mensual de péptido C (VarMonth).** Como tenemos las fechas en las que se ha medido el péptido C, podemos conocer la cantidad de meses de evolución desde la primera medida hasta la última. Aunque se trata de una aproximación un poco burda, si dividimos la variación global, VarC, entre el número de meses, obtenemos una variación mensual media del paciente.
- **Variaciones parciales de péptido C.** Como tenemos pacientes con varias medidas de péptido C, se puede calcular la caída de péptido C en ciertos intervalos temporales. Tras realizar un análisis exploratorio de los pacientes que presentan un evolutivo del péptido C, se han establecido 3 intervalos temporales: variación de péptido C a los 6-18 meses (Var1), variación de péptido C a los 18-30 meses

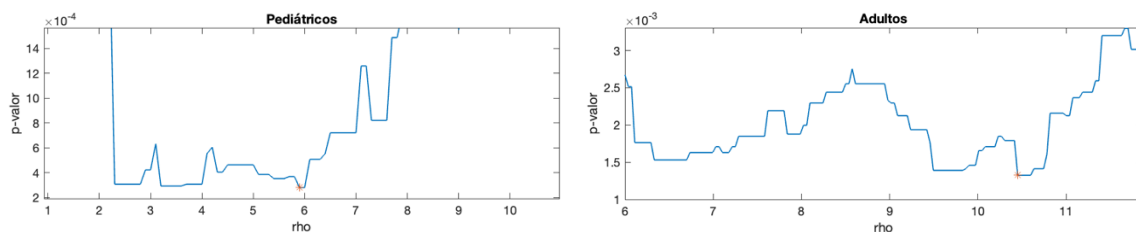


Figura 10. *P*-valor del test U de Mann-Whitney entre el grupo de péptido C detectable e indetectable en función del valor del hiperparámetro r (ρ). A la izquierda se observa el grupo de sujetos pediátricos y a la derecha el de adultos. El radio óptimo en cada caso está marcado con un asterisco.

(Var2) y variación de péptido C a los 30-36 meses (Var3). Nuevamente, todas estas medidas están normalizadas por el valor basal. Realmente, la Var3, en los pacientes que presentan valores a tan largo plazo, siempre coincide con la variación global, VarC, por lo que va a generar mucho ruido y puede alterar los resultados obtenidos. Por este motivo, finalmente esta última variable se ha decidido no incluirla en el estudio.

Es importante recalcar que todas las variaciones, a pesar de ser pérdidas, se expresan como diferencias positivas. Esto es debido a que se restan las diversas medidas de las analíticas con respecto al valor basal, que es el más elevado porque el nivel de péptido C disminuye con el tiempo.

4.4 Imputación de datos

Aunque la mayoría de los datos han sido extraídos directamente de las curvas de glucemia y, por tanto, no hay muchos datos perdidos, sí que hay algunos pacientes que no presentan HbA1c en sus analíticas en el momento en el que se midió el péptido C. Además, las variables Var1 y Var2 presentan valores perdidos, ya que no todos los pacientes con evolutivo presentan medidas de péptido C en ambos intervalos de tiempo considerados. Por este motivo, es conveniente realizar un curado de la base de datos, en el que se van a imputar las medidas faltantes para evitar que esto afecte a los análisis posteriores.

Para llevar a cabo la imputación, se ha empleado el algoritmo de imputación múltiple por cadenas de ecuaciones (*multiple imputation by chained equations*, MICE), que viene implementado en la librería “mice” del lenguaje de programación R (Buuren & Groothuis-Oudshoorn, 2011). En este algoritmo, se parte de la base de datos sin imputar y se generan diversas bases de datos con las variables imputadas con valores que sean factibles. Los valores imputados son extraídos de una distribución que es generada de manera específica para cada valor perdido. A continuación, se hace un análisis de las diferentes imputaciones, en el que se calcula un parámetro Q , que denota la cantidad de interés científico. Finalmente, todos los resultados se agrupan en uno único asumiendo que la distribución de las Q estimadas es aproximadamente normal. Es importante tener en cuenta que, para poder imputar una variable, es necesario que el porcentaje de valores perdidos sea inferior al 20% (Haji-Maghsoudi et al., 2013). Para comprobar que la imputación se ha realizado correctamente, se observa el histograma antes y después de la

imputación y, sobre ellos, se calculan el coeficiente de correlación de Pearson y la divergencia de Kullback-Leibler. El primero calcula el grado de asociación entre las dos distribuciones; mientras que la segunda mide el grado de diferencia entre ellas. Si la imputación es adecuada, la distribución de los histogramas debería ser similar, el coeficiente de correlación de Pearson debería ser cercano a 1 y la divergencia de Kullback-Leibler debería ser cercana a 0.

4.5 Análisis de correlaciones: redes de asociación

Una vez se han extraído todas las variables pertinentes y se han imputado los valores perdidos existentes, se va a realizar un análisis de correlaciones por medio de las redes de asociación. En nuestro caso, estas redes se van a generar con la ayuda del paquete “qgraph” de R (Epskamp et al., 2012). En él, los enlaces entre nodos presentan un color verde en caso de que las correlaciones sean positivas y rojo si las correlaciones son negativas. Como ya se ha comentado, el grosor de los enlaces nos indica la fuerza de la correlación, de modo que cuanto más gruesos, más cerca del 1 o -1 está la correlación entre dos variables. El coeficiente de correlación que se calcula se puede incluir como argumento de entrada a la red; en este estudio, se ha escogido el coeficiente de correlación de Spearman porque es capaz de identificar relaciones tanto lineales como no lineales. La forma en la que se disponen los nodos a la salida también es modificable, aunque por defecto salen todos formando un círculo y las correlaciones en el interior. Si se escoge la disposición “spring”, las distancias entre los nodos son también proporcionales a la fuerza de las correlaciones entre ellos. Otro de los parámetros que es útil es el umbral, que nos permite visualizar solo las correlaciones que están por encima de un cierto valor de corte; mientras que el resto no se muestran.

En todas las redes generadas se van a incluir la edad y el sexo como covariables para eliminar su posible influencia de la red de correlaciones presente. Inicialmente, se va a generar una red con todos los parámetros de VG, tanto los que son característicos del ámbito clínico como las medidas no lineales. Así, podremos hacer un análisis exploratorio de las relaciones entre ellos, para comprobar cuáles aportan información redundante y cuáles son los diferentes clústeres en los que se agrupan los parámetros de VG. Estos grupos se intentarán asociar con los diferentes aspectos de la VG en base a los conocimientos previos sobre cada variable que se han encontrado en la literatura. Tras este análisis inicial, se van a generar redes que, además de todos estos parámetros,

incluyan variables de clasificación de los pacientes en función de la reserva pancreática y medidas del nivel y la variación de péptido C. Con esta segunda aproximación, se intentará establecer qué variables presentan más correlación con nuestra variable *target*. Al final, estos análisis se hacen desde un punto de vista cualitativo, ya que nunca se han evaluado todos los parámetros de VG de forma global, simplemente se han establecido correlaciones de cada uno de ellos por separado con diversos parámetros clínicos (Bowler et al., 2024; El-Laboudi et al., 2016; Leksic et al., 2022; Tokutsu et al., 2022; Wang et al., 2022).

4.6 Selección de características

Tras realizar el estudio exploratorio con las redes de asociación, es necesario llevar a cabo una selección de características. Este paso es esencial para reducir la dimensionalidad inicial del problema y obtener un subespacio de características más manejable, útil cuando queremos llevar a cabo una clasificación de los sujetos (Shafiee et al., 2021). Como es la primera vez que se aborda este problema, se plantea un caso sencillo, en el que se busca clasificar los pacientes en función de si tienen reserva pancreática (péptido C > 0.1 ng/mL) o si no la tienen (péptido C < 0.1 ng/mL). Para ello, primero se ha planteado seleccionar las variables con un menor p -valor de cada uno de los clústeres que se han obtenido con las redes de asociación. Posteriormente, se han empleado algoritmos un poco más complejos como *Sequential Forward Selection* (SFS) y *Fast Correlation-Based Filter* (FCBF). Finalmente, se han intentado seleccionar aquellas variables que están menos correlacionadas con el resto de las variables de la red de asociación, por medio del establecimiento de umbrales más elevados. El valor de estos umbrales será estricto, de 0.7 y 0.8.

SFS está encuadrado dentro de los métodos de selección secuencial de características, concretamente se considera un algoritmo basado en una metodología “greedy”. Esto quiere decir que, en cada iteración del método, se elige la opción más adecuada en el momento, sin tener en cuenta las consecuencias al final del algoritmo. En concreto, SFS parte de una única característica, que suele ser aquella que presenta un rendimiento superior a las demás. Esto se puede evaluar de varias maneras, aunque la más habitual es calcular la precisión del modelo que se quiere obtener con cada una de las variables. Después, se van añadiendo variables, una en cada iteración del algoritmo, de tal forma que siempre se selecciona aquella que proporciona una precisión mayor a la salida. Este

ciclo se repite hasta que el modelo presenta un número de características que es establecido primeramente por el usuario (Shafiee et al., 2021). Como en cada iteración se evalúa la utilidad del modelo, es decir, si tiene una mayor precisión comparado con otros, también se le considera un método *wrapper*. En nuestro caso, se evaluará la precisión del modelo y se busca obtener un modelo final que presente $n = 5$ características.

FCBF es un método de selección de características que, al contrario que SFS, es independiente del modelo que se quiera generar posteriormente, por lo que se le considera un método de filtrado. En él, se seleccionan las características que son más relevantes y, al mismo tiempo, que no sean redundantes (Ying et al., 2021). Primero se evalúa la relevancia de todas las variables por medio de la *symmetric uncertainty* (SU), que establece la cantidad de información que comparten dos variables, de modo que vale 1 si las dos variables están totalmente correlacionadas y 0 si son independientes. Aquellas variables que son relevantes tendrán un valor de SU con la variable *target* cercano a 1, de tal forma que podemos ordenar las variables en orden decreciente de SU. Para eliminar las variables no relevantes, se establece un umbral de SU que se debe superar. Esta es la primera etapa del algoritmo; en la segunda etapa, se van a eliminar las variables redundantes. Para ello, se empieza analizando la característica más relevante y se calcula su SU con todas las otras características. Si alguna variable tiene una SU con esta primera característica más elevada que su SU con la variable *target*, esto quiere decir que comparte más información con las otras variables que con la que se quiere predecir. En consecuencia, esta variable se considera redundante y se elimina. Esto se repite con la siguiente variable más relevante y así sucesivamente hasta llegar a la última variable de la lista. En nuestro estudio, además, se generan 1000 muestras de *Bootstrap* a partir de la base de datos inicial, de modo que se hace la selección de variables con este algoritmo durante 1000 veces. Con este método, incrementamos la robustez del proceso de selección, ya que tenemos una estadística de cuántas veces se ha seleccionado para cada variable para muestras que, en principio, son “diferentes”, aunque realmente todas son extracciones hechas a partir de la base de datos global. Así, para generar el modelo podemos tomar solo aquellas variables que superan un cierto umbral de veces de selección dentro de las muestras de *Bootstrap*. En nuestro estudio, los umbrales seleccionados fueron de 50, 100 y 125, lo que nos permite seleccionar 5, 3 y 2 variables, respectivamente.

4.7 Clasificadores péptido C detectable vs. indetectable

Partiendo de la selección de variables realizada con cada uno de los métodos y algoritmos descritos anteriormente, se van a aplicar diversos modelos de *Machine Learning* que nos permiten realizar una clasificación entre dos categorías. Estas variables seleccionadas son las variables predictoras del modelo; mientras que la variable *target* es la denominada ‘isReserve’, que nos indica si un individuo tiene reserva pancreática (=1) o no la tiene (=0). Los diversos clasificadores empleados son el Análisis Discriminante Lineal (*Linear Discriminant Analysis*, LDA), la Regresión Logística (*Logistic Regression*, LR) y el Análisis Discriminante Cuadrático (*Quadratic Discriminant Analysis*, QDA).

LDA es uno de los métodos de clasificación más sencillos que están englobados dentro de la categoría de *Machine Learning*. Este método busca obtener una combinación lineal de las variables predictoras para poder determinar la variable *target*. Para ello, transforma el espacio N -dimensional de características inicial en un subespacio de menor dimensión, con lo que busca maximizar la distancia interclase y minimizar las distancias intraclase (Jessica et al., 2021). Por este motivo, es un algoritmo muy útil en problemas de clasificación, como es nuestro caso. LR es muy similar, porque también es un mecanismo basado en una combinación lineal, aunque en este caso se aplica una transformación logarítmica para obtener un valor entre 0 y 1 (Guerrero et al., 2021). Por último, QDA también realiza un análisis discriminante, pero en este caso la combinación de variables ya no es lineal. Al contrario que LDA, no asume que todas las covarianzas son iguales, por lo que se calcula una matriz de covarianzas (Chand & Vishwakarma, 2022). Ante ello, nos aporta una mayor cantidad de parámetros y, con ellos, una mayor flexibilidad. Por supuesto, todo esto a costa de una menor robustez del modelo.

En todos estos modelos, el valor que se obtiene a la salida se encuentra entre 0 y 1, ya que hablamos de una clasificación binaria. Por tanto, no obtenemos una categorización de los datos directamente, sino que lo que nos devuelven LDA, LR y QDA es una probabilidad de pertenencia a cada clase. Por tanto, cuanto más cerca estemos del 0, más probabilidades hay de que el individuo no tenga reserva pancreática y, a medida que nos acercamos al 1, esta probabilidad disminuye y aumenta la de pertenencia a la clase con reserva pancreática. En este contexto, es necesario establecer un umbral y, *a priori*, no está claro cuál es el óptimo. Por este motivo, se ha creado un vector de umbrales entre 0

y 1, con una cadencia entre ellos de 0.01, a partir de los que se extraen las medidas de especificidad y sensibilidad del modelo. Para generar estos valores de rendimiento del modelo, es importante que las instancias o sujetos que se usen sean distintos a los que se han empleado en la generación del propio modelo. Esto es así porque se trata de modelos de aprendizaje supervisado, en los que las etiquetas de cada sujeto son conocidas, por lo que validar el rendimiento del modelo con las mismas instancias puede suponer la obtención de unos valores excesivamente optimistas. Por este motivo, se va a emplear la metodología de validación cruzada dejando uno fuera (*leave-one-out cross-validation*), en la que, en cada iteración del entrenamiento, una de las instancias de la base de datos queda fuera en la etapa de generación del modelo y se emplea posteriormente para determinar el rendimiento del modelo creado (T.-T. Wong, 2015). De esta manera, aseguramos que todas las mediciones del rendimiento en cada iteración son independientes entre sí.

4.8 Análisis estadístico

Durante todo el trabajo, la estadística es una herramienta imprescindible que nos permite resumir y describir los datos de los que partimos, así como establecer la validez de las herramientas empleadas para analizar estos datos. Antes de iniciar el análisis de correlaciones, es necesario llevar a cabo un análisis poblacional de los datos, como una aproximación inicial al proceso de clasificación. Para ello, nos vamos a valer de dos herramientas: el test U de Mann-Whitney y los *box plots/violin plots*. El test U de Mann-Whitney o test de Wilcoxon determina si hay una diferencia estadísticamente significativa entre las distribuciones de dos variables para diferentes grupos poblacionales (El Hallani et al., 2025). En nuestro caso concreto, se busca observar si hay diferencia significativa en la distribución de los parámetros de VG entre los pacientes con péptido C detectable e indetectable. Esta es una prueba no paramétrica, por lo que no asume la normalidad de los datos, lo que es beneficioso, ya que la distribución de variables en el ámbito de la diabetes no suele alinearse con esto – visto con la distribución de los niveles de glucosa. Para determinar si hay una diferencia significativa, el test nos devuelve un *p*-valor, que será significativo cuando sea inferior a un cierto umbral. Esta prueba se va a aplicar por separado a los grupos de adultos y de sujetos pediátricos, ya que pueden existir diferencias. Aparte de este test estadístico, podemos hacer un análisis de las distribuciones, tanto para adultos como para niños, para sujetos con péptido C detectable e indetectable. Lo más sencillo es establecer la media $\pm$ SD de cada uno de los parámetros

de VG, de tal forma que den más consistencia a los resultados obtenidos con el test U de Mann-Whitney. Un abordaje más visual se puede obtener gracias a la herramienta de MATLAB® “*grpandplot*”, que nos permite observar de manera conjunta la distribución de los datos con un *scatter plot*, acompañada de un *box plot* y un *violin plot* (M. H. Wong, 2022). Gracias al *scatter plot*, podemos observar los valores de todos los sujetos para una cierta variable de manera conjunta. A mayores, el *box plot* o diagrama de cajas nos marca la mediana de la distribución de datos, así como el primer y tercer cuartil. Los bigotes de este gráfico se extienden 1.5 veces el IQR por debajo del primer cuartil y por encima del tercero, de tal forma que los valores que quedan fuera se consideran *outliers* (Frigge et al., 1989). Algo más de información aporta el *violin plot*, que consiste en una traza de densidad de la distribución de los datos dispuesta de manera similar al *box plot* (Hintze & Nelson, 1998). De esta forma, se obtiene algo más de información de dónde se concentran los datos. Todos estos gráficos combinados nos permiten caracterizar correctamente la distribución de cada uno de los parámetros bajo estudio.

En la etapa de las redes de asociación, es necesario establecer qué correlaciones son realmente significativas y cuáles no. Para que una correlación se considere significativa, debe ser similar independientemente de la muestra concreta bajo estudio. El umbral de corte no se encuentra siempre en el mismo valor de asociación, por lo que es necesario realizar un análisis de estabilidad de cada red en concreto. Esto se hace por medio de la generación de 1000 muestras de *Bootstrap* partiendo de la red original, de tal forma que se calcula la matriz de correlaciones para cada una de ellas. Después, se extraen las matrices de correlación de la mediana y los percentiles 2.5% y 97.5%, de tal manera que esto se observa en un gráfico junto con los valores de la red de asociación estándar. En el eje de abscisas se observan los valores de correlación y en el eje de ordenadas el número de veces que aparece cada uno de esos valores para las matrices obtenidas mediante *bootstrapping*. Gracias a las matrices de correlación del 2.5% y el 97.5%, podemos observar el intervalo de confianza del 95%, de modo que la red será estable en el punto en el que este intervalo de confianza no corte el valor cero de correlación. El valor de correlación de la matriz mediana (y de la matriz original) que se corresponda con este intervalo de confianza será el umbral que es necesario establecer en la red de asociación para que solo se muestren las correlaciones estables.

Por último, para determinar el rendimiento de los modelos de clasificación empleados, se va a hacer uso de las curvas ROC (*receiver operating characteristic*). Aquí, se emplean

los valores de especificidad y sensibilidad calculados anteriormente para cada uno de los umbrales, que se representan en los ejes de coordenadas: $1 - \text{especificidad}$ en el eje de abscisas y la sensibilidad en el eje de ordenadas. De estos gráficos se puede extraer el área bajo la curva ROC (AUC), que se utiliza como un estimador del rendimiento del modelo, de tal forma que el modelo ideal es aquel en el que la AUC es 1; mientras que un modelo que clasifica de manera completamente azarosa presenta una AUC de 0.5 (Çorbacioğlu & Aksel, 2023). Gracias a esto, podremos comprobar si los clasificadores que se han escogido para este trabajo realizan una clasificación más precisa cuando se incluyen varias variables que cuando solo se incluye un parámetro de VG.

Capítulo 5. Resultados

4.1 Introducción.....	74
4.2 Imputación de datos.....	74
4.3 Análisis poblacional.....	75
4.4 Análisis de correlaciones.....	86
4.5 Estudio de clasificación.....	90

5.1 Introducción

En este capítulo se presentan los principales resultados obtenidos a raíz de los análisis realizados a lo largo de este TFG. Podemos distinguir tres etapas, donde la primera es un análisis poblacional en el que se explora la base de datos, de tal forma que se pueda comprobar la potencialidad de los parámetros de VG extraídos como clasificadores de la reserva pancreática de los pacientes con DM1. En la segunda se lleva a cabo un análisis de correlaciones por medio de las redes de asociación, en el que se pretenden establecer clústeres de manera visual y comprobar la relación de los parámetros de VG con las variables relacionadas con el nivel o la variación de péptido C. Por último, se realiza un análisis de clasificación por medio de diversos algoritmos de *Machine Learning* – LDA, LR y QDA –, de tal forma que se determina el rendimiento de cada uno de los modelos generados y se compara con el rendimiento de los parámetros de VG de forma individual.

5.2 Imputación de datos

Antes de empezar a realizar los diferentes análisis, es necesario realizar la imputación de los datos perdidos en la base de datos. Se observa que, en este caso, las variables con valores perdidos son HbA1c (2.43%), MODD (0.47%), Var1 (17.65%) y Var2 (23.53%). Por tanto, en principio se pudieron imputar todas las variables excepto Var2, que supera el porcentaje de datos perdidos permitido. Para ello, se emplea la librería “mice” de R sobre los valores perdidos de la base de datos.

Una vez realizada la imputación, se obtuvieron las distribuciones de los histogramas de las tres variables sin imputar y ya imputadas, que se observan en la Figura 11. Como se puede ver, la distribución de los tres histogramas apenas se ve alterada. A partir de las distribuciones de los histogramas sin imputar e imputados, se calculó el coeficiente de correlación de Pearson y la divergencia de Kullback-Leibler entre ellas. Para la HbA1c, se obtuvo un coeficiente de correlación de Pearson de 1 y una divergencia de Kullback-Leibler de 7.8133×10^{-4} . Para MODD, se obtuvo un coeficiente de correlación de Pearson de 1 y una divergencia de Kullback-Leibler de 4.9731×10^{-4} . Para Var1, se obtuvo un coeficiente de correlación de Pearson de 0.9898 y una divergencia de Kullback-Leibler de 0.0172. Ante estos resultados, la imputación de variables fue aceptada.

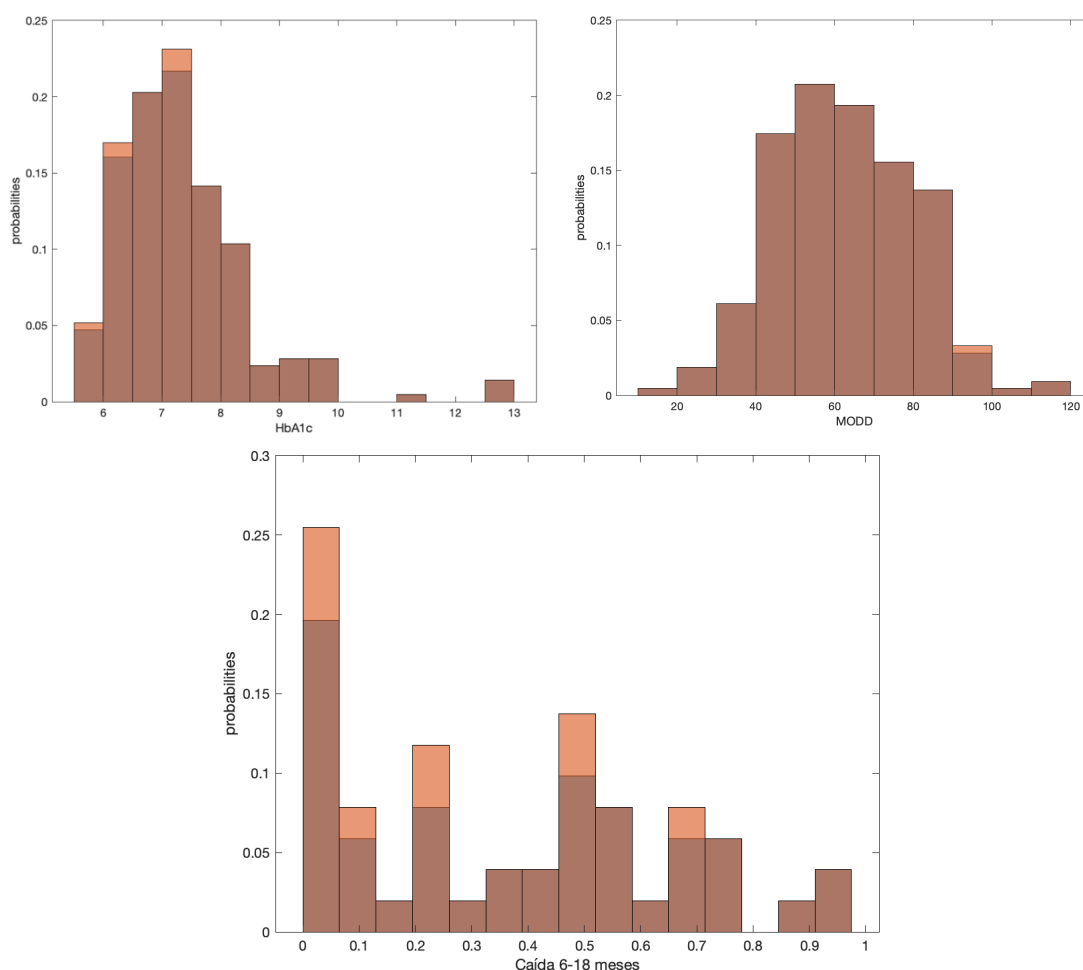


Figura 11. Distribución de los histogramas de las variables con datos perdidos antes de la imputación (en gris) y una vez imputadas (en naranja). Arriba a la derecha se observan los histogramas de HbA1c, arriba a la izquierda, los de MODD y abajo, los de Var1 (Caída 6-18 meses).

5.3 Análisis poblacional

A partir de los diversos pacientes de la base de datos, se ha llevado a cabo un análisis poblacional, en el que se ha realizado una caracterización de las diversas variables bajo estudio. Cabe mencionar que estos análisis se han llevado a cabo de manera independiente en el grupo de sujetos adultos y pediátricos. En la Tabla 5 se puede observar para el grupo de adultos la media $\pm$ SD de los pacientes con reserva pancreática en la primera columna, la media $\pm$ SD de los pacientes sin reserva pancreática en la segunda columna, y el *p*-valor resultado del test U de Mann-Whitney en la tercera columna. Lo mismo se observa en la Tabla 6 del grupo de sujetos pediátricos. En general, podemos ver que los pacientes sin reserva pancreática presentan unos valores promedio de los parámetros de VG más elevados que los pacientes con reserva, excepto HbA1c, MinGly, TIR, TITR, TOTR, LZC y CTM. Esto ocurre tanto en adultos como en sujetos pediátricos. La gran mayoría de los

variable	Con reserva (media $\pm$ SD)	Sin reserva (media $\pm$ SD)	<i>p</i>-valor
HbA1c	7.50 $\pm$ 0.95	7.37 $\pm$ 0.90	0.5656
TBR2	0.45 $\pm$ 0.65	0.64 $\pm$ 0.95	0.1661
TBR1	2.46 $\pm$ 2.32	3.43 $\pm$ 2.55	0.0255
TIR	66.29 $\pm$ 17.36	60.43 $\pm$ 16.77	0.0498
TAR1	21.84 $\pm$ 9.44	23.71 $\pm$ 8.41	0.3976
TAR2	8.56 $\pm$ 10.33	11.57 $\pm$ 11.37	0.0342
TITR	42.42 $\pm$ 16.19	37.73 $\pm$ 14.61	0.1198
TOTR	23.87 $\pm$ 5.8	22.70 $\pm$ 5.28	0.3273
CHypo	2.42 $\pm$ 2.33	3.39 $\pm$ 2.75	0.0336
CHyper	19.48 $\pm$ 13.85	23.42 $\pm$ 14.06	0.0697
GRI	38.43 $\pm$ 21.45	47.65 $\pm$ 21.86	0.0160
SD	55.12 $\pm$ 15.16	62.43 $\pm$ 13.91	0.0053
CV	34.39 $\pm$ 5.12	37.81 $\pm$ 5.57	0.0015
MAGE	169.37 $\pm$ 48.99	198.32 $\pm$ 56.47	0.0071
MODD	57.28 $\pm$ 17.83	66.97 $\pm$ 15.55	0.0018
IQR	73.94 $\pm$ 22.74	83.86 $\pm$ 22.13	0.0143
LBGI	2.93 $\pm$ 1.44	3.75 $\pm$ 1.61	0.0158
HBGI	8.94 $\pm$ 5.04	10.67 $\pm$ 5.19	0.0265
ADRR	34.74 $\pm$ 13.36	41.17 $\pm$ 12.49	0.0023
MinGly	52.46 $\pm$ 7.78	51.75 $\pm$ 7.35	0.3099
MaxGly	471.49 $\pm$ 114.98	524.66 $\pm$ 131.89	0.0323
J	47.46 $\pm$ 19.94	53.46 $\pm$ 20.24	0.0400
LAGE	280.53 $\pm$ 84.75	339.44 $\pm$ 120.79	0.0050
IGC	11.86 $\pm$ 5.66	14.42 $\pm$ 5.73	0.0079
GVP	38.92 $\pm$ 15.62	43.86 $\pm$ 12.03	0.0031
MAG	42.34 $\pm$ 13.98	48.86 $\pm$ 17.11	0.0024
CONGA	0.39 $\pm$ 0.08	0.39 $\pm$ 0.08	0.5353
LZC	0.38 $\pm$ 0.06	0.38 $\pm$ 0.08	0.6850
CTM	0.31 $\pm$ 0.09	0.27 $\pm$ 0.07	0.0015

Tabla 5. Análisis poblacional de los parámetros de VG para el grupo de sujetos adultos. En la primera columna se observan las variables bajo estudio, en la segunda las medias $\pm$ SD de cada parámetro para los pacientes con péptido C detectable (> 0.1 ng/mL), en la tercera las medias $\pm$ SD de cada parámetro para los pacientes con péptido C indetectable (< 0.1 ng/mL) y en la última el *p*-valor obtenido del test U de Mann-Whitney.

variable	Con reserva (media $\pm$ SD)	Sin reserva (media $\pm$ SD)	<i>p</i> -valor
HbA1c	7.32 $\pm$ 2.14	6.80 $\pm$ 0.61	0.4411
TBR2	0.58 $\pm$ 0.99	1.51 $\pm$ 2.07	0.0029
TBR1	4.98 $\pm$ 4.54	6.15 $\pm$ 4.04	0.0826
TIR	76.65 $\pm$ 14.38	65.64 $\pm$ 10.51	0.0004
TAR1	12.71 $\pm$ 9.97	18.63 $\pm$ 7.67	0.0028
TAR2	4.20 $\pm$ 6.09	7.42 $\pm$ 5.64	0.0009
TITR	57.30 $\pm$ 18.73	45.87 $\pm$ 11.11	0.0025
TOTR	19.35 $\pm$ 8.53	19.77 $\pm$ 4.25	0.3773
CHypo	4.56 $\pm$ 4.40	6.44 $\pm$ 4.94	0.0361
CHyper	10.55 $\pm$ 10.17	16.74 $\pm$ 8.42	0.0011
GRI	30.56 $\pm$ 18.55	46.09 $\pm$ 15.86	0.0005
SD	48.48 $\pm$ 17.52	62.88 $\pm$ 14.07	0.0001
CV	35.31 $\pm$ 7.63	42.71 $\pm$ 7.39	0.0003
MAGE	151.22 $\pm$ 58.65	187.21 $\pm$ 70.25	0.0417
MODD	49.86 $\pm$ 17.93	66.79 $\pm$ 14.17	5.86 $\times 10^{-5}$
IQR	60.94 $\pm$ 26.04	82.32 $\pm$ 20.32	0.0001
LBGI	3.33 $\pm$ 1.36	4.64 $\pm$ 1.73	0.0010
HBGI	6.23 $\pm$ 4.05	9.23 $\pm$ 3.31	0.0003
ADRR	30.48 $\pm$ 14.84	37.55 $\pm$ 13.74	0.0481
MinGly	51.79 $\pm$ 6.85	45.75 $\pm$ 5.65	0.0004
MaxGly	459.03 $\pm$ 230.50	518.69 $\pm$ 170.99	0.0442
J	35.28 $\pm$ 17.00	44.96 $\pm$ 12.83	0.0013
LAGE	284.30 $\pm$ 183.16	322.85 $\pm$ 123.75	0.0635
IGC	9.56 $\pm$ 4.60	13.87 $\pm$ 3.94	3.32 $\times 10^{-5}$
GVP	42.93 $\pm$ 22.68	48.14 $\pm$ 25.05	0.2094
MAG	48.75 $\pm$ 25.70	52.24 $\pm$ 39.92	0.9091
CONGA	0.34 $\pm$ 0.09	0.38 $\pm$ 0.07	0.0481
LZC	0.45 $\pm$ 0.08	0.41 $\pm$ 0.10	0.0140
CTM	0.34 $\pm$ 0.09	0.27 $\pm$ 0.07	0.0007

Tabla 6. Análisis poblacional de los parámetros de VG para el grupo de sujetos pediátricos. En la primera columna se observan las variables bajo estudio, en la segunda las medias $\pm$ SD de cada parámetro para los pacientes con péptido C detectable (> 0.1 ng/mL), en la tercera las medias $\pm$ SD de cada parámetro para los pacientes con péptido C indetectable (< 0.1 ng/mL) y en la última el *p*-valor obtenido del test U de Mann-Whitney.

p -valores obtenidos a raíz del test U de Mann-Whitney son estadísticamente significativos, aunque se observa que los resultados son ligeramente mejores en el grupo de sujetos pediátricos. En el caso de los adultos, los parámetros SD, MAGE, LAGE e IGC son significativos a nivel 0.01, y los parámetros CV, MODD, ADRR, GVP, MAG y CTM son significativos a nivel 0.005. Por su parte, en el grupo de sujetos pediátricos, los parámetros TBR2, TAR1, TITR, CHyper, LBGI y J son significativos a nivel 0.005 y los parámetros TIR, TAR2, GRI, SD, CV, MODD, IQR, HBGI, MinGly, IGC y CTM son significativos a nivel 0.001.

Podemos comparar las medias de ambos grupos con los umbrales que se han establecido para algunos de los parámetros de VG. El más importante es el CV, cuyo corte se establece en 36% para determinar un buen control glucémico. En ambos grupos, adultos y pediátricos, las medias de grupo de los pacientes con péptido C detectable se encuentran por debajo del corte y las de los que tienen péptido C indetectable están por encima. En el caso del TIR, en promedio, ambos grupos de adultos se encuentran por debajo del 70%; mientras que los sujetos pediátricos con reserva pancreática sí que se encuentran por encima del 70% en promedio. No obstante, este parámetro tiene mucha variabilidad, por lo que este corte no es perfecto. Lo mismo ocurre con el TITR, donde el corte se encuentra en el 50%.

Para completar el análisis, se presentan las distribuciones de los parámetros de VG, donde se comparan los pacientes con péptido C detectable (> 0.1 ng/mL) e indetectable (< 0.1 ng/mL), tanto para sujetos pediátricos como adultos. Para ello, se ha usado la función “grplot” de MATLAB®, que nos permite superponer en el mismo gráfico un *scatter plot*, un *violin plot* y un *box plot*. Además, también nos permite comparar las distribuciones entre varios grupos (con reserva vs. sin reserva o adultos vs. pediátricos). Con ellos, se corroboran los resultados observados con el test U de Mann-Whitney, aunque, en la mayoría de los casos, existe un solapamiento de las distribuciones entre los grupos de péptido C detectable e indetectable. Todo esto se muestra en las Figuras 12-40.

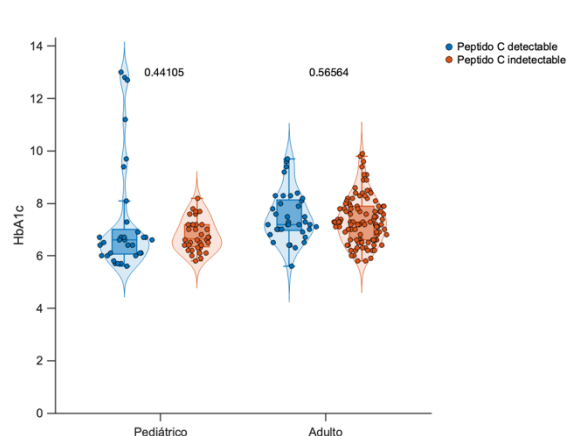


Figura 12. Distribución del parámetro HbA1c, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

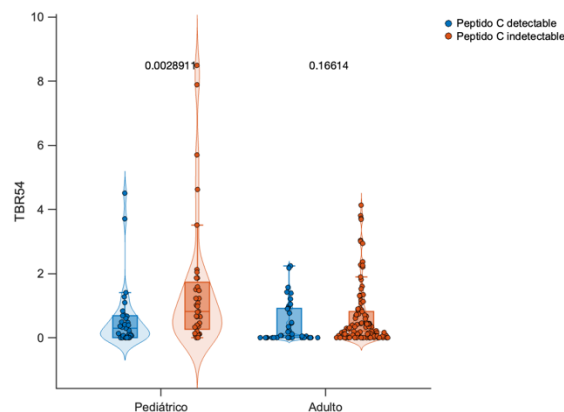


Figura 13. Distribución del parámetro TBR2 (TBR54), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

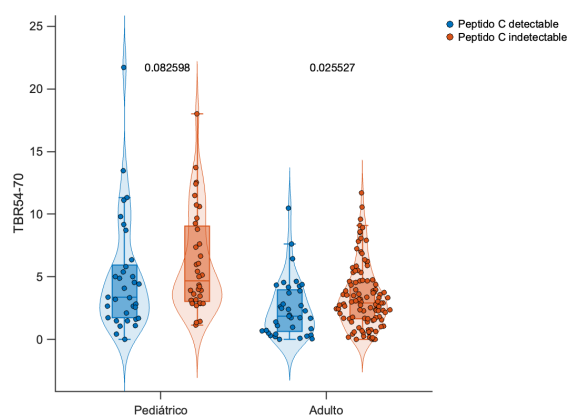


Figura 14. Distribución del parámetro TBR1 (TBR54-70), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

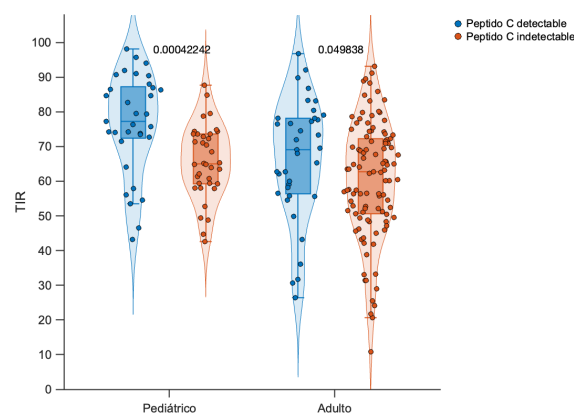


Figura 15. Distribución del parámetro TIR, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

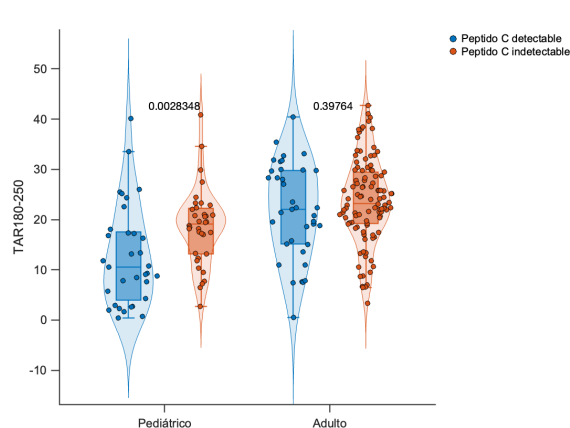


Figura 16. Distribución del parámetro TAR1 (TAR180-250), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

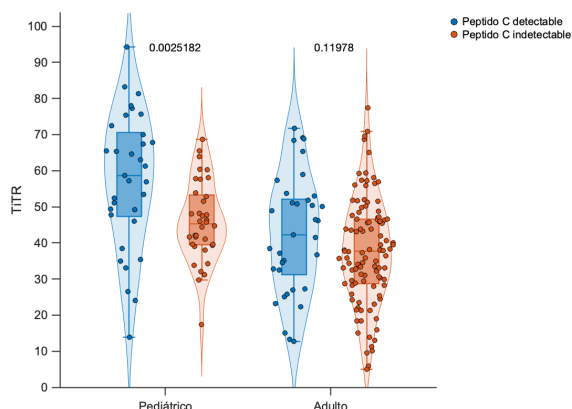


Figura 18. Distribución del parámetro TITR, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

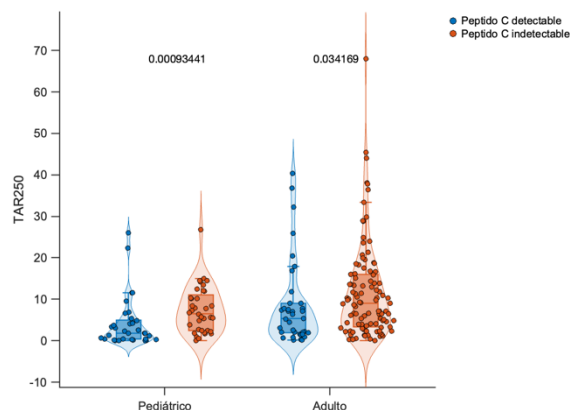


Figura 17. Distribución del parámetro TAR2 (TAR250), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

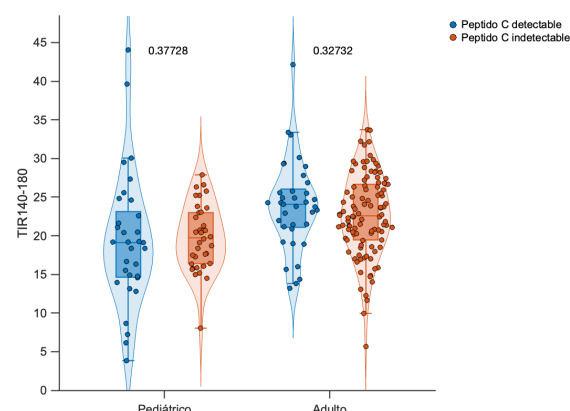


Figura 19.. Distribución del parámetro TOTR (TIR140-180), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

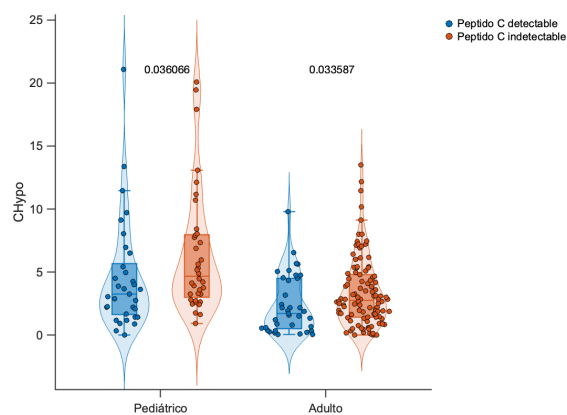


Figura 20. Distribución del parámetro CHypo, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

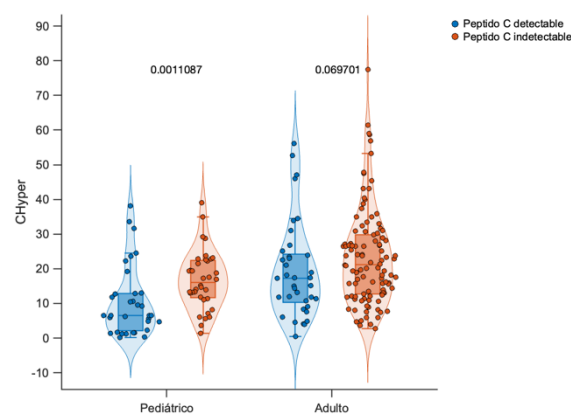


Figura 21. Distribución del parámetro CHyper, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

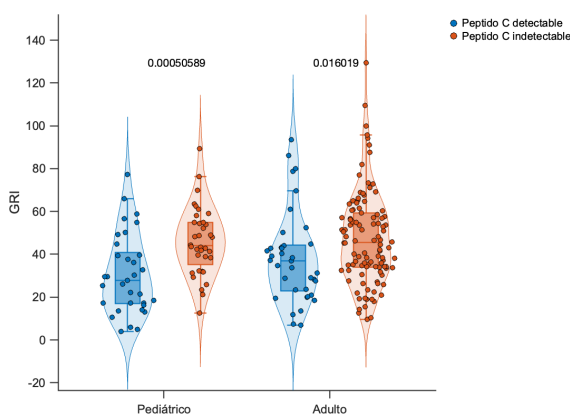


Figura 22. Distribución del parámetro GRI, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

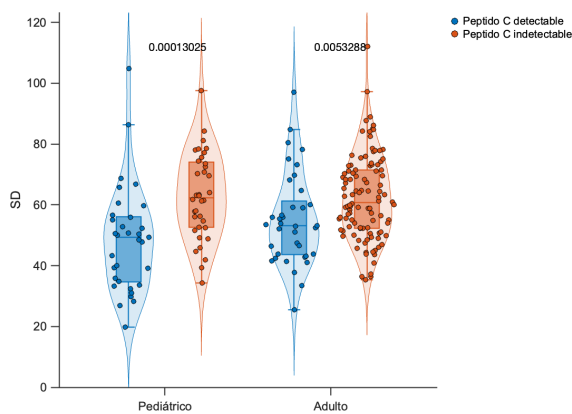


Figura 23. Distribución del parámetro SD, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

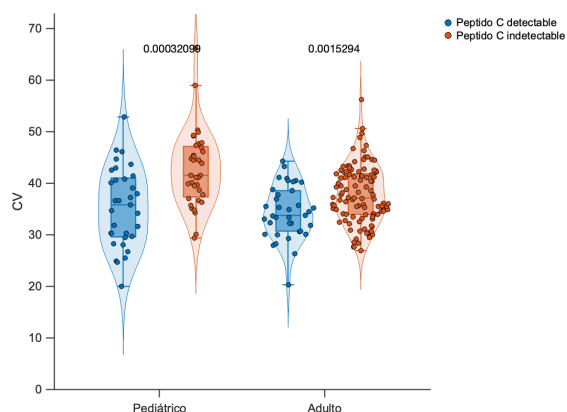


Figura 24. Distribución del parámetro CV, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

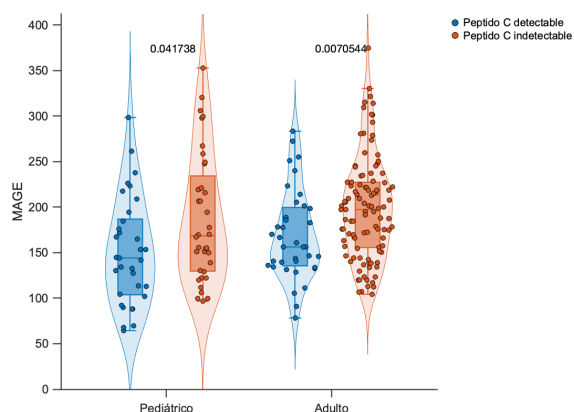


Figura 25. Distribución del parámetro MAGE, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

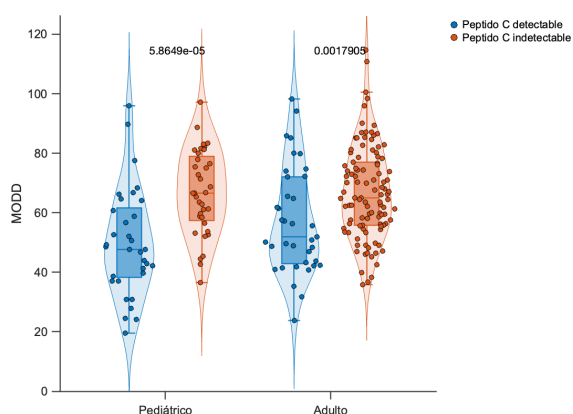


Figura 26. Distribución del parámetro MODD, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

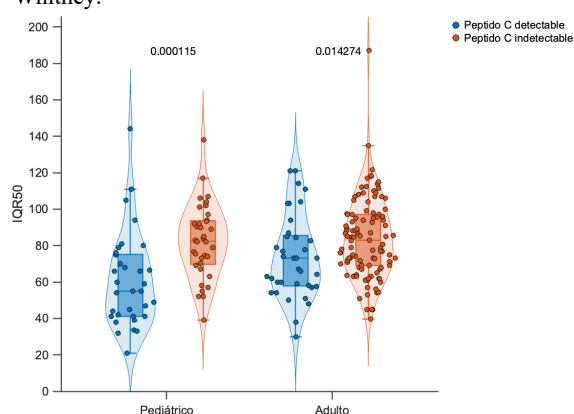


Figura 27. Distribución del parámetro IQR (IQR50), a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

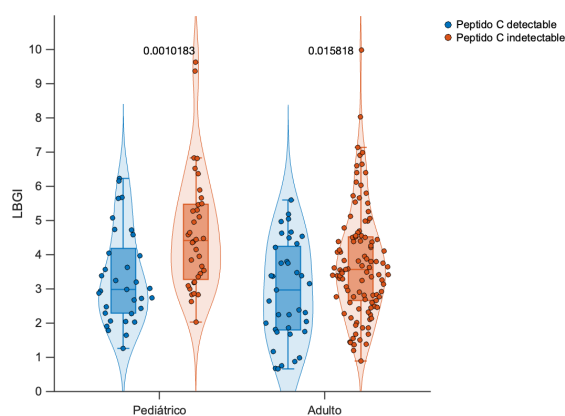


Figura 28. Distribución del parámetro LBGI, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

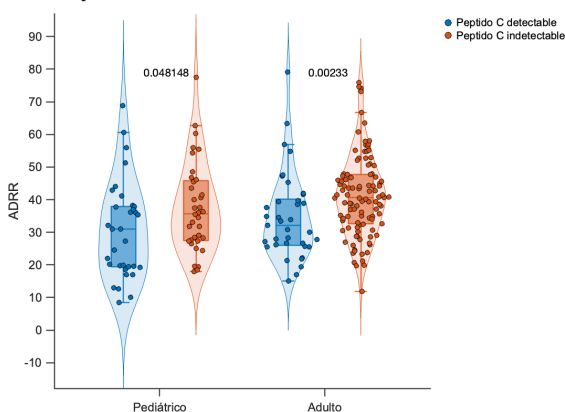


Figura 30. Distribución del parámetro ADRR, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

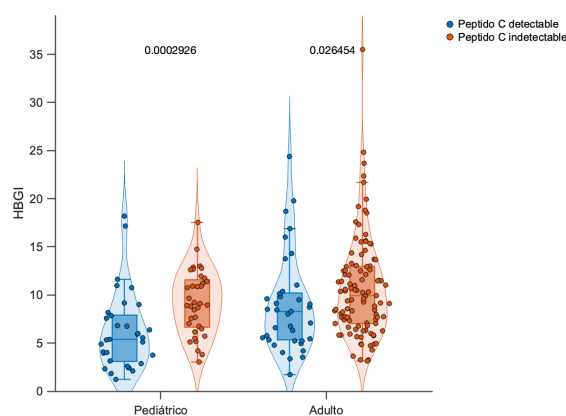


Figura 29. Distribución del parámetro HBGI, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

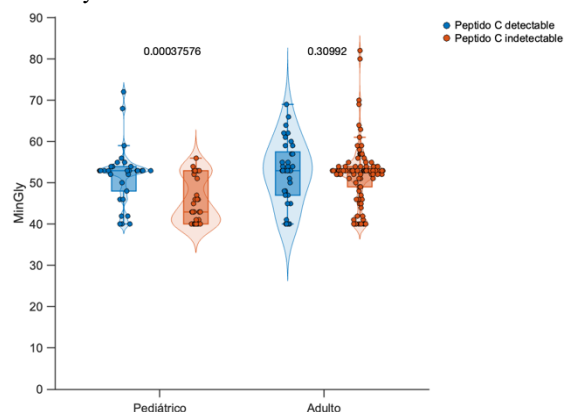


Figura 31. Distribución del parámetro MinGly, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

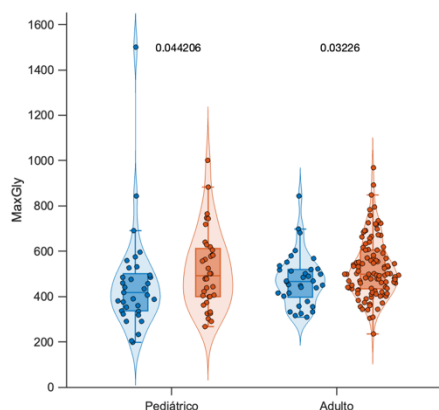


Figura 32. Distribución del parámetro MaxGly, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

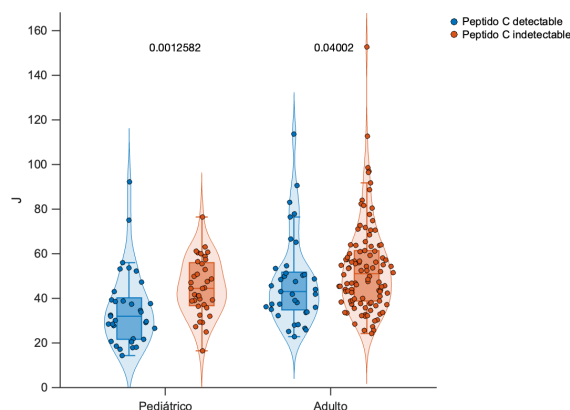


Figura 33. Distribución del parámetro J, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

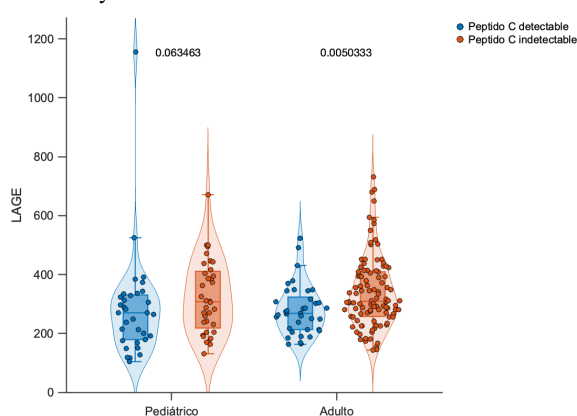


Figura 34. Distribución del parámetro LAGE, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

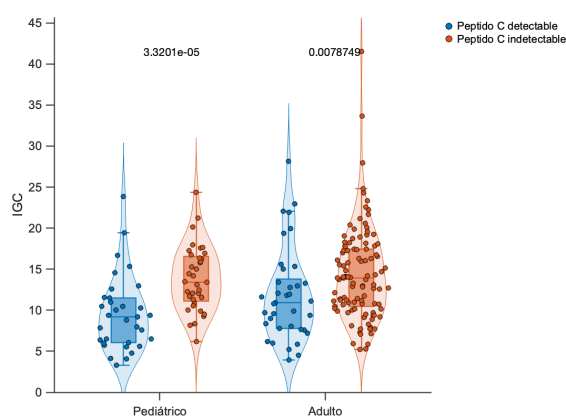


Figura 35. Distribución del parámetro IGC, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

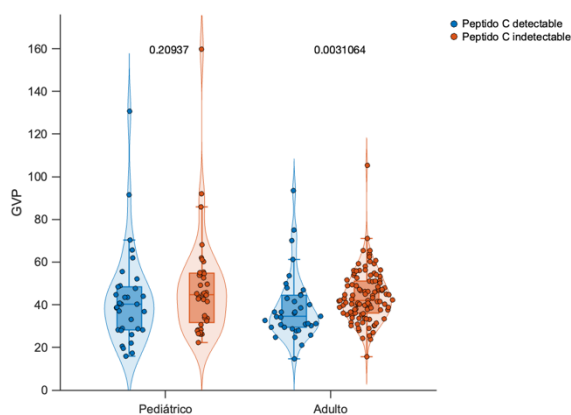


Figura 36. Distribución del parámetro GVP, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

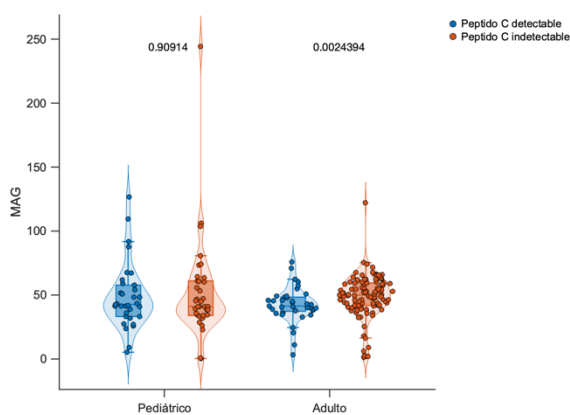


Figura 37. Distribución del parámetro MAG, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

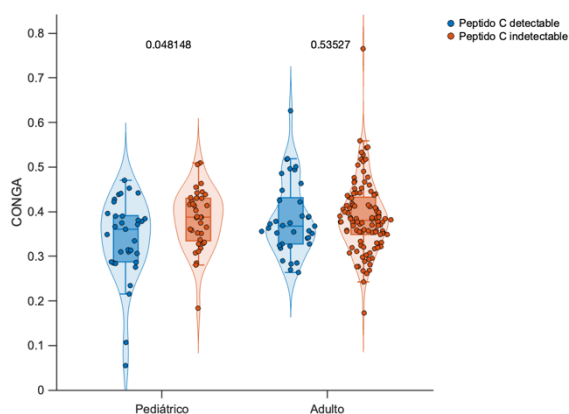


Figura 38. Distribución del parámetro CONGA, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

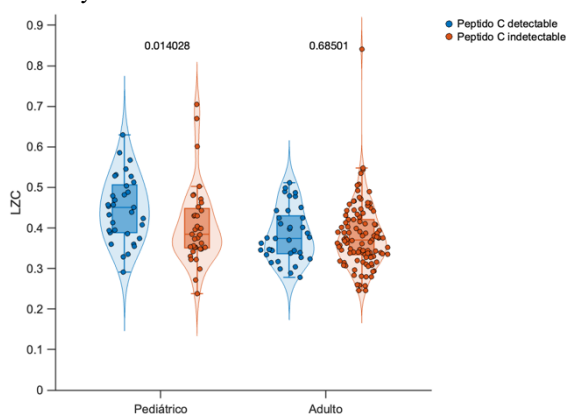


Figura 39. Distribución del parámetro LZC, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

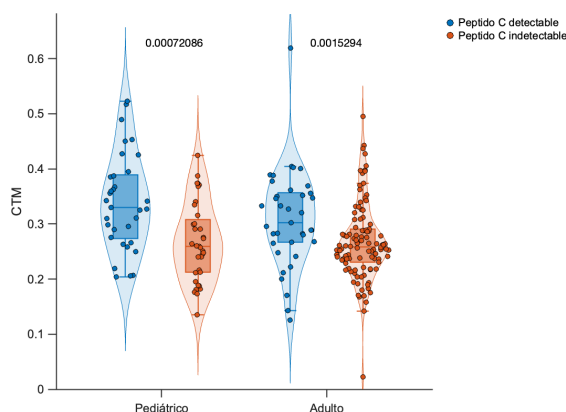


Figura 40. Distribución del parámetro CTM, a la izquierda para sujetos pediátricos y a la derecha para adultos. En naranja se observa la distribución en pacientes sin reserva pancreática y, en azul, los pacientes con reserva. Encima de cada comparación de distribuciones se encuentran los p-valores obtenidos con el test U de Mann-Whitney.

5.4 Análisis de correlaciones

Cada una de las redes de asociación que se muestran en este apartado aparecen en conjunto con su gráfico de estabilidad, de modo que se observa el corte que se ha establecido como umbral para cada una de ellas. Primeramente, observamos en la Figura 41 la red de asociación con todos los parámetros de VG. Junto a ella, su gráfica de estabilidad, en la que el intervalo de confianza del 95% corta en el cero para un valor de correlación mediana de ± 0.2 , de modo que la red es estable con un umbral de 0.2. Se ha generado la misma red, pero separando a los pacientes en dos grupos: aquellos con péptido C detectable y los que tienen péptido C indetectable. Estas dos redes, junto con sus gráficos de estabilidad, también se muestran en la Figura 41. De nuevo, el punto de corte con el cero se produce más o menos para un valor de correlación de ± 0.2 .

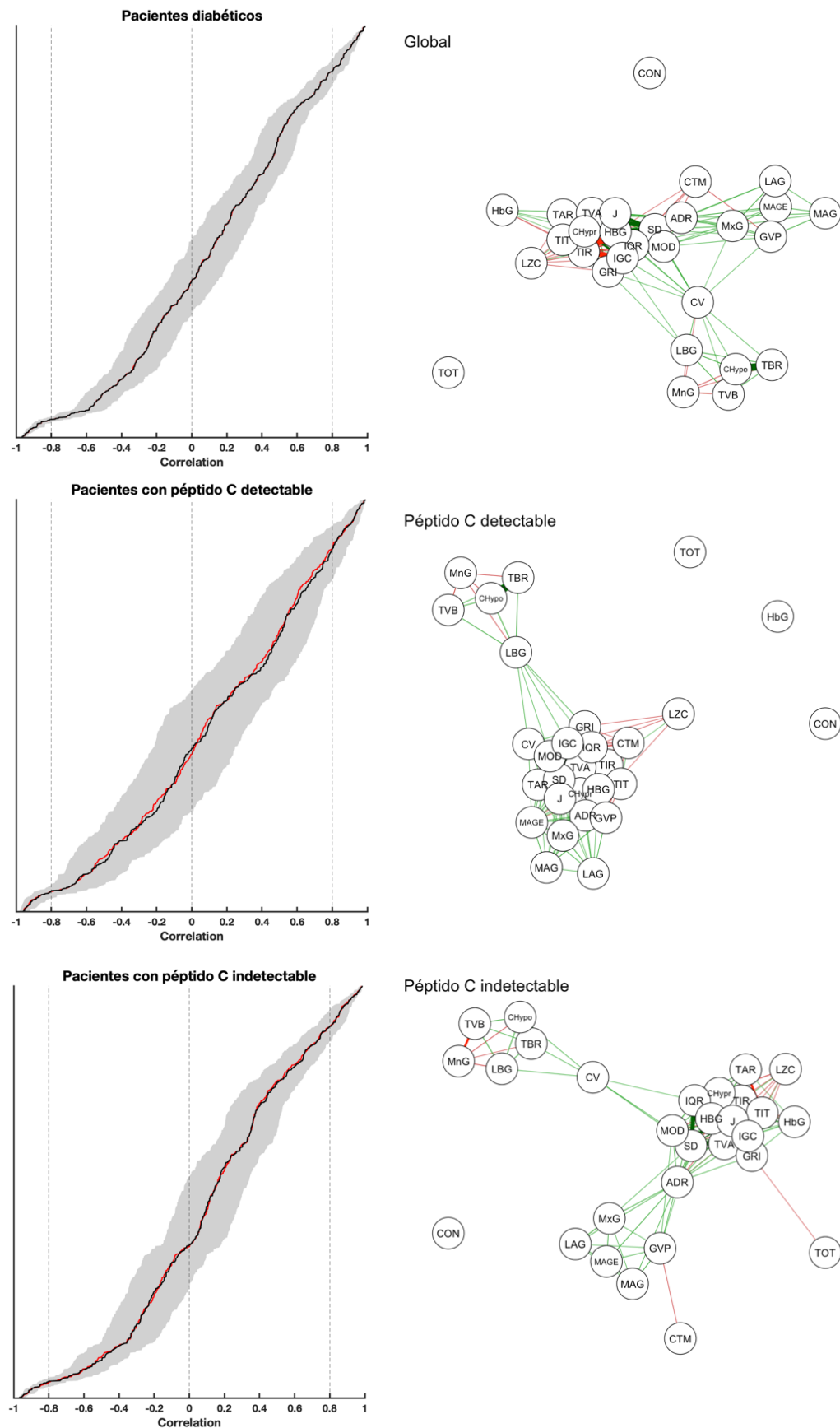
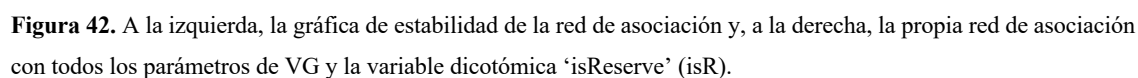


Figura 41. A la izquierda, se muestran los gráficos de estabilidad y, a la derecha, las redes de asociación con un umbral de 0.5. Arriba del todo se observa la red global de toda la base de datos, en medio la red que solo incluye los pacientes con péptido C detectable y abajo la red con los pacientes que tienen péptido C indetectable.

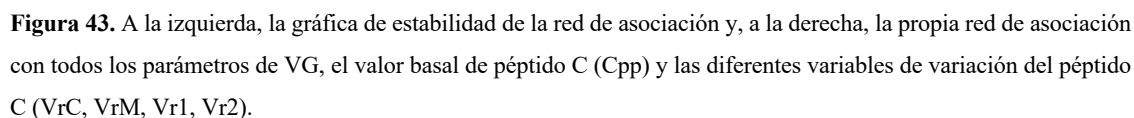
Para observar solo las correlaciones más fuertes, el umbral se ha establecido un poco más arriba, en 0.5, para las tres redes de asociación. Esto nos permite observar ciertos clústeres, que son más o menos los mismos en todas ellas. Por un lado, hay un grupo de parámetros que nos indican la carga o tiempo en hipoglucemia del paciente, donde se incluyen TBR2 (TVB), TBR1 (TBR), MinGly (MnG), CHypo y LBGI. Se observa otro pequeño grupo con parámetros que miden las oscilaciones de glucosa, donde englobamos MAGE, LAGE (LAG), MAG, GVP y MaxGly (MxG). En el centro de la red, se observa un grupo de parámetros relacionados con la carga de hiperglucemia y el riesgo glucémico. Por un lado, en relación con la hiperglucemia están HBGI, CHyper (CHypr), TAR1 (TAR) y TAR2 (TVA). Por otro, parámetros de riesgo o control glucémico son IGC, GRI, J, TIR, TITR (TIT), MODD (MOD) y ADRR (ADR). Además, se incluyen dos medidas clásicas de dispersión, SD e IQR. Entre medias del grupo de la carga de hipoglucemia y el de la carga de hiperglucemia y control glucémico se encuentra el parámetro CV, que es el que se emplea en la práctica clínica. Por último, hay otros parámetros que no presentan una pertenencia clara a ningún grupo, que son HbA1c (HbG), TOTR (TOT), CONGA (CON), CTM y LZC. Entre ellos, se encuentran todos los parámetros de análisis no lineal.

A continuación, se ha generado la misma red, pero incluyendo la variable dicotómica 'isReserve' (isR), que vale 0 si el paciente tiene un nivel de péptido C indetectable y 1 si el nivel de péptido C es detectable. Esta, junto con su gráfico de estabilidad, se puede observar en la Figura 42. El corte del intervalo de confianza del 95% con el cero se produce para un valor de correlación de ± 0.2 , por lo que podemos establecer el umbral de la red en 0.2. Aunque se observa que la variable isR no queda englobada cerca de ninguno de los clústeres definidos, sí que presenta correlaciones más fuertes de 0.2 con la mayoría de las variables. De ellas, las correlaciones más fuertes se producen con 6 variables, que se encuentran por encima de 0.3: -0.342 con CV, -0.340 con IGC, -0.337 con SD, 0.328 con CTM, -0.321 con IQR y -0.306 con GRI.

Como último análisis, se incluyeron en la red las diferentes medidas de variación del péptido C: VarC (VrC), VarMonth (VrM), Var1 (Vr1) y Var2 (Vr2), junto con el valor basal de péptido C cada paciente, Cpep (Cpp). Como se observa en la Figura 43, en este caso la gráfica de estabilidad corta en el cero para un valor de correlación de alrededor de



0.35, por lo que hay que subir ligeramente el umbral en la red de asociación. Ante ello, podemos observar que todas las correlaciones de los parámetros de variación del péptido C con parámetros de VG son inferiores a este corte. Sin embargo, hay algunas correlaciones más fuertes entre el valor basal de péptido C y los parámetros de VG. De hecho, hay 6 variables cuyas correlaciones se encuentran por encima de 0.5: -0.612 con MODD, -0.572 con CV, -0.571 con IQR, -0.534 con IGC, -0.516 con SD y -0.504 con LBGI. Las variables más fuertemente relacionadas con VarC son MAGE (0.290), CTM (-0.229), HbA1c (0.218), MinGly (-0.215), LZC (-0.181) e IQR (0.181). En el caso de



Var1 son MinGly (-0.167), LBG1 (0.1513), MaxGly (-0.141), CHypo (0.131), TBR1 (0.129) e IQR (0.121). Más fuertemente correlacionadas con Var2, se encuentran MAGE (0.182), IQR (0.150), LAGE (0.105), CONGA (0.104), TAR1 (0.104) y MaxGly (-0.100). Por último, en relación con VarMonth encontramos IQR (0.189), CONGA (0.166), TBR1 (0.165), LBG1 (0.159), CV (0.153) y CHypo (0.139).

5.5 Estudio de clasificación

Para intentar establecer la clasificación de los pacientes en aquellos que presentan reserva pancreática y los que no, se han empleado diversos algoritmos de *Machine Learning* – LDA, LR y QDA. Tanto las redes de asociación como otros algoritmos de selección de variables – SFS y FCBF – se han aplicado para elegir los parámetros que se iban a incluir en los modelos. Primeramente, se muestran en la Figura 44 las curvas ROC de los modelos que emplean aquellos parámetros de VG pertenecientes a cada uno de los clústeres con un menor p -valor. Estos son IGC, CTM, CV, MAGE y MinGly.

Después, se ha empleado el algoritmo de selección SFS para escoger las variables, de tal manera que el criterio de parada era llegar a $n = 5$ variables. En este caso, las variables escogidas dependen del modelo de clasificación posterior, de modo que para LDA son IGC, GVP, MODD y CTM; para LR son las mismas, y para QDA son MODD, TOTR, LBG1, TBR y MinGly. Las curvas resultantes se muestran en la Figura 45.

El otro algoritmo automático que se va a emplear es FCBF combinado con *bootstrapping*, de modo que se van a incluir las variables que haya seleccionado FCBF en más de una cierta proporción de las iteraciones de *Bootstrap* realizadas. En este caso, las variables seleccionadas son independientes del modelo de clasificación posterior, como se observa en la Figura 46. El primer corte se establece en 50 iteraciones (5%), que

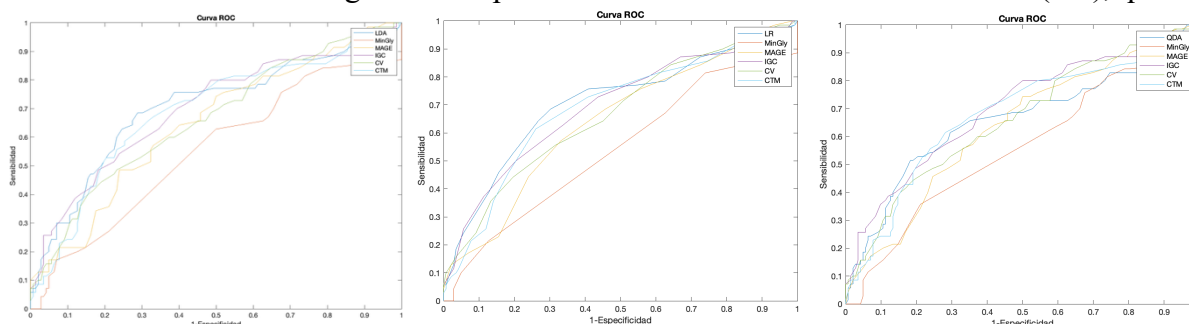


Figura 44. Curvas ROC de los modelos de clasificación generados a partir de las variables seleccionadas de los clústeres de las redes de asociación. A la izquierda se observa el modelo de LDA, en el medio el de LR y a la derecha el de QDA.

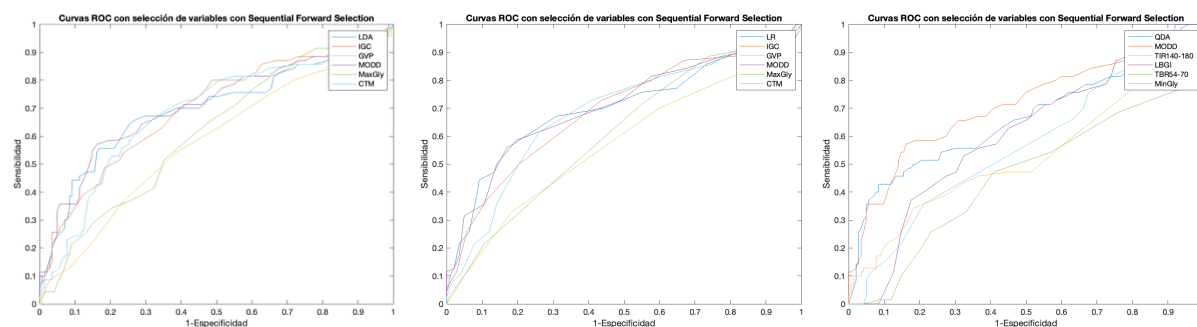


Figura 45. Curvas ROC de los modelos de clasificación generados a partir de las variables seleccionadas con el algoritmo *Sequential Forward Selection* (SFS). A la izquierda se observa el modelo de LDA, en el medio el de LR y a la derecha el de QDA.

selecciona 5 variables: SD, MODD, IQR, MinGly y CTM. Si ponemos un corte más estricto en 100 iteraciones (10%), solo se seleccionan 3 variables de las anteriores:

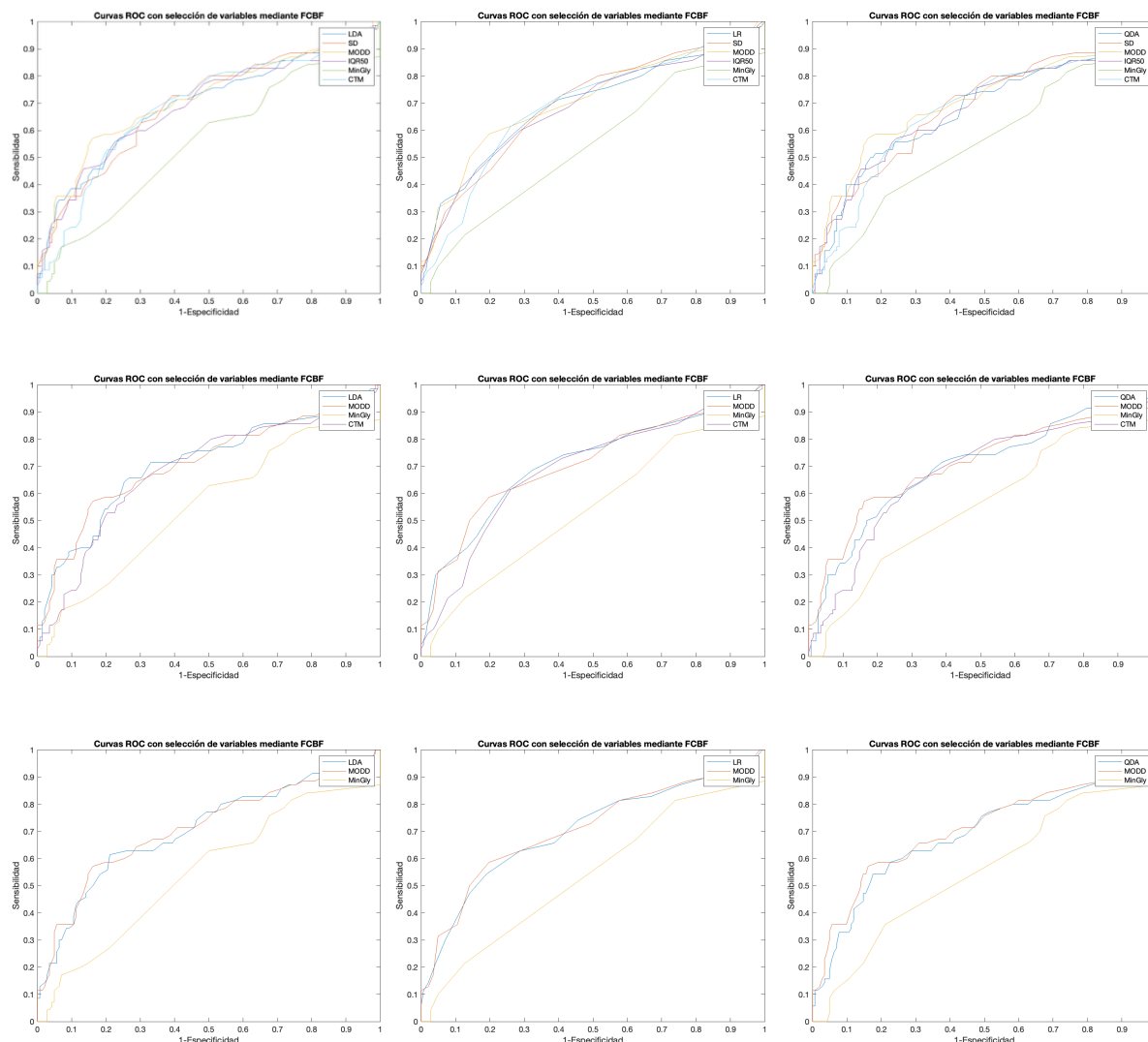


Figura 46. Curvas ROC de los modelos de clasificación generados a partir de las variables seleccionadas con el algoritmo FCBF. En la primera fila, se emplean todas las variables seleccionadas en más de 50 iteraciones; en la segunda, en más de 100, y en la tercera, en más de 125. A la izquierda se observa el modelo de LDA, en el medio el de LR y a la derecha el de QDA.

MODD, MinGly y CTM. Finalmente, con un corte de 125 iteraciones (12.5%), solamente quedan seleccionadas 2 variables: MODD y MinGly.

Por último, se van a escoger aquellas variables que presenten correlaciones más débiles con el resto de los parámetros, lo que se observa si subimos el umbral de las redes de asociación. Con esto conseguimos eliminar la redundancia, pero queremos variables relevantes, así que de entre las variables que quedan desconectadas se van a tomar las que tengan un mayor valor de SU con la variable *target*. Si tomamos un umbral de 0.7, las variables seleccionadas son CTM, LZC, CONGA, TOTR y HbA1c. Si subimos el umbral a 0.8, tenemos más variables no redundantes de donde seleccionar, de tal modo que se seleccionan son CTM, MAGE, CV, ADRR y GVP. Las curvas para ambos casos se observan en la Figura 47.

Finalmente, se busca determinar el rendimiento de estos modelos por medio del área bajo la curva ROC (AUC), de tal forma que se compara el AUC del modelo de clasificación con el AUC que se obtendría si el modelo tuviese solo una de las variables que lo conforman. Los resultados de este análisis se pueden observar resumidos en la Tabla 7. Así, con los tres modelos de clasificación, la opción que nos permite obtener un

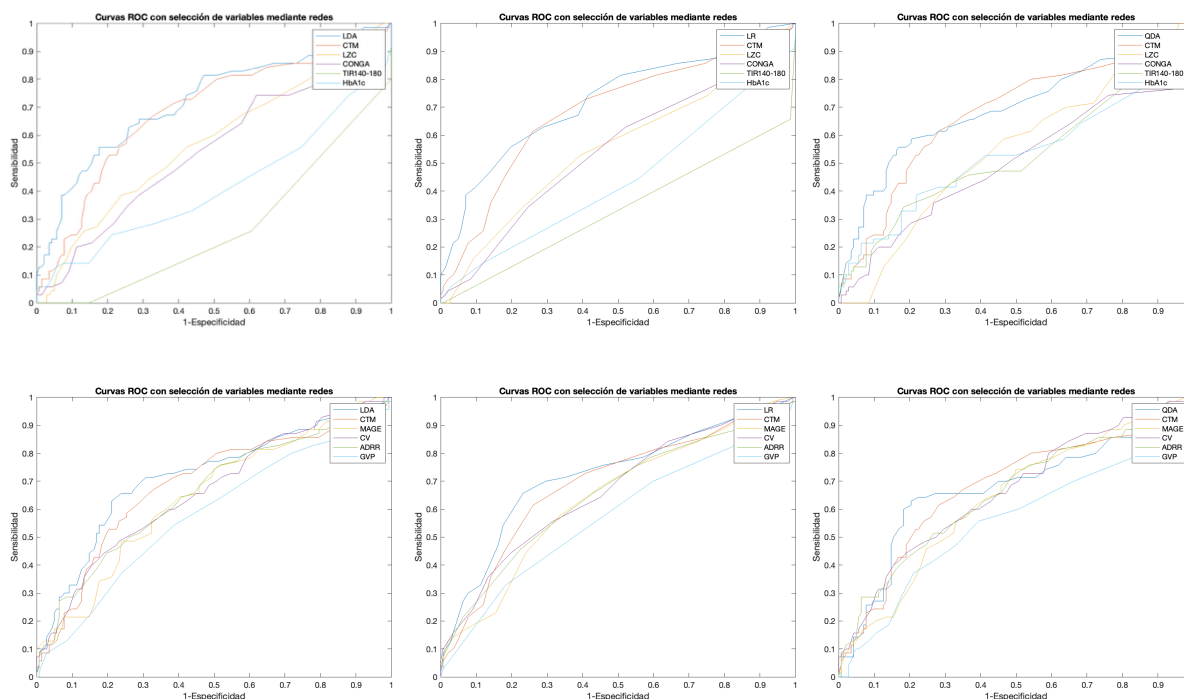


Figura 47. Curvas ROC de los modelos de clasificación generados a partir de las variables seleccionadas jugando con el umbral de las redes de asociación. En la primera fila, se emplean las variables más relevantes que no presentan correlaciones más fuertes de 0.7 con ningún otro parámetro y, en la segunda fila, aquellos sin correlaciones más fuertes de 0.8. A la izquierda se observa el modelo de LDA, en el medio el de LR y a la derecha el de QDA.

mayor AUC es aquella en la que la selección de variables se ha hecho tomando aquellas que no presenten correlaciones más fuertes de 0.7 con ningún otro parámetro y que tengan mayor valor de SU con la variable *target*. De los tres, el mayor es LDA, con un AUC de 0.7222; luego LR, con 0.7214, y, por último, QDA, con 0.6895. Los valores de AUC de los modelos son casi siempre similares al de la variable con mayor AUC e incluso, a veces, son más bajos. También podemos observar que el parámetro que proporciona un mayor valor de AUC con cualquiera de los tres modelos es MODD, con un AUC de 0.7112 si se emplea LDA, de 0.7073 si se emplea LR y de 0.7029 si se emplea QDA.

Método	LDA	Var1	AUC1	Var2	AUC2	Var3	AUC3	Var4	AUC4	Var5	AUC5
Selección manual	0.7041	IGC	0.6969	CTM	0.6899	CV	0.6647	MAGE	0.6481	MinGly	0.5512
<i>Sequential Forward Selection (SFS)</i>	0.7006	MODD	0.7112	IGC	0.6969	CTM	0.6899	MaxGly	0.6139	GVP	0.5766
FCBF (5%)	0.6943	MODD	0.7112	SD	0.6949	CTM	0.6899	IQR	0.6883	MinGly	0.5512
FCBF (10%)	0.7090	MODD	0.7112	CTM	0.6899	MinGly	0.5512		–		–
FCBF (12.5%)	0.7063	MODD	0.7112	MinGly	0.5512		–		–		–
Redes de asociación (umbral de 0.7)	0.7222	CTM	0.6899	LZC	0.5760	CONGA	0.5480	HbA1c	0.4136	TOTR	0.2673
Redes de asociación (umbral de 0.8)	0.7189	CTM	0.6899	CV	0.6647	ADRR	0.6565	MAGE	0.6481	GVP	0.5766

Método	LR	Var1	AUC1	Var2	AUC2	Var3	AUC3	Var4	AUC4	Var5	AUC5
Selección manual	0.7072	IGC	0.6920	CTM	0.6899	CV	0.6668	MAGE	0.6436	MinGly	0.5403
<i>Sequential Forward Selection (SFS)</i>	0.7042	MODD	0.7073	IGC	0.6920	CTM	0.6899	MaxGly	0.6104	GVP	0.5733
FCBF (5%)	0.6963	MODD	0.7073	SD	0.6980	CTM	0.6899	IQR	0.6863	MinGly	0.5403
FCBF (10%)	0.7095	MODD	0.7073	CTM	0.6899	MinGly	0.5403		–		–
FCBF (12.5%)	0.7031	MODD	0.7073	MinGly	0.5403		–		–		–
Redes de asociación (umbral de 0.7)	0.7214	CTM	0.6899	LZC	0.5561	CONGA	0.5488	HbA1c	0.4391	TOTR	0.3330
Redes de asociación (umbral de 0.8)	0.7193	CTM	0.6899	CV	0.6668	ADRR	0.6518	MAGE	0.6436	GVP	0.5733

Método	QDA	Var1	AUC1	Var2	AUC2	Var3	AUC3	Var4	AUC4	Var5	AUC5
Selección manual	0.6571	IGC	0.6929	CTM	0.6697	CV	0.6629	MAGE	0.6379	MinGly	0.5496
<i>Sequential Forward Selection (SFS)</i>	0.6596	MODD	0.7029	LBGI	0.6069	TOTR	0.5275	MinGly	0.5496	TBR	0.4539
FCBF (5%)	0.6622	MODD	0.7029	SD	0.6863	IQR	0.6737	CTM	0.6697	MinGly	0.5496
FCBF (10%)	0.6901	MODD	0.7029	CTM	0.6697	MinGly	0.5496		–		–
FCBF (12.5%)	0.6848	MODD	0.7029	MinGly	0.5496		–		–		–
Redes de asociación (umbral de 0.7)	0.6895	CTM	0.6697	LZC	0.5486	HbA1c	0.5309	TOTR	0.5275	CONGA	0.4948
Redes de asociación (umbral de 0.8)	0.6672	CTM	0.6697	CV	0.6629	ADRR	0.6492	MAGE	0.6379	GVP	0.5486

Tabla 7. Valores del área bajo la curva ROC (AUC) obtenidos con los diversos modelos de clasificación para cada uno de los métodos de selección de variables aplicados. Junto al AUC de los modelos, se muestran los AUC de las variables que lo conforman si cada una de ellas fuera la única que se emplea para el modelo, en orden decreciente de valor de AUC.

Capítulo 6. **Discusión y limitaciones**

6.1 Introducción.....	97
6.2 Análisis de correlaciones.....	97
6.3 Estudio de clasificación.....	100
6.3 Limitaciones.....	101

6.1 Introducción

La DM1 es una patología crónica que presenta numerosas complicaciones microvasculares y macrovasculares, y la aparición y el desarrollo de la mayoría de ellas depende del control glucémico que presente el paciente. En este ámbito, el concepto de VG es crucial, ya que se espera que un paciente bien controlado presente una menor variabilidad (Lazar et al., 2023). Para poder caracterizar correctamente el concepto de VG, no se puede emplear únicamente un parámetro, ya que presenta una dimensionalidad muy elevada que engloba numerosos aspectos.

En este TFG se ha llevado a cabo un análisis exploratorio de los diferentes parámetros de VG empleados a nivel clínico, así como algunas características derivadas del análisis no lineal de series temporales complejas. Esto ha permitido establecer las correlaciones entre ellos, de tal forma que se ve cuáles son más relevantes para la clasificación de pacientes en función de su nivel de péptido C (detectable vs. indetectable) y cuántos nos aportan realmente información complementaria entre sí. A partir de ello, se ha hecho uso de diversos modelos de clasificación para combinar los parámetros más relevantes y menos redundantes para el caso de estudio.

En este capítulo se van a abordar y analizar los principales resultados obtenidos a partir de las redes de asociación y los diferentes modelos de clasificación. Además, se van a exponer los aspectos a mejorar de este proyecto.

6.2 Análisis de correlaciones

Los clústeres establecidos tras la observación de las redes de asociación constatan los aspectos en común que tienen ciertos de los parámetros de VG. En el caso del clúster formado por TBR1, TBR2, MinGly, CHypo y LBGI, todos estos parámetros son indicadores del tiempo de hipoglucemia de una manera o de otra. De hecho, ya se ha visto en estudios previos, como el de (El-Laboudi et al., 2016) o el de (Gómez et al., 2019), que LBGI es uno de los parámetros de VG que mejor caracteriza el riesgo de hipoglucemia y que más relación tiene con el tiempo de hipoglucemia (TBR1 y TBR2). También el CV se observó que tenía correlación con el tiempo de hipoglucemia en el estudio de (El-Laboudi et al., 2016), lo que también se muestra claramente en las redes de asociación. De hecho, este parámetro presenta correlaciones más fuertes de 0.5 con todos los parámetros de hipoglucemia.

El grupo formado por MAGE, LAGE, MAG, GVP y MaxGly es un conjunto de parámetros que miden grandes oscilaciones en la curva de glucemia. Al contrario que en el estudio de (Kohnert et al., 2013), MAGE aparece relacionada con MAG, probablemente por las diferencias en el algoritmo de cálculo de MAGE. No obstante, la correlación entre ellos es lógica, ya que al final MAGE y LAGE miden oscilaciones de manera absoluta; mientras que MAG y GVP tienen en cuenta el intervalo temporal. En el caso de MaxGly se puede considerar una oscilación que no es comparada con el valor basal, ya que solo mide el máximo. De esta forma, la información que nos aportan es muy similar.

Por su parte, el grupo de parámetros más extenso incluye parámetros de hiperglucemia como HBGI, CHyper, TAR1 y TAR2; parámetros de control o riesgo glucémico como IGC, GRI, J, TIR, TITR y ADRR, y parámetros de dispersión como SD, IQR y MODD. En el estudio de (Rodbard, 2009) ya se identificó la relación entre los principales parámetros de dispersión, como SD, MODD e IQR. Aunque aquí también se incluye el parámetro MAGE, que en nuestro estudio queda englobado en otro clúster, si bien es verdad que está en proximidad a los otros parámetros de dispersión. En este mismo estudio se muestra que IGC presenta una alta correlación con HBGI/LBGI, lo que se corrobora en nuestro estudio. Y esto tiene mucho sentido, puesto que es calculado a partir de estos dos índices. También se ha observado en estudios previos el puente entre los parámetros de dispersión y los parámetros de control glucémico, como en el de (Tokutsu et al., 2022), donde se observa una fuerte correlación negativa entre IQR y TIR. Asimismo, en el estudio de (Pérez-López et al., 2023) se determina que el GRI está asociado con parámetros como HbA1c, SD, TAR2 y TBR2. Al contrario, en nuestro estudio, GRI se incluye en el clúster de los parámetros de hiperglucemia y control glucémico, de tal forma que no presenta correlaciones fuertes ni con HbA1c ni con TBR2. Por último, (Beck et al., 2019) observaron una alta correlación entre parámetros de hiperglucemia como HBGI y parámetros de control como TIR. Así, podemos decir que hay una base sólida para dictaminar que la dispersión de los datos nos puede indicar tanto la presencia de hiperglucemia como el grado de control glucémico que presenta un paciente, que viene descrito como la probabilidad de eventos extremos fuera del rango fisiológicos.

Como podemos comprobar, en la mayoría de los estudios realizados, las correlaciones que se establecen son entre parámetros de VG y descriptores clásicos en la clínica, como

la HbA1c, el TIR o el tiempo en hipoglucemia. Esto se debe principalmente a que son parámetros más interpretables, de modo que se establecen relaciones con variables conocidas, cuya implicación dentro de la patología es comprendida por los profesionales de la salud. Por tanto, el foco está en conseguir una mayor interpretabilidad de los parámetros más complejos que permita asemejarlos a aquellos que manejan los médicos en la práctica habitual. En este contexto, es necesario que exista una interpretabilidad de parámetros de análisis no lineal, como CONGA, LZC y CTM. Estos no presentan una pertenencia clara a ninguno de los grupos, lo que nos indica que añaden cierta dimensionalidad al problema de la VG.

Normalmente, en vez de evaluar el control glucémico de los pacientes, en la mayor parte de los estudios se estudian las correlaciones con la aparición de ciertas complicaciones asociadas a la DM. Un ejemplo es el estudio de (Akaza et al., 2018), en el que se establece la asociación entre MAGE y la neuropatía diabética periférica. El problema de esto es que simplemente evitamos la aparición de patologías asociadas. En cambio, en este estudio se pretende establecer la asociación con el péptido C, que es un marcador de protección en pacientes diabéticos. De esta manera, lo que se observa es que parámetros nos permiten determinar la funcionalidad de las células β que mantiene cada paciente. Los índices que se observan correlacionados con la variable de clasificación 'isReserve' son IGC, CTM, IQR, GRI, SD y CV. De ellos, CV ya se ha visto que tiene una alta correlación con los niveles de péptido C en el estudio de (Christensen et al., 2020), aunque en este caso es en pacientes con DM2. La mayor parte son índices de dispersión (IQR) o nos indican el riesgo del paciente (IGC y GRI), pero todos ellos son índices sencillos. En cambio, es interesante la aparición de CTM en este grupo, que además es significativo en el test U de Mann-Whitney, tanto para adultos como para sujetos pediátricos. A pesar de no ser una métrica específicamente calculada para los registros de glucemia, se observa que puede tener una gran utilidad en estudios posteriores, ya que es una de las que mayor relación tienen con el péptido C y, por tanto, con el control de los pacientes diabéticos. También se observa que todas las correlaciones son negativas, ya que cuanto mayor valor de los parámetros de VG, menor nivel de péptido C tiene el paciente. Esto ya se había concluido en estudios anteriores, como el de (Crăciun et al., 2018). Al contrario, CTM presenta una correlación positiva, ya que un mayor valor de esta variable nos indica una menor variabilidad y, por ende, un mayor nivel de péptido C.

Correlacionadas con el nivel basal de péptido C se encuentran IQR, IGC, LBGI, MODD, CV y SD. MODD y CV. Ya se había identificado que parámetros clásicos de dispersión como MODD, CV y también MAGE reducen su valor cuando aumenta la inmunorreactividad a péptido C del paciente en el estudio de (Ohara et al., 2019). No obstante, los resultados de las redes de asociación nos ofrecen una mayor variedad de parámetros que comparten información con el péptido C. De nuevo, todas las correlaciones son negativas. Todos estos parámetros que aparecen con una fuerte correlación con alguna de las variables del péptido C fueron significativos en el test U de Mann-Whitney. No obstante, no todos los parámetros significativos en esta prueba inicial han demostrado una fuerte correlación, lo que nos indica que, aunque las distribuciones de estas variables en los dos grupos (péptido C detectable vs. indetectable) son diferentes, esto puede ser más fruto del azar que de una capacidad de diferenciación entre grupos. Cabe destacar que con las variables de variación de péptido C, ninguna de las correlaciones ha sido significativa, de modo que la predicción de la evolución del nivel de péptido C no es tan evidente a partir de los diversos parámetros de VG.

6.3 Estudio de clasificación

En otros estudios, como en el de (Cichosz et al., 2025), se emplean índices de variabilidad glucémica como el TAR, TBR o GRI para separar entre pacientes sanos, con prediabetes y con DM1. Por tanto, solo se distingue entre los pacientes sanos y enfermos. En cambio, este estudio de clasificación busca separar entre aquellos pacientes que presentan un mejor control de la patología y los que tienen un control glucémico peor. Por tanto, se va un paso más allá, para poder ayudar a los profesionales sanitarios a la hora de establecer una monitorización más o menos estricta para un paciente.

En los resultados de los modelos de clasificación, se observa un mejor valor de AUC con LDA y LR que el obtenido con QDA. Esto se debe a que, aunque QDA es un modelo con una mayor flexibilidad, es menos robusto precisamente por este motivo (Chand & Vishwakarma, 2022). De hecho, al tener una gran cantidad de parámetros, es más sensible a cambios en los valores de entrada y su rendimiento es más bajo. A pesar de esta leve diferencia, los valores de AUC de los tres clasificadores no mejoran mucho en comparación con la clasificación que se puede obtener a partir de los parámetros de variabilidad de manera individual. Por tanto, existe un amplio rango de mejora en la

clasificación de pacientes con péptido C detectable vs. indetectable, puesto que la mejor AUC obtenida es de 0.7222.

Uno de los parámetros de VG que más se ha seleccionado es MODD, que presenta una AUC individual de en torno a 0.70-0.71 con los tres modelos de clasificación. Esta medida nos establece patrones de comidas de un paciente gracias a la variabilidad interdiaria y ya se ha visto que está correlacionada con el nivel de péptido C en el estudio de (Ohara et al., 2019). Su gran relación con la efectividad del tratamiento con CGM en los pacientes podría explicar este alto valor de AUC, ya que un paciente con péptido C detectable ($isR = 1$) se puede beneficiar mucho más con un tratamiento y conseguir un buen control glucémico. En cambio, un paciente con péptido C indetectable ($isR = 0$) no se va a beneficiar tanto y le va a costar más alcanzar un buen control a pesar de presentar CGM. De nuevo, otro de los parámetros que aparece en la mayoría de los modelos de clasificación es CTM, cuya implicación dentro de la VG está aún por determinar, pero que no está directamente relacionado con ninguno de los aspectos descritos por parámetros de VG clásicos. Otro de los parámetros con elevada AUC es IGC, que se selecciona en varios modelos y nos da un valor del riesgo glucémico del paciente, ya que es una combinación de LBGi y HBGI. Otros que también han sido seleccionados en muchos modelos son CV, MinGly o MAGE. CV y MAGE son parámetros que han sido muy estudiados y que se mantienen como una dimensión imprescindible para la caracterización del control de los pacientes, aunque probablemente no sean los más exactos.

En cuanto a los métodos de selección, el uso de las redes de asociación para valorar la redundancia excede el rendimiento de FCBF en la mayor parte de los modelos, excepto en el caso de QDA. Por ello, este abordaje presenta un gran potencial para poder interrelacionar los diferentes parámetros de VG y seleccionar aquellos que no son redundantes. De hecho, en este estudio parece primar este criterio por encima de la relevancia, ya que al establecer el umbral en 0.8, se pudieron incluir variables en el modelo con una mayor relevancia (caracterizada por el valor de SU con la variable *target*), pero esto empeoró el resultado de AUC de los modelos.

6.4 Limitaciones

Una de las principales limitaciones de este trabajo es que existen muchas variables que influyen en el control glucémico del paciente que no se pueden caracterizar. Entre

ellas están los momentos de las comidas y la cantidad de carbohidratos ingeridos, el momento, duración y dosis de los bolos de insulina que recibe el paciente o la presencia, duración e intensidad de períodos de ejercicio. Todos estos factores tienen relevancia para un buen control del paciente diabético, pero no se monitorizan de manera habitual en la clínica. De esta forma, se pierde mucha información útil y los análisis realizados acaban siendo una simplificación de la situación real.

A esto se suma que nuestra variable *target*, que es el péptido C, no es una medida totalmente fiable (Leighton et al., 2017; Venugopal et al., 2023). A pesar de que siempre se intenta realizar en condiciones de ayuno para que no haya alteraciones y que refleje lo más fielmente posible la reserva pancreática del paciente, su determinación no es perfecta. Por tanto, la medida que se obtiene de péptido C en sangre es solo aproximada y, además, puede haber alteraciones posteriores al hacer el análisis de la muestra de sangre. Esto se observa en diversos pacientes de la base de datos, que presentan repuntes de péptido C o incluso alguno que tiene subidas en el nivel de péptido C en vez de bajadas. Por tanto, emplear este parámetro como *gold standard* del control glucémico no es la mejor opción si queremos ser rigurosos, aunque es la única medida que se puede realizar de manera directa actualmente. Además, el número de pacientes que presentan más de una medición de péptido C son pocos, y generalmente esta medida se hace a bastante corto plazo. Por este motivo, no hay datos suficientes para realizar una correcta predicción de la evolución del péptido C en los pacientes.

A pesar de haber realiza una búsqueda exhaustiva de los parámetros de VG que se emplean a nivel clínico, no todos ellos se han incluido en el estudio. Por un lado, parámetros como GRADE (*glycemic risk assessment diabetic equation*), LI (*liability index*) y el índice M no se han implementado porque no aparecían en las revisiones bibliográficas que se consultaron inicialmente (Lazar et al., 2023; Leksic et al., 2022). Otros como MIME (*mean indices of meal excursions*) no han podido implementarse debido a que se necesitan datos sobre el pico posprandial o el momento de inicio de las comidas (Breyton et al., 2021). De entre los parámetros que sí se han implementado, todos se han calculado para una ventana de 2 semanas, aunque en algunos estudios, como en el de (Bowler et al., 2024), se emplean períodos de tiempo más cortos para algunas variables. Por ejemplo, SD y MAGE se calculan para períodos de 1 día, ya que miden la variabilidad intradiaria, y MODD para períodos de varios días (en torno a 4), puesto que mide la variabilidad interdiaria.

Capítulo 7. Conclusiones y líneas futuras

7.1 Introducción.....	105
7.2 Cumplimiento de los objetivos del trabajo de fin de grado.....	105
7.3 Conclusiones y aportaciones originales.....	106
7.4 Líneas futuras.....	108

7.1 Introducción

En el último capítulo de este TFG se resumen los aspectos más relevantes que se han tratado a lo largo del trabajo. Para ello, se inicia con una evaluación del cumplimiento de los objetivos establecidos al inicio del TFG. Después, se extraen las principales conclusiones que se han obtenido a raíz del proyecto y las líneas de investigación que se pueden explotar en un futuro.

7.2 Cumplimiento de los objetivos del trabajo de fin de grado

Debido a la ausencia de una medida *gold standard* para caracterizar la VG en los pacientes con DM1, este TFG se ha centrado en realizar un análisis exploratorio de los parámetros existentes para determinar qué combinación de variables refleja mejor los diferentes aspectos de la VG y es capaz de estimar el estado general del paciente diabético. Para ello, se ha hecho uso de las redes de asociación, que han permitido caracterizar las relaciones entre los parámetros de VG y de estos con la variable *target*, que es el péptido C. Además, se han utilizado varios algoritmos de clasificación englobados dentro de las técnicas de *Machine Learning* para establecer un modelo de clasificación de pacientes en aquellos con péptido C detectable vs. indetectable. Con este objetivo en mente, se plantearon una serie de objetivos específicos en el capítulo 1. En este último capítulo, se valorará el grado de consecución de cada uno de ellos:

- i. Se ha llevado a cabo una búsqueda bibliográfica sobre la DM, que sobre todo se ha centrado en la DM1. Gracias a ello, se han podido entender mejor los factores que influyen tanto en su aparición como en el desarrollo de las diversas complicaciones microvasculares y macrovasculares.
- ii. Se han recogido los datos clínicos de los pacientes y se han descargado sus datos de glucemia, de tal forma que se ha podido crear la base de datos de partida. Con estos datos, se han conseguido extraer todos los parámetros.
- iii. Se ha realizado una revisión de los parámetros de VG a nivel clínico, así como aquellos derivados del análisis no lineal de series temporales complejas. Esto nos permite conocer las fórmulas para implementar estos parámetros y llevar a cabo la extracción de variables a partir de las curvas de glucemia.

- iv. Se han codificado correctamente las variables cualitativas y se han imputado los valores perdidos existentes. Así, se ha logrado obtener la base de datos completa sobre la que se han extraído los resultados.
- v. Se ha llevado a cabo un análisis poblacional para observar la distribución de las diversas variables extraídas. Esto nos permite hacer un análisis preliminar para determinar si los parámetros de VG van a clasificar bien entre pacientes con y sin reserva pancreática.
- vi. Se ha realizado un análisis de correlaciones por medio de redes de asociación. Con ello, se han establecido los principales clústeres de variables que aportan información similar sobre la VG, y se han detectado aquellas que se asocian en mayor medida con el nivel de péptido C.
- vii. Se han empleado diversos modelos de clasificación para distinguir entre pacientes con péptido C detectable e indetectable, en las que se han incluido diversos parámetros de VG. Con ello, se pretende determinar si una combinación de variables es capaz de clasificar los pacientes con un rendimiento superior a cualquiera de estas variables de manera individual.
- viii. Se han comentado los resultados expuestos en el TFG, que se han comparado con resultados de estudios anteriores. A partir de ellos, se han extraído las conclusiones oportunas, además de exponer las principales limitaciones y líneas futuras del trabajo.

En conclusión, podemos afirmar que los objetivos específicos planteados en un inicio se han abordado correctamente en este TFG.

7.3 Conclusiones y aportaciones originales

A partir de los análisis llevados a cabo en este TFG, se puede concluir que **una combinación de los parámetros de VG no es capaz de mejorar la clasificación en péptido C detectable vs. indetectable con respecto a las medidas individuales y, por tanto, no es capaz de determinar el estado de control glucémico que presenta el paciente.**

En base a los resultados y la discusión expuestos, podemos extraer las siguientes conclusiones:

1. La mayoría de los parámetros clásicos de VG comparten entre ellos una gran cantidad de información, como se observa en las redes de asociación, si bien se puede detectar la presencia de ciertos clústeres. Esto nos indica que los diversos aspectos de la VG están bastante correlacionados entre ellos, de manera que unos afectan a otros de manera directa y separarlos con diversas métricas no es tan sencillo.
2. Algunos de los parámetros de análisis no lineal, especialmente CTM, parecen tener una estrecha correlación con los niveles de péptido C y un buen rendimiento de clasificación. De hecho, es necesario establecer qué aspectos de la VG están midiendo exactamente, ya que sus correlaciones con el resto de los parámetros clínicos son más débiles. Por tanto, análisis posteriores que incluyan parámetros no específicos del ámbito de la diabetes pueden ser potencialmente interesantes.
3. El empleo de algoritmos de *Machine Learning* como LDA, LR o QDA no parece ofrecer un buen resultado de clasificación. De hecho, en ocasiones la combinación de varios parámetros empeora en vez de mejorar el valor de AUC. Probablemente debido a la gran información que comparten, el uso de diversos parámetros incluye más información redundante que aspectos complementarios sobre el control glucémico.
4. Hay parámetros que, por si solos, presentan un buen rendimiento caracterizado por la curva de AUC, como MODD, IGC o CTM, de tal forma que son capaces de clasificar bastante bien a los pacientes. Aunque no sea de forma perfecta, nos indican que la VG sí que tiene mucha relación con el control glucémico.

Este análisis exploratorio, más que ofrecer unos resultados concluyentes, es un punto de partida en el análisis del control glucémico de pacientes diabéticos. La VG es un concepto muy complejo que parece establecer una relación estrecha con este, pero no queda completamente claro cómo caracterizarla correctamente. Por tanto, este TFG sirve

de base para llevar a cabo más análisis en relación con el péptido C, el control glucémico y la VG.

7.4 Líneas futuras

A partir de este trabajo, se pueden establecer diversas líneas futuras que intenten paliar las diversas limitaciones encontradas. Primero, es necesario obtener una mayor cantidad de datos para generar una base de datos con una mayor potencia. Para ello, es necesario que exista una mayoría de pacientes con diversas medidas del péptido C. Esto es tan sencillo como incluir esta medida en los análisis de sangre de los pacientes de manera periódica. Además, sería beneficioso incluir otros factores que influyen en el control glucémico de los pacientes. Para las dosis de insulina, una buena solución es el uso de plumas inteligentes, cuyos bolos quedan registrados en la CGM. Para obtener información sobre las comidas, es necesario depender de la responsabilidad del propio paciente. En cuanto a la caracterización del ejercicio, se podrían emplear acelerómetros para cuantificar su intensidad.

En posteriores estudios, se debería ampliar la cantidad de medidas de VG incluidas, de tal forma que se puedan observar más aspectos de la variabilidad. Entre ellas se pueden incluir medidas como MIME, que aporta información sobre los picos posprandiales y, en consecuencia, sobre los patrones de comida de manera más directa. También existen nuevos abordajes relacionados con la VG que no están basados en parámetros. Un ejemplo es el uso del concepto de glucodensidad. Este consiste en establecer un histograma continuo de la frecuencia de aparición de cada nivel de glucemia en un paciente durante un período de tiempo concreto (Cui et al., 2023; Matabuena et al., 2024). Con estos gráficos, se consigue una caracterización más dinámica de la variabilidad del paciente que con un simple parámetro y nos permite llevar a cabo diversos análisis a partir de él.

Aparte, el análisis con redes de asociación se ha hecho de manera cualitativa, pero una mejor manera de establecer los clústeres es llevar a cabo un análisis de modularidad (Rubinov & Sporns, 2010). Además, esto se puede combinar con un análisis de centralidad para determinar la distancia de cada nodo a los otros nodos pertenecientes a su mismo grupo. Con esto, se pueden establecer los grupos en las redes de asociación de manera más clara y separar los diversos aspectos de la VG.

Por último, gracias a la presencia de un análisis más completo y con mayor cantidad de parámetros, se podría plantear un estudio de regresión. En él, se pueden incluir diversos algoritmos de regresión que nos permitan estimar la evolución del péptido C en los pacientes diabéticos. Esto puede ser de gran utilidad para conocer en cuánto tiempo un paciente con DM1 va a perder toda su reserva pancreática y así poder establecer un tratamiento más personalizado para cada uno de ellos.

Bibliografía

Ai, D., Pan, H., Li, X., Wu, M., & Xia, L. C. (2019). Association network analysis identifies enzymatic components of gut microbiota that significantly differ between colorectal cancer patients and healthy controls. *PeerJ*, 7, e7315. <https://doi.org/10.7717/peerj.7315>

Armonaite, K., Conti, L., Olejarczyk, E., & Tecchio, F. (2024). Insights on neural signal analysis with Higuchi fractal dimension. *Communications in Applied and Industrial Mathematics*, 15(2), 17–27. <https://doi.org/10.2478/caim-2024-0009>

Arokiasamy, P., Salvi, S., & Selvamani, Y. (2021). Global Burden of Diabetes Mellitus. In *Handbook of Global Health* (pp. 1–44). Springer International Publishing. https://doi.org/10.1007/978-3-030-05325-3_28-1

Azizi, T. (2022). On the fractal geometry of different heart rhythms. *Chaos, Solitons & Fractals: X*, 9, 100085. <https://doi.org/10.1016/j.csfx.2022.100085>

Baghurst, P. A. (2011). Calculating the Mean Amplitude of Glycemic Excursion from Continuous Glucose Monitoring Data: An Automated Algorithm. *Diabetes Technology & Therapeutics*, 13(3), 296–302. <https://doi.org/10.1089/dia.2010.0090>

Bahillo-Currieses, P., Fernández Velasco, P., Pérez-López, P., Vidueira Martínez, A. M., Nieto de la Marca, M. de la O., & Díaz-Soto, G. (2024). Utility of time in tight range (TITR) in evaluating metabolic control in pediatric and adult patients with type 1 diabetes in treatment with advanced hybrid closed-loop systems. *Endocrine*, 86(2), 539–545. <https://doi.org/10.1007/s12020-024-03881-6>

Barabási, A.-L. (2016). *Network Science* (1st ed.). Cambridge University Press. <http://networksciencebook.com/>

Battelino, T., Danne, T., Bergenstal, R. M., Amiel, S. A., Beck, R., Biester, T., Bosi, E., Buckingham, B. A., Cefalu, W. T., Close, K. L., Cobelli, C., Dassau, E., DeVries, J. H., Donaghue, K. C., Dovc, K., Doyle, F. J., Garg, S., Grunberger, G., Heller, S., Heinemann, L., Hirsch, I. B., Hovorka, R., Jia, W., Kordonouri, O., Kovatchev, B., Kowalski, A., Laffel, L., Levine, B., Mayorov, A., Mathieu, C., Murphy, H. R., Nimri, R., Nørgaard, K., Parkin, C. G., Renard, E., Rodbard, D., Saboo, B., Schatz, D., Stoner,

K., Urakami, T., Weinzimer, S. A., & Phillip, M. (2019). Clinical Targets for Continuous Glucose Monitoring Data Interpretation: Recommendations from the International Consensus on Time in Range. *Diabetes Care*, 42(8), 1593–1603. <https://doi.org/10.2337/dci19-0028>

Beck, R. W., Bergenstal, R. M., Cheng, P., Kollman, C., Carlson, A. L., Johnson, M. L., & Rodbard, D. (2019). The Relationships Between Time in Range, Hyperglycemia Metrics, and HbA1c. *Journal of Diabetes Science and Technology*, 13(4), 614–626. <https://doi.org/10.1177/1932296818822496>

Bergenstal, R. M. (2015). Glycemic Variability and Diabetes Complications: Does It Matter? Simply Put, There Are Better Glycemic Markers! *Diabetes Care*, 38(8), 1615–1621. <https://doi.org/10.2337/dc15-0099>

Bidonde, J., Fagerlund, B. C., Frønsdal, K. B., Lund, U. H., & Robberstad, B. (2017). FreeStyle Libre Flash Glucose Self-Monitoring System: A Single-Technology Assessment. *Norwegian Institute of Public Health*

Blum, A. (2018). Freestyle Libre Glucose Monitoring System. *Clinical Diabetes*, 36(2), 203–204. <https://doi.org/10.2337/cd17-0130>

Borowska, M. (2021). Multiscale Permutation Lempel–Ziv Complexity Measure for Biomedical Signal Analysis: Interpretation and Application to Focal EEG Signals. *Entropy*, 23(7), 832. <https://doi.org/10.3390/e23070832>

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16(1), 5–13. <https://doi.org/10.1002/wps.20375>

Borsboom, D., & Cramer, A. O. J. (2013). Network Analysis: An Integrative Approach to the Structure of Psychopathology. *Annual Review of Clinical Psychology*, 9(1), 91–121. <https://doi.org/10.1146/annurev-clinpsy-050212-185608>

Bowler, A.-L. M., Burke, L. M., Coffey, V. G., & Cox, G. R. (2024). Day-to-Day Glycemic Variability Using Continuous Glucose Monitors in Endurance Athletes. *Journal of Diabetes Science and Technology*. <https://doi.org/10.1177/19322968241250355>

Breyton, A.-E., Lambert-Porcheron, S., Laville, M., Vinoy, S., & Nazare, J.-A. (2021). CGMS and Glycemic Variability, Relevance in Clinical Research to Evaluate Interventions in T2D, a Literature Review. *Frontiers in Endocrinology*, 12. <https://doi.org/10.3389/fendo.2021.666008>

Broll, S., Urbanek, J., Buchanan, D., Chun, E., Muschelli, J., Punjabi, N. M., & Gaynanova, I. (2021). Interpreting blood GLUcose data with R package iglu. *PLOS ONE*, 16(4), e0248560. <https://doi.org/10.1371/journal.pone.0248560>

Buuren, S. van, & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45(3). <https://doi.org/10.18637/jss.v045.i03>

Chand, S., & Vishwakarma, V. P. (2022). Application of quadratic discriminant analysis algorithm for the classification of acute leukemia using microscopic image data. *Advances in Applied Mathematics Science*, 21(5), 2737–2750.

Christensen, M. B., Gæde, P., Hommel, E., Gotfredsen, A., & Nørgaard, K. (2020). Glycaemic variability and hypoglycaemia are associated with C-peptide levels in insulin-treated type 2 diabetes. *Diabetes & Metabolism*, 46(1), 61–65. <https://doi.org/10.1016/j.diabet.2019.02.002>

Christou, M. A., Christou, P. A., Kyriakopoulos, C., Christou, G. A., & Tigas, S. (2023). Effects of Hypoglycemia on Cardiovascular Function in Patients with Diabetes. *International Journal of Molecular Sciences*, 24(11), 9357. <https://doi.org/10.3390/ijms24119357>

Cichosz, S. L., Kronborg, T., Laugesen, E., Hangaard, S., Fleischer, J., Hansen, T. K., Jensen, M. H., Poulsen, P. L., & Vestergaard, P. (2025). From Stability to Variability: Classification of Healthy Individuals, Prediabetes, and Type 2 Diabetes Using Glycemic Variability Indices from Continuous Glucose Monitoring Data. *Diabetes Technology & Therapeutics*, 27(1), 34–44. <https://doi.org/10.1089/dia.2024.0226>

Çorbacıoğlu, Ş. K., & Aksel, G. (2023). Receiver operating characteristic curve analysis in diagnostic accuracy studies. *Turkish Journal of Emergency Medicine*, 23(4), 195–198. https://doi.org/10.4103/tjem.tjem_182_23

Cui, E. H., Goldfine, A. B., Quinlan, M., James, D. A., & Sverdlov, O. (2023). Investigating the value of glucodensity analysis of continuous glucose monitoring data in type 1 diabetes: an exploratory analysis. *Frontiers in Clinical Diabetes and Healthcare*, 4. <https://doi.org/10.3389/fcdhc.2023.1244613>

Czupryniak, L., Barkai, L., Bolgarska, S., Bronisz, A., Broz, J., Cypryk, K., Honka, M., Janez, A., Krnic, M., Lalic, N., Martinka, E., Rahelic, D., Roman, G., Tankova, T., Várkonyi, T., Wolnik, B., & Zherdova, N. (2014). Self-Monitoring of Blood Glucose in Diabetes: From Evidence to Clinical Reality in Central and Eastern Europe—Recommendations from the International Central-Eastern European Expert Group. *Diabetes Technology & Therapeutics*, 16(7), 460–475. <https://doi.org/10.1089/dia.2013.0302>

Crăciun, C. I., Craciun, A.-E., Bocsan, I. C., Rusu, A., Bala, C., Coadă, C., Roman, G., & Buzoianu, A. (2018). The C-peptide correlations and effect on the glycemic variability parameters in patients with type 2 diabetes. *Human and Veterinary Medicine*, 10, 111–116.

Díaz-Soto, G., Pérez-López, P., Fernández-Velasco, P., Nieto de la Marca, M. de la O., Delgado, E., del Amo, S., de Luis, D., & Bahillo-Currieses, P. (2024). Glycemia Risk Index Assessment in a Pediatric and Adult Patient Cohort with Type 1 Diabetes Mellitus. *Journal of Diabetes Science and Technology*, 18(5), 1063–1069. <https://doi.org/10.1177/19322968231154561>

Dunn, T. C., Ajjan, R. A., Bergenstal, R. M., & Xu, Y. (2024). Is It Time to Move Beyond TIR to TITR? Real-World Data from Over 20,000 Users of Continuous Glucose Monitoring in Patients with Type 1 and Type 2 Diabetes. *Diabetes Technology & Therapeutics*, 26(3), 203–210. <https://doi.org/10.1089/dia.2023.0565>

El Hallani, A., Chakhtouna, A., & Adib, A. (2025). Alzheimer's Detection via Cepstral Domain Feature Fusion and Mann-Whitney U-Test Statistical Filtering. *2025 5th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, 1–7. <https://doi.org/10.1109/IRASET64571.2025.11007957>

El-Laboudi, A. H., Godsland, I. F., Johnston, D. G., & Oliver, N. S. (2016). Measures of Glycemic Variability in Type 1 Diabetes and the Effect of Real-Time Continuous

Glucose Monitoring. *Diabetes Technology & Therapeutics*, 18(12), 806–812. <https://doi.org/10.1089/dia.2016.0146>

ElSayed, N. A., McCoy, R. G., Aleppo, G., Balapattabi, K., Beverly, E. A., Briggs Early, K., Bruemmer, D., Ebekozien, O., Echouffo-Tcheugui, J. B., Ekhlaspour, L., Gaglia, J. L., Garg, R., Khunti, K., Lal, R., Lingvay, I., Matfin, G., Pandya, N., Pekas, E. J., Pilla, S. J., Polsky, S., Segal, A. R., Seley, J. J., Stanton, R. C., & Bannuru, R. R. (2025). 2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes—2025. *Diabetes Care*, 48(Supplement_1), S27–S49. <https://doi.org/10.2337/dc25-S002>

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, 48(4). <https://doi.org/10.18637/jss.v048.i04>

Fabris, C., Patek, S. D., & Breton, M. D. (2016). Are Risk Indices Derived from CGM Interchangeable With SMBG-Based Indices? *Journal of Diabetes Science and Technology*, 10(1), 50–59. <https://doi.org/10.1177/1932296815599177>

Fernandes, N. J., Nguyen, N., Chun, E., Punjabi, N. M., & Gaynanova, I. (2022). Open-Source Algorithm to Calculate Mean Amplitude of Glycemic Excursions Using Short and Long Moving Averages. *Journal of Diabetes Science and Technology*, 16(2), 576–577. <https://doi.org/10.1177/19322968211061165>

Flatt, A. J. S., Greenbaum, C. J., Shaw, J. A. M., & Rickels, M. R. (2021). Pancreatic islet reserve in type 1 diabetes. *Annals of the New York Academy of Sciences*, 1495(1), 40–54. <https://doi.org/10.1111/nyas.14572>

Frigge, M., Hoaglin, D. C., & Iglewicz, B. (1989). Some Implementations of the Boxplot. *The American Statistician*, 43(1), 50–54. <https://doi.org/10.1080/00031305.1989.10475612>

Gao, Z., Fu, G., Ouyang, C., Tsutsui, S., Liu, X., Yang, J., Gessner, C., Foote, B., Wild, D., Ding, Y., & Yu, Q. (2019). edge2vec: Representation learning using edge semantics for biomedical knowledge discovery. *BMC Bioinformatics*, 20(1), 306. <https://doi.org/10.1186/s12859-019-2914-2>

Gómez, A. M., Henao, D. C., Imitola Madero, A., Taboada, L. B., Cruz, V., Robledo Gómez, M. A., Rondón, M., Muñoz-Velandia, O., García-Jaramillo, M., & León Vargas,

F. M. (2019). Defining High Glycemic Variability in Type 1 Diabetes: Comparison of Multiple Indexes to Identify Patients at Risk of Hypoglycemia. *Diabetes Technology & Therapeutics*, 21(8), 430–439. <https://doi.org/10.1089/dia.2019.0075>

Gong, B., Yang, W., Xing, Y., Lai, Y., & Shan, Z. (2025). Global, regional, and national burden of type 1 diabetes in adolescents and young adults. *Pediatric Research*, 97(2), 568–576. <https://doi.org/10.1038/s41390-024-03107-5>

Gregory, J. M., Moore, D. J., & Simmons, J. H. (2013). Type 1 Diabetes Mellitus. *Pediatrics in Review*, 34(5), 203–215. <https://doi.org/10.1542/pir.34-5-203>

Guerrero, M. C., Parada, J. S., & Espitia, H. E. (2021). EEG signal analysis using classification techniques: Logistic regression, artificial neural networks, support vector machines, and convolutional neural networks. *Heliyon*, 7(6), e07258. <https://doi.org/10.1016/j.heliyon.2021.e07258>

Gutiérrez-Tobal, G. C., Gomez-Pilar, J., Kheirandish-Gozal, L., Martín-Montero, A., Poza, J., Álvarez, D., del Campo, F., Gozal, D., & Hornero, R. (2021). Pediatric Sleep Apnea: The Overnight Electroencephalogram as a Phenotypic Biomarker. *Frontiers in Neuroscience*, 15. <https://doi.org/10.3389/fnins.2021.644697>

Haji-Maghsoudi, S., Haghdoost, A., Rastegari, A., & Baneshi, M. R. (2013). Influence of Pattern of Missing Data on Performance of Imputation Methods: An Example Using National Data on Drug Injection in Prisons. *International Journal of Health Policy and Management*, 1(1), 69–77. <https://doi.org/10.15171/ijhpm.2013.11>

Hintze, J. L., & Nelson, R. D. (1998). Violin Plots: A Box Plot-Density Trace Synergism. *The American Statistician*, 52(2), 181–184. <https://doi.org/10.1080/00031305.1998.10480559>

Jessica, E. O., Hamada, M., Yusuf, S. I., & Hassan, M. (2021). The Role of Linear Discriminant Analysis for Accurate Prediction of Breast Cancer. *2021 IEEE 14th International Symposium on Embedded Multicore/Many-Core Systems-on-Chip (MCSoc)*, 340–344. <https://doi.org/10.1109/MCSoc51149.2021.00057>

Jimeno, N., Gomez-Pilar, J., Poza, J., Hornero, R., Vogeley, K., Meisenzahl, E., Haidl, T., Rosen, M., Klosterkötter, J., & Schultze-Lutter, F. (2020). Main Symptomatic Treatment Targets in Suspected and Early Psychosis: New Insights from Network Analysis. *Schizophrenia Bulletin*, 46(4), 884–895. <https://doi.org/10.1093/schbul/sbz140>

Jones, L. V, Ray, A., Moy, F. M., & Buckley, B. S. (2019). Techniques of monitoring blood glucose during pregnancy for women with pre-existing diabetes. *Cochrane Database of Systematic Reviews*, 2019(5). <https://doi.org/10.1002/14651858.CD009613.pub4>

Kahn, C. R., Weir, G. C., King, G. L., Moses, A. C., Smith, R. J., & Jacobson, A. M. (2004). *Joslin's Diabetes Mellitus (14th ed., Vol. 23)*. Lippincott Williams & Wilkins.

Kamusheva, M., Tachkov, K., Dimitrova, M., Mitkova, Z., García-Sáez, G., Hernando, M. E., Goettsch, W., & Petrova, G. (2021). A Systematic Review of Collective Evidences Investigating the Effect of Diabetes Monitoring Systems and Their Application in Health Care. *Frontiers in Endocrinology*, 12. <https://doi.org/10.3389/fendo.2021.636959>

Karakuş, K. E., Sakarya, S., Yeşiltepe Mutlu, G., Berkkan, M., Muradoğlu, S., Can, E., Gökçe, T., Eviz, E., & Hatun, Ş. (2021). Benefits and Drawbacks of Continuous Glucose Monitoring (CGM) Use in Young Children With Type 1 Diabetes: A Qualitative Study From a Country Where the CGM Is Not Reimbursed. *Journal of Patient Experience*, 8. <https://doi.org/10.1177/23743735211056523>

Katsarou, A., Gudbjörnsdottir, S., Rawshani, A., Dabelea, D., Bonifacio, E., Anderson, B. J., Jacobsen, L. M., Schatz, D. A., & Lernmark, Å. (2017). Type 1 diabetes mellitus. *Nature Reviews Disease Primers*, 3(1), 17016. <https://doi.org/10.1038/nrdp.2017.16>

Khurana, V., Nigam, A., & De, A. (2025). Exploring glycemic variability and time in range: Emerging indicators of at-risk pregnancy in type 2 diabetes mellitus. *International Journal of Gynecology & Obstetrics*, 169(2), 728–734. <https://doi.org/10.1002/ijgo.16140>

Klonoff, D. C., Wang, J., Rodbard, D., Kohn, M. A., Li, C., Liepmann, D., Kerr, D., Ahn, D., Peters, A. L., Umpierrez, G. E., Seley, J. J., Xu, N. Y., Nguyen, K. T., Simonson,

G., Agus, M. S. D., Al-Sofiani, M. E., Armaiz-Pena, G., Bailey, T. S., Basu, A., Battelino, T., Bekel, S. Y., Benhamou, P.-I., Bequette, B. W., Blevins, T., Breton, M. D., Castle, J. R., Chase, J. G., Chen, K. Y., Choudhary, P., Clements, M. A., Close, K. L., Cook, C. B., Danne, T., Doyle 3rd, F. J., Drincic, A., Dungan, K. M., Edelman, S. V., Ejlskjær, N., Espinoza, J. C., Fleming, G. A., Forlenza, G. P., Freckmann, G., Galindo, R. J., Gómez, A. M., Gutow, H. A., Heinemann, L., Hirsch, I. B., Hoang, T. D., Hovorka, R., Jendle, J. H., Ji, L., Joshi, S. R., Joubert, M., Koliwad, S. K., Lal, R. A., Lansang, M. C., Andy Lee, W.-A., Leelarathna, L., Leiter, L. A., Lind, M., Litchman, M. L., Mader, J. K., Mahoney, K. M., Mankovsky, B., Masharani, U., Mathioudakis, N. N., Mayorov, A., Messler, J., Miller, J. D., Mohan, V., Nichols, J. H., Nørgaard, K., O’Neal, D. N., Pasquel, F. J., Philis-Tsimikas, A., Pieber, T., Phillip, M., Polonsky, W. H., Pop-Busui, R., Rayman, G., Rhee, E.-J., Russell, S. J., Shah, V. N., Sherr, J. L., Sode, K., Spanakis, E. K., Wake, D. J., Waki, K., Wallia, A., Weinberg, M. E., Wolpert, H., Wright, E. E., Zilbermint, M., & Kovatchev, B. (2023). A Glycemia Risk Index (GRI) of Hypoglycemia and Hyperglycemia for Continuous Glucose Monitoring Validated by Clinician Ratings. *Journal of Diabetes Science and Technology*, 17(5), 1226–1242. <https://doi.org/10.1177/19322968221085273>

Köhlmoos, A., & Dittmar, M. (2025). Glycemic Variability and Control by CGM in Healthy Older and Young Adults and Their Relationship with Diet. *Journal of the Endocrine Society*, 9(7). <https://doi.org/10.1210/jendso/bvaf081>

Kohnert, K.-D., Heinke, P., Fritzsche, G., Vogt, L., Augstein, P., & Salzsieder, E. (2013). Evaluation of the Mean Absolute Glucose Change as a Measure of Glycemic Variability Using Continuous Glucose Monitoring Data. *Diabetes Technology & Therapeutics*, 15(6), 448–454. <https://doi.org/10.1089/dia.2012.0303>

Lazar, S., Ionita, I., Reurean-Pintilei, D., & Timar, B. (2023). How to Measure Glycemic Variability? A Literature Review. *Medicina*, 60(1), 61. <https://doi.org/10.3390/medicina60010061>

Leighton, E., Sainsbury, C. A., & Jones, G. C. (2017). A Practical Review of C-Peptide Testing in Diabetes. *Diabetes Therapy*, 8(3), 475–487. <https://doi.org/10.1007/s13300-017-0265-4>

Leksic, G., Baretić, M., Gudelj, L., Radic, M., Milicic, I., Ivanišević, M., & Jurisic-Erzen, D. (2022). Glycemic Variability in Type 1 Diabetes Mellitus Pregnancies—Novel Parameters in Predicting Large-for-Gestational-Age Neonates: A Prospective Cohort Study. *Biomedicines*, *10*(9), 2175. <https://doi.org/10.3390/biomedicines10092175>

Liu, J., Ren, Z.-H., Qiang, H., Wu, J., Shen, M., Zhang, L., & Lyu, J. (2020). Trends in the incidence of diabetes mellitus: results from the Global Burden of Disease Study 2017 and implications for diabetes mellitus prevention. *BMC Public Health*, *20*(1), 1415. <https://doi.org/10.1186/s12889-020-09502-x>

Marks, B. E., Williams, K. M., Sherwood, J. S., & Putman, M. S. (2022). Practical aspects of diabetes technology use: Continuous glucose monitors, insulin pumps, and automated insulin delivery systems. *Journal of Clinical & Translational Endocrinology*, *27*, 100282. <https://doi.org/10.1016/j.jcte.2021.100282>

Masierek, M., Nabrdalik, K., Janota, O., Kwiendacz, H., Macherski, M., & Gumprecht, J. (2022). The Review of Insulin Pens—Past, Present, and Look to the Future. *Frontiers in Endocrinology*, *13*. <https://doi.org/10.3389/fendo.2022.827484>

Matabuena, M., Ghosal, R., Aguilar, J., Wagner, R., Merino, C., Castro, J., Zipunnikov, V., Onnela, J.-P., & Gude, F. (2024). Glucodensity Functional Profiles Outperform Traditional Continuous Glucose Monitoring Metrics. <https://doi.org/10.48550/arXiv.2410.00912>

Mo, Y., Lu, J., & Zhou, J. (2024). Glycemic variability: Measurement, target, impact on complications of diabetes and does it really matter? *Journal of Diabetes Investigation*, *15*(1), 5–14. <https://doi.org/10.1111/jdi.14112>

Ohara, M., Hiromura, M., Nagaike, H., Kohata, Y., Fujikawa, T., Goto, S., Sato, N., Kushima, H., Terasaki, M., Yamamoto, T., Mori, Y., Hayashi, T., Fukui, T., Yamagishi, S., & Hirano, T. (2019). Relationship between glucose variability evaluated by continuous glucose monitoring and clinical factors, including glucagon-stimulated insulin secretion in patients with type 2 diabetes. *Diabetes Research and Clinical Practice*, *158*, 107904. <https://doi.org/10.1016/j.diabres.2019.107904>

Paschou, S. A., Papadopoulou-Marketou, N., Chrousos, G. P., & Kanaka-Gantenbein, C. (2018). On type 1 diabetes mellitus pathogenesis. *Endocrine Connections*, 7(1), R38–R46. <https://doi.org/10.1530/EC-17-0347>

Peng, W., Yuan, J., Chiavaroli, V., Dong, G., Huang, K., Wu, W., Ullah, R., Jin, B., Lin, H., Derraik, J. G. B., & Fu, J. (2021). 10-Year Incidence of Diabetic Ketoacidosis at Type 1 Diabetes Diagnosis in Children Aged Less Than 16 Years from a Large Regional Center (Hangzhou, China). *Frontiers in Endocrinology*, 12. <https://doi.org/10.3389/fendo.2021.653519>

Pérez-López, P., Fernández-Velasco, P., Bahillo-Curienes, P., de Luis, D., & Díaz-Soto, G. (2023). Impact of glucose variability on the assessment of the glycemia risk index (GRI) and classic glycemic metrics. *Endocrine*, 82(3), 560–568. <https://doi.org/10.1007/s12020-023-03511-7>

Peyser, T. A., Balo, A. K., Buckingham, B. A., Hirsch, I. B., & Garcia, A. (2018). Glycemic Variability Percentage: A Novel Method for Assessing Glycemic Variability from Continuous Glucose Monitor Data. *Diabetes Technology & Therapeutics*, 20(1), 6–16. <https://doi.org/10.1089/dia.2017.0187>

Powers, A. C. (2021). Type 1 diabetes mellitus: much progress, many opportunities. *Journal of Clinical Investigation*, 131(8), e142242. <https://doi.org/10.1172/JCI142242>

Qiu, J., Yin, W., Wang, R., Luo, S., & Zhou, Z. (2024). Fulminant type 1 diabetes: Focusing on triggering factors. *Diabetes/Metabolism Research and Reviews*, 40(2), e3731. <https://doi.org/10.1002/dmrr.3731>

Quattrin, T., Mastrandrea, L. D., & Walker, L. S. K. (2023). Type 1 diabetes. *The Lancet*, 401(10394), 2149–2162. [https://doi.org/10.1016/S0140-6736\(23\)00223-4](https://doi.org/10.1016/S0140-6736(23)00223-4)

Rahman, M. S., Hossain, K. S., Das, S., Kundu, S., Adegoke, E. O., Rahman, Md. A., Hannan, Md. A., Uddin, M. J., & Pang, M.-G. (2021). Role of Insulin in Health and Disease: An Update. *International Journal of Molecular Sciences*, 22(12), 6403. <https://doi.org/10.3390/ijms22126403>

Ramos-Fernández, G., Boyer, D., Aureli, F., & Vick, L. G. (2009). Association networks in spider monkeys (*Ateles geoffroyi*). *Behavioral Ecology and Sociobiology*, 63(7), 999–1013. <https://doi.org/10.1007/s00265-009-0719-4>

Ranea, J. A. G., Perkins, J., Chagoyen, M., Díaz-Santiago, E., & Pazos, F. (2022). Network-Based Methods for Approaching Human Pathologies from a Phenotypic Point of View. *Genes*, 13(6), 1081. <https://doi.org/10.3390/genes13061081>

Rickels, M. R., Evans-Molina, C., Bahnson, H. T., Ylescupidez, A., Nadeau, K. J., Hao, W., Clements, M. A., Sherr, J. L., Pratley, R. E., Hannon, T. S., Shah, V. N., Miller, K. M., & Greenbaum, C. J. (2020). High residual C-peptide likely contributes to glycemic control in type 1 diabetes. *Journal of Clinical Investigation*, 130(4), 1850–1862. <https://doi.org/10.1172/JCI134057>

Roche, E. F., Menon, A., Gill, D., & Hoey, H. (2005). Clinical presentation of type 1 diabetes. *Pediatric Diabetes*, 6(2), 75–78. <https://doi.org/10.1111/j.1399-543X.2005.00110.x>

Rodbard, D. (2009). Interpretation of Continuous Glucose Monitoring Data: Glycemic Variability and Quality of Glycemic Control. *Diabetes Technology & Therapeutics*, 11(S1), S55–S67. <https://doi.org/10.1089/dia.2008.0132>

Rodbard, D. (2018). Glucose Variability: A Review of Clinical Applications and Research Developments. *Diabetes Technology & Therapeutics*, 20(S2), S25–S215. <https://doi.org/10.1089/dia.2018.0092>

Rodbard, D. (2021). The Ambulatory Glucose Profile: Opportunities for Enhancement. *Diabetes Technology & Therapeutics*, 23(5), 332–341. <https://doi.org/10.1089/dia.2020.0524>

Roze, S., Isitt, J. J., Smith-Palmer, J., Lynch, P., Klinkenbijn, B., Zammit, G., & Benhamou, P.-Y. (2021). Long-Term Cost-Effectiveness the Dexcom G6 Real-Time Continuous Glucose Monitoring System Compared with Self-Monitoring of Blood Glucose in People with Type 1 Diabetes in France. *Diabetes Therapy*, 12(1), 235–246. <https://doi.org/10.1007/s13300-020-00959-y>

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage*, 52(3), 1059–1069. <https://doi.org/10.1016/j.neuroimage.2009.10.003>

Ruiz-Grao, M. C., Díez-Fernández, A., Mesas, A. E., Martínez-Vizcaíno, V., Sequí-Domínguez, I., Sebastián-Valles, F., & Garrido-Miguel, M. (2024). Trends in the Incidence of Type 1 Diabetes in European Children and Adolescents from 1994 to 2022: A Systematic Review and Meta-Analysis. *Pediatric Diabetes*, 2024, 1–13. <https://doi.org/10.1155/2024/2338922>

Salankar, N., Qaisar, S. M., Pławiak, P., Tadeusiewicz, R., & Hammad, M. (2022). EEG based alcoholism detection by oscillatory modes decomposition second order difference plots and machine learning. *Biocybernetics and Biomedical Engineering*, 42(1), 173–186. <https://doi.org/10.1016/j.bbe.2021.12.009>

Selvin, E. (2024). The Glucose Management Indicator: Time to Change Course? *Diabetes Care*, 47(6), 906–914. <https://doi.org/10.2337/dci23-0086>

Shafiee, S., Lied, L. M., Burud, I., Dieseth, J. A., Alsheikh, M., & Lillemo, M. (2021). Sequential forward selection and support vector regression in comparison to LASSO regression for spring wheat yield prediction based on UAV imagery. *Computers and Electronics in Agriculture*, 183, 106036. <https://doi.org/10.1016/j.compag.2021.106036>

Sparks, J. R., Sarzynski, M. A., Davis, J. M., Grandjean, P. W., & Wang, X. (2022). Cross-Sectional and Individual Relationships between Physical Activity and Glycemic Variability. *Translational Journal of the American College of Sports Medicine*, 7(4), 1–12. <https://doi.org/10.1249/TJX.0000000000000207>

Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*, 51(D1), D638–D646. <https://doi.org/10.1093/nar/gkac1000>

Tokutsu, A., Okada, Y., Mita, T., Torimoto, K., Wakasugi, S., Katakami, N., Yoshii, H., Uryu, K., Nishida, K., Arao, T., Tanaka, Y., Goshō, M., Shimomura, I., & Watada,

H. (2022). Relationship between blood glucose variability in ambulatory glucose profile and standardized continuous glucose monitoring metrics: Subanalysis of a prospective cohort study. *Diabetes, Obesity and Metabolism*, 24(1), 82–93. <https://doi.org/10.1111/dom.14550>

Tomic, D., Harding, J. L., Jenkins, A. J., Shaw, J. E., & Magliano, D. J. (2025). The epidemiology of type 1 diabetes mellitus in older adults. *Nature Reviews Endocrinology*, 21(2), 92–104. <https://doi.org/10.1038/s41574-024-01046-z>

Untung Halajur, & Irfan Setiawan. (2021). The Relationship of Knowledge with the Self-Management of Diabetes Clients Diabetes Mellitus. *International Journal of Science, Technology & Management*, 2(1), 363–367. <https://doi.org/10.46729/ijstm.v2i1.108>

Venugopal, S. K., Mowery, M. L., & Jialal, I. (2023). Biochemistry, C peptide. *StatPearls* [Internet].

Waite, M., Martin, C., Franklin, R., Duce, D., & Harrison, R. (2018). Human Factors and Data Logging Processes with the Use of Advanced Technology for Adults with Type 1 Diabetes: Systematic Integrative Review. *JMIR Human Factors*, 5(1), e11. <https://doi.org/10.2196/humanfactors.9049>

Wang, S., Tan, Z., Wu, T., Shen, Q., Huang, P., Wang, L., Liu, W., Song, H., Lin, M., Shi, X., & Li, X. (2022). Largest Amplitude of Glycemic Excursion Calculating from Self-Monitoring Blood Glucose Predicted the Episodes of Nocturnal Asymptomatic Hypoglycemia Detecting by Continuous Glucose Monitoring in Outpatients with Type 2 Diabetes. *Frontiers in Endocrinology*, 13. <https://doi.org/10.3389/fendo.2022.858912>

WHO (2019). Classification of diabetes mellitus. *World Health Organization*.

Wong, M. H. (2022). grpandplot: An open-source MATLAB tool for drawing box plot and violin plot with automatic multi-way data grouping (1.0.0). *Zenodo*. <https://doi.org/10.5281/zenodo.7295877>

Wong, T.-T. (2015). Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48(9), 2839–2846. <https://doi.org/10.1016/j.patcog.2015.03.009>

Xing, D., Kollman, C., Beck, R. W., Tamborlane, W. V., Laffel, L., Buckingham, B. A., Wilson, D. M., Weinzimer, S., Fiallo-Scharer, R., & Ruedy for the Juvenile Diabetes Res, K. J. (2011). Optimal Sampling Intervals to Assess Long-Term Glycemic Control Using Continuous Glucose Monitoring. *Diabetes Technology & Therapeutics*, 13(3), 351–358. <https://doi.org/10.1089/dia.2010.0156>

Ying, C., Wang, G., & Li, H. (2021). Design of Feature Selection Algorithm Based on Improved FCBF. *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 323–327. <https://doi.org/10.1109/ICSP51882.2021.9408729>

Zhang, J., Yang, J., Liu, L., Li, L., Cui, J., Wu, S., & Tang, K. (2021). Significant abnormal glycemic variability increased the risk for arrhythmias in elderly type 2 diabetic patients. *BMC Endocrine Disorders*, 21(1), 83. <https://doi.org/10.1186/s12902-021-00753-2>

Glosario de siglas y acrónimos

AAPC	Porcentaje promedio de cambio anual, <i>average anual percent change</i>
ADA	Asociación Americana de Diabetes, <i>American Diabetes Association</i>
ADRR	Promedio de ratios de riesgo diarios, <i>average daily risk ratio</i>
AGP	Perfil ambulatorio de glucosa, <i>ambulatory glucose profile</i>
AIRE	Regulador autoinmune, <i>autoimmune regulator</i>
APS-I	Síndrome autoinmune poliglandular tipo 1, <i>Autoimmune Polyendocrine Syndrome type I</i>
APS-II	Síndrome autoinmune poliglandular tipo 2, <i>Autoimmune Polyendocrine Syndrome type II</i> , APS-II
AUC	Área bajo la curva ROC
BCR	Receptor de los linfocitos B, <i>B cell receptor</i>
CAD	Cetoacidosis diabética
CGM	Monitorización continua de glucosa, <i>continuous glucose monitoring</i>
CHyper	Carga de hiperglucemia
CHypo	Carga de hipoglucemia
CONGA	Red de solapamiento continua de acción glucémica, <i>continuous overlapping net glycemic action</i>
Cpep	Nivel basal de péptido C
CSII	Infusión continua de insulina subcutánea, <i>continuous subcutaneous infusion of insulin</i>
CTLA-4	Antígeno-4 de los linfocitos T citotóxicos, <i>cytotoxic T lymphocyte antigen-4</i>
CTM	Medida de tendencia central, <i>central tendency measure</i>
CV	Coeficiente de variación
DC	Células dendríticas, <i>dendritic cells</i>
DDT	Dosis diaria total

DM	Diabetes mellitus
DM1	Diabetes mellitus tipo 1
DM1A	Diabetes mellitus tipo 1A
DM1B	Diabetes mellitus tipo 1B
DM2	Diabetes mellitus tipo 2
DMG	Diabetes mellitus gestacional
DMPT	Diabetes mellitus postrasplante
ECA	Enzima convertidora de angiotensina
ERBB3	Receptor Erb-B2 de tirosina quinasa 3, <i>Erb-B2 receptor tyrosine kinase 3</i>
FCBF	<i>Fast Correlation-Based Filter</i>
FDA	Administración de Alimentos y Medicamentos, <i>Food and Drug Administration</i>
GAD	Ácido glutámico descarboxilasa, <i>glutamic acid decarboxylase</i>
GBD	Carga Global de Enfermedades, <i>Global Burden of Disease</i>
GLP-1	Péptido similar al glucagón tipo 1, <i>glucagon-like peptide-1</i>
GMI	Indicador de manejo de la glucosa, <i>glucose management indicator</i>
GRADE	<i>Glycemic risk assessment diabetic equation</i>
GRI	Índice de riesgo glucémico, <i>glycemia risk index</i>
GVP	Porcentaje de variabilidad glucémica, <i>glycemic variability percentage</i>
H	Hombres
HbA1c	Hemoglobina glicosilada
HBGI	Índice de alta glucosa en sangre, <i>high blood glucose index</i>
HCL	Sistemas híbridos de asa cerrada, <i>hybrid closed loop</i>
HCUV	Hospital Clínico Universitario de Valladolid
HIP14	Proteína 14 de interacción Huntingtin, <i>Huntingtin-interacting protein 14</i>
HLA	Antígeno leucocitario humano, <i>human leukocyte antigen</i>

IA-2	Antígeno de los islotes 2, <i>islet antigen 2</i>
I:C	Ratio insulina/carbohidratos
IFIH1	Dominio 1 de helicasa inducida por interferón, <i>interferon-induced helicase domain 1</i>
IGC	Índice de control glucémico, <i>index of glycemic control</i>
IL-1	Interleucina 1, <i>interleukin 1</i>
IMC	Índice de masa corporal
INF-γ	Interferón γ , <i>interferon γ</i>
IQR	Rango intercuartil, <i>interquartile range</i>
J	Índice J, <i>J-index</i>
LADA	Diabetes latente autoinmune en adultos, <i>latent autoimmune diabetes in adults</i>
LAGE	Amplitud máxima de excursiones de glucosa, <i>large amplitude of glucose excursions</i>
LBGI	Índice de baja glucosa en sangre, <i>low blood glucose index</i>
LDA	Análisis discriminante lineal, <i>Linear Discriminant Analysis</i>
LI	<i>Liability index</i>
LR	Regresión logística, <i>Logistic Regression</i>
LZC	Complejidad de Lempel-Ziv, <i>Lempel-Ziv Complexity</i>
m	Media
M	Mujeres
MAG	Cambio de glucosa absoluta media por unidad de tiempo, <i>mean absolute glucose change per unit of time</i>
MAGE	Amplitud media de excursiones de glucosa, <i>mean amplitude of glucose excursions</i>
MaxGly	Máximo de glucemia, <i>maximum of glycemia</i>
MDI	Inyección diaria múltiple de insulina, <i>multiple daily injection</i>
MG	Glucosa media
MHC	Complejo Mayor de Histocompatibilidad, <i>Major Hystocompatibility Complex</i>

MICE	Imputación múltiple por cadenas de ecuaciones, <i>multiple imputation by chained equations</i>
MIME	<i>Mean indices of meal excursions</i>
MinGly	Mínimo de glucemia, <i>minimum of glycemia</i>
MODD	Media de diferencias diarias, <i>mean of daily differences</i>
MODY	<i>Maturity-Onset Diabetes of the Young</i>
NHC	Número de historia clínica
NIH	Instituto Nacional de Salud, <i>National Institute of Health</i>
OMS	Organización Mundial de la Salud
PTGO	Prueba de tolerancia a la glucosa oral
PTPN22	Proteína tirosina fosfatasa no receptora tipo 22, <i>protein tyrosine phosphatase non-receptor type 22</i>
QDA	Análisis discriminante cuadrático, <i>Quadratic Discriminant Analysis</i>
ROC	<i>Receiver operating characteristic</i>
SACYL	Sanidad de Castilla y León
SD	Desviación estándar, <i>standard deviation</i>
SFS	<i>Sequential Forward Selection</i>
SGLT-2	Cotransportador de sodio-glucosa tipo 2, <i>Sodium glucose cotransporter-2</i>
SMBG	Monitorización de glucosa en sangre mediante punción digital, <i>self-monitoring blood glucose</i>
SNPs	Polimorfismos de nucleótido único, <i>single nucleotide polymorphisms</i>
SU	<i>Symmetric uncertainty</i>
TAR	Tiempo por encima de rango, <i>time above range</i>
TAR1	Tiempo por encima de rango nivel 1, <i>time above range level 1</i>
TAR2	Tiempo por encima de rango nivel 2, <i>time above range level 2</i>
TBR	Tiempo por debajo de rango, <i>time below range</i>
TBR1	Tiempo por debajo de rango nivel 1, <i>time below range level 1</i>

TBR2	Tiempo por debajo de rango nivel 2, <i>time below range level 2</i>
TCR	Receptor de los linfocitos T, <i>T cell receptor</i>
TIR	Tiempo en rango, <i>time in range</i>
TITR	Tiempo en rango estrecho, <i>time in tight range</i>
TNF-α	Factor de necrosis tumoral α , <i>tumor necrosis factor α</i>
TOTR	Tiempo fuera de rango estrecho, <i>time out of tight range</i>
Var1	Variación de péptido C a los 6-18 meses
Var2	Variación de péptido C a los 18-30 meses
Var3	Variación de péptido C a los 30-36 meses
VarC	Variación global de péptido C
VarMonth	Variación promedio mensual de péptido C
VEGF	Factor de crecimiento endotelial vascular, <i>vascular endothelial growth factor</i>
VG	Variabilidad glucémica
VNTRs	Repeticiones en tándem variables, <i>variable number of tandem repeats</i>
XPID	Síndrome de inmunodesregulación, poliendocrinopatía y enteropatía ligada al cromosoma X, <i>X-linked syndrome of polyendocrinopathy, immune dysfunction and diarrhea</i>
ZnT8	Transportador de zinc 8, <i>zinc transporter 8</i>