Universidad de Concepción

Facultad de Farmacia

Doctorado en Ciencias y Tecnología Analítica

DESARROLLO Y EVALUACIÓN DE MÉTODOS MULTIVARIADOS BASADOS EN DATOS GINECO-OBSTÉTRICOS Y ESPECTRALES DE SUERO PARA LA PREDICCIÓN DE DIABETES GESTACIONAL EN EL PRIMER TRIMESTRE DEL EMBARAZO

Tesis presentada a la Facultad de Farmacia de la Universidad de Concepción para optar al grado académico de Doctora en Ciencias y Tecnología Analítica

POR: DANIELA FRANCISCA MENNICKENT BARROS

Profesor Guía: Dr. Enrique Alberto Guzmán Gutiérrez

Profesor Co-Guía: Dr. Juan Andrés Araya Quintana

Julio, 2023

Concepción, Chile

Porque de Él, por Él y para Él son todas las cosas.

A Él sea la gloria para siempre. Amén.

AGRADECIMIENTOS

A la Agencia Nacional de Investigación y Desarrollo, por financiar este trabajo a través de la Beca de Doctorado Nacional 21190736 y Beneficios Complementarios de Gastos Operacionales y Pasantía Doctoral en el Extranjero; y los proyectos FONDECYT 11170710, FONDECYT 11181153 y FOVI 210057.

A la Universidad de Concepción y el Ministerio de Educación, por su apoyo económico mediante los Fondos “Asistencia a Cursos de Especialización de Corta Duración en Sociedades Científicas o Instituciones Académicas en el Extranjero” y “Pasantía Internacional de Doctorado” del Proyecto UCO 1866.

A Enrique y Juan, por su confianza y apoyo. Por animarme a salir de mi zona de confort. Por todas las oportunidades que han compartido conmigo para potenciar mi carrera académica.

A Federico, por su amabilidad y generosidad durante mi pasantía en la Sapienza Università di Roma.

A papá, mamá, Rocío y Camila, por su amor y respaldo constantes.

A Dios, por su provisión y guía. Él ha sido fiel y, sin duda, lo seguirá siendo.

TABLA DE CONTENIDO

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

RESUMEN

Introducción: La diabetes gestacional (DG) es un estado de hiperglicemia que se diagnostica en el segundo o tercer trimestre del embarazo, sin embargo para ese entonces el fenotipo fetal ya está alterado. En literatura se reportan estrategias para predecir DG en etapas más tempranas del embarazo; no obstante presentan un bajo poder predictivo o requieren largos tiempos de análisis. **Objetivo general:** Evaluar métodos basados en el análisis multivariado de datos gineco-obstétricos y/o espectrales de suero para la predicción de DG en el primer trimestre del embarazo. **Estrategia analítica:** Se desarrollaron modelos predictivos para DG a partir de técnicas de análisis multivariado *single-block* y *multi-block*, y (1) data gineco-obstétrica (53 parámetros clínicos y 10 bioquímicos); (2) data espectral de suero del infrarrojo cercano (NIR); o (3) ambas (28 parámetros clínicos, y NIR). **Resultados:** Los mejores modelos de (1), (2) y (3) lograron un área bajo la curva ROC de 0,867, 0,577 y 0,522, respectivamente. Solo el primero aumentó el poder predictivo respecto al promedio de literatura. Los mejores modelos de (2) y (3) lograron un tiempo de análisis de 32 minutos. Ambos disminuyeron el tiempo de análisis respecto al promedio de literatura. **Conclusión:** Las estrategias basadas en el análisis multivariado de datos gineco-obstétricos y/o espectrales de suero del primer trimestre de gestación mejoran la predicción de DG en etapas tempranas del embarazo, ya sea incrementando el poder predictivo, o reduciendo los tiempos de análisis.

ABSTRACT

Introduction: Gestational diabetes mellitus (GDM) is a hyperglycemia state that is diagnosed in the second or third trimester of pregnancy, however by that time the fetal phenotype is already altered. Strategies to predict GDM in earlier stages of pregnancy are reported in literature; nevertheless they present low predictive power or require long times of analysis. **General objective:** To evaluate methods based on the multivariate analysis of gyneco-obstetrical and/or serum spectral data for the prediction of GDM in the first trimester of pregnancy. **Analytical strategy:** GDM predictive models were developed using single- and multi-block multivariate analysis techniques, and (1) gyneco-obstetrical data (53 clinical and 10 biochemical parameters); (2) near-infrared (NIR) serum spectral data; or (3) both (28 clinical parameters, and NIR). **Results:** The best models of (1), (2) and (3) achieved an area under the ROC curve of 0.867, 0.577 and 0.522, respectively. Only the first one increased the predictive power in comparison to the literature average. The best models of (2) and (3) achieved an analysis time of 32 minutes. Both decreased the time of analysis in comparison to the literature average. **Conclusion:** Strategies based on the multivariate analysis of first trimester gyneco-obstetrical and/or serum spectral data improve the prediction of GDM in early pregnancy, either by increasing the predictive power or by reducing the time of analysis.

1. INTRODUCCIÓN

1.1. Necesidad de predecir diabetes gestacional antes de las 24-28 semanas de gestación

1.1.1. Diabetes gestacional

La Asociación Americana de Diabetes define la diabetes gestacional (DG) como una patología diagnosticada en el segundo o tercer trimestre del embarazo, que no constituye claramente una diabetes pregestacional [1]. En otras palabras, corresponde a un estado de hiperglicemia que es diagnosticado por primera vez durante la gestación, con niveles de glucosa que son inferiores a los que se utilizan para el diagnóstico de diabetes fuera del embarazo [2]. Esta definición incluye alteraciones tanto de la glicemia en ayunas como de la tolerancia a la glucosa. En efecto, en Chile el diagnóstico de DG se lleva a cabo de acuerdo a los siguientes criterios: glicemia en ayunas entre 100 y 125 mg/dl en dos días diferentes y/o glicemia a las 2 horas de una carga de 75 g de glucosa mayor o igual a 140 mg/dl en el segundo o tercer trimestre del embarazo [3].

La DG es actualmente la complicación más común del embarazo a nivel mundial [2]. De acuerdo a una estimación realizada por la Federación Internacional de Diabetes, en 2021 el 16,7% de los nacidos vivos en el mundo padecieron de

algún grado de hiperglicemia en el embarazo; de ellos, un 80,3% correspondió a DG [4]. Es decir, la prevalencia mundial de DG en 2021 fue de un 13,4%. Asimismo, Garmendia y colaboradores reportaron que la prevalencia de DG en Chile fue de 13,0% en 2015 [5]. De manera interesante, ambos estudios dan cuenta de que la prevalencia de DG está aumentando. El primero proyecta un incremento de los casos de DG a nivel internacional para los años 2030 y 2045 [4], en tanto el segundo informa un incremento de un 195% en la prevalencia nacional de esta patología entre los años 2002 y 2015 [5].

Desde un punto de vista fisiopatológico, la DG puede ser explicada principalmente por dos alteraciones: un incremento de la resistencia a la insulina y una disfunción de las células β pancreáticas. En un embarazo normal, la placenta secreta hormonas que promueven un estado fisiológico de resistencia a la insulina, lo que conlleva un incremento en los niveles plasmáticos de glucosa y su posterior transporte a través de la placenta para ser utilizada como fuente de energía en el crecimiento fetal; y la consecuente hiperplasia e hipertrofia de las células β pancreáticas para mantener la homeostasis de la glucosa durante el embarazo [6]. En muchos casos de DG existen alteraciones previas a la concepción de la sensibilidad a la insulina y de la función de las células β pancreáticas, defectos que solo se manifiestan clínicamente en el marco de las demandas metabólicas propias del embarazo [2]. Así, en un embarazo con DG, las células β pancreáticas fallan en compensar el incremento fisiológico de la

glicemia inducido por las hormonas placentarias lo que, combinado con una resistencia a la insulina exacerbada, resulta en hiperglicemia [6].

La hiperglicemia en el embarazo tiene consecuencias negativas para la salud de la madre y el hijo, tanto a corto como a largo plazo [7]. Por ejemplo, a corto plazo, la madre con DG tiene un mayor riesgo de desarrollar preeclampsia (OR 1,28; 95% CI: 1,20-1,37) y de culminar su embarazo con una cesárea de emergencia (OR 1,08; 95% CI: 1,03-1,12) [8]; mientras que a largo plazo es más propensa a desarrollar prediabetes (OR 3,07; 95% CI: 2,53-3,73) y diabetes mellitus tipo 2 (DM2, OR 5,44; 95% CI: 3,68-8,08) [9]. Asimismo, el hijo de una madre con DG presenta un mayor riesgo a corto plazo de presentar macrosomía (OR 1,38; 95% CI: 1,32-1,44), distocia de hombros (OR 1,22; 95% CI: 1,09-1,37), hiperbilirrubinemia (OR 1,08; 95% CI: 1,02-1,13) e hipoglicemia neonatal (OR 1,10; 95% CI: 1,00-1,12), además de nacer por parto prematuro (OR 1,16; 95% CI: 1,10-1,23) y requerir cuidados intensivos (OR 1,09; 95% CI: 1,03-1,14) [8]; en tanto a largo plazo es más proclive a desarrollar obesidad (OR 1,58; 95% CI: 1,24-2,01) [9] e intolerancia a la glucosa (OR 1,96; 95% CI: 1,41-2,73) [10].

1.1.2. Necesidad de predecir diabetes gestacional en etapas tempranas del embarazo

El diagnóstico de DG se realiza en el segundo o tercer trimestre del embarazo; particularmente entre las semanas 24 y 28 de gestación [1,3]. Sin embargo, existe evidencia reciente que sugiere que su detección en este periodo podría ser demasiado tardía. A saber, considerando que la macrosomía es una de las alteraciones más comunes en los fetos de embarazos que cursan con DG, Sovio y colaboradores evaluaron distintos indicadores de crecimiento fetal en diferentes etapas del embarazo de mujeres con esta patología. Si bien los autores no encontraron diferencias estadísticamente significativas en la biometría fetal respecto a los fetos de embarazos normales a la semana 20 de gestación, sí encontraron alteraciones importantes a la semana 28, lo que les permitió concluir que el diagnóstico de DG es precedido por un crecimiento excesivo del feto entre las semanas 20 y 28 de gestación, lo que se manifiesta principalmente en un incremento de la circunferencia abdominal fetal [11]. Asimismo, y en vista de que los hijos de madres con DG no solo desarrollan macrosomía fetal, sino que también tienen un mayor riesgo de desarrollar obesidad a largo plazo, Venkataraman y colaboradores examinaron la adiposidad de fetos de mujeres con esta patología en etapas tempranas del embarazo. Los autores encontraron que los fetos de mujeres con DG presentan un grosor de pared abdominal anterior significativamente mayor que los fetos de embarazos normales a las 20

semanas de gestación, concluyendo que al tiempo del diagnóstico de esta enfermedad, la adiposidad abdominal fetal ya se encuentra desproporcionadamente incrementada [12]. Por su parte, y tomando en cuenta que la hiperglicemia en el embarazo se asocia a alteraciones cardiacas en la descendencia, Yovera y colaboradores estudiaron la morfología y función del corazón de fetos de embarazos afectados por DG en diferentes ventanas temporales. Los autores encontraron que, en comparación con el grupo control, los fetos de mujeres con DG presentan una alteración significativa en la función del ventrículo derecho a las 24-32 semanas de embarazo, periodo que se superpone con el intervalo temporal en que se realiza el diagnóstico de esta patología [13]. En conjunto, estos estudios demuestran que, al tiempo en que se realiza el diagnóstico de DG, el fenotipo fetal ya está alterado, lo que sugiere que los eventos fisiopatológicos que conducen a estos efectos adversos comienzan antes de la manifestación clínica de esta patología.

Por lo tanto, dado que: (1) la prevalencia de DG está en aumento, (2) esta patología tiene consecuencias negativas para la madre y el hijo, a corto y largo plazo, y (3) el fenotipo fetal ya se encuentra alterado al tiempo del diagnóstico de esta enfermedad, es necesario implementar métodos que permitan la predicción de DG antes de las 24-28 semanas de gestación. El detectar esta patología en etapas más tempranas del embarazo permitiría iniciar su tratamiento de manera oportuna y así prevenir sus efectos adversos para la salud materna y filial a corto

y largo plazo, lo que también podría reducir los costos que éstos conllevan. Solo en términos de cuidado prenatal y postparto, esta enfermedad se asocia a un incremento del costo por caso de €263-€13.680 [14], lo que equivale a $237.396-$12.348.222. De manera interesante, en 2018 tanto el Instituto Nacional de Diabetes y Enfermedades Digestivas y Renales de Estados Unidos como la Federación Internacional de Ginecología y Obstetricia identificaron a la búsqueda de marcadores precoces de DG como un área prioritaria para investigaciones futuras [15,16].

1.2. Estrategias univariadas para la predicción de diabetes gestacional antes de las 24-28 semanas de gestación

Se han llevado a cabo múltiples intentos para encontrar un biomarcador que permita predecir, por sí solo y con máxima sensibilidad y especificidad, un cuadro de DG antes de las 24-28 semanas de gestación [17–20]. Algunas estrategias son determinaciones que se relacionan con la alteración de la homeostasis de la glucosa propia de esta enfermedad. Por ejemplo, se ha descrito que la hemoglobina glicosilada (HbA1c) en el primer trimestre de gestación, tiene una gran especificidad (94-100%) para la predicción de DG, sin embargo presenta una muy baja sensibilidad (13-29%) [17], lo que limita su uso como método de detección temprana para esta patología. De la misma manera, la glicemia en ayunas en el primer trimestre de embarazo, que permite la predicción de DG con

una especificidad de 77% y una sensibilidad de 47%, suele tener asociada una alta tasa de falsos positivos [17], lo que también restringe su aplicación con este fin. Por su parte, la prueba de tolerancia a la glucosa oral (PTGO) con 75 g de glucosa en el primer trimestre de gestación tiene una alta sensibilidad (87%) y especificidad (100%) para la detección de DG [17], sin embargo presenta una baja reproducibilidad [21,22] lo que, sumado a que es un método que consume mucho tiempo y que es desagradable para algunas pacientes [23]; coarta su potencial como método de predicción para esta enfermedad. Además, cabe señalar que no existen criterios claros para establecer la presencia de hiperglicemia en etapas tempranas del embarazo [15,16], por lo que los métodos que se basan en marcadores glicémicos para la detección precoz de DG otorgan resultados inciertos y pueden conducir a conclusiones erróneas.

Otras aproximaciones para la detección temprana de DG se basan en el análisis de biomarcadores proteicos en suero durante el primer trimestre del embarazo. Algunos de estos marcadores son la adiponectina, la proteína C reactiva (hs-CRP), la globulina fijadora de hormonas sexuales (SHBG), la fibronectina glicosilada, el receptor soluble de pro-renina (s(Pro)RR) y la ciclofilina A (CyPA), que permiten la predicción de esta patología con sensibilidades de 28%, 89%, 85%, 81%, 70% y 78%, respectivamente; y especificidades de 90%, 55%, 37%, 90%, 70% y 75%, respectivamente [20]. Asimismo, se han llevado a cabo estrategias diagnósticas basadas en el análisis de micro-ARN (miR) en suero;

por ejemplo, con miR-16-5p, miR-17-5p y miR-20a-5p. Si bien estos miR permiten la predicción de DG con una alta especificidad (95-98%) en etapas tempranas del segundo trimestre de gestación, su sensibilidad es muy baja (18-42%) [20]. De igual modo, se han evaluado otros biomarcadores para la detección temprana de DG, entre ellos: insulina en ayunas, leptina, ferritina, la enzima alanina aminotransferasa (ALT), el factor de necrosis tumoral α (TNF-α) y algunas moléculas derivadas de la placenta como la proteína 3 relacionada con folistatina (FSTL3) y el factor de crecimiento placentario (PLGF) [18]. Si bien algunas de estas estrategias permiten predecir DG con alta sensibilidad y especificidad, ninguna supera el desempeño de la PTGO en el primer trimestre de embarazo, cuyas múltiples desventajas ya fueron señaladas.

En suma, las estrategias univariadas que se han desarrollado para la predicción de DG antes de las 24-28 semanas de gestación tienen distintas desventajas que limitan su aplicación con dicho fin, siendo la más común el presentar un bajo poder predictivo. Ciertamente, debido a la complejidad de la fisiopatología de la DG, es muy poco probable que un único parámetro tenga la sensibilidad y especificidad suficientes para su detección temprana. Dado que la combinación de biomarcadores suele mejorar la exactitud de predicción respecto a estrategias que utilizan biomarcadores individuales [18], es más probable que un modelo construido en base a la conjunción de múltiples variables logre el desempeño necesario para su predicción [23].

1.3. Estrategias multivariadas para la predicción de diabetes gestacional antes de las 24-28 semanas de gestación

1.3.1. Análisis multivariado

El análisis multivariado es una rama de la estadística que comprende el estudio de múltiples mediciones o variables, correspondientes a una o varias muestras [24]. Dependiendo del área en que se aplique, puede tener distintos nombres: en química, *chemometrics*; en biología, *biometrics*; en biomedicina, *machine learning;* en psicología, *psychometrics*; en economía, *econometrics* [25]. El análisis multivariado involucra distintas técnicas, que pueden ser no supervisadas o supervisadas (**Figura 1-1**). Las técnicas no supervisadas tienen por objetivo explorar la estructura que subyace a la data en estudio. En contraste, las técnicas supervisadas pretenden predecir una propiedad de interés, ya sea la clase a la cual pertenece una muestra, o bien un parámetro continuo asociado a la misma. En el primer caso, se utilizan técnicas de clasificación y, en el segundo, de regresión. Las técnicas no supervisadas no utilizan información sobre la propiedad que se desea predecir y, por ende, permiten evaluar la agrupación espontánea de las muestras que conforman la data en estudio, según los patrones de variables que las caracterizan. Las técnicas supervisadas sí utilizan dicha información, y producen la distribución de las muestras en base a los patrones de variables que se correlacionan con la propiedad a predecir, con lo

que se genera un modelo que posteriormente permite predecir la misma propiedad en muestras incógnitas.

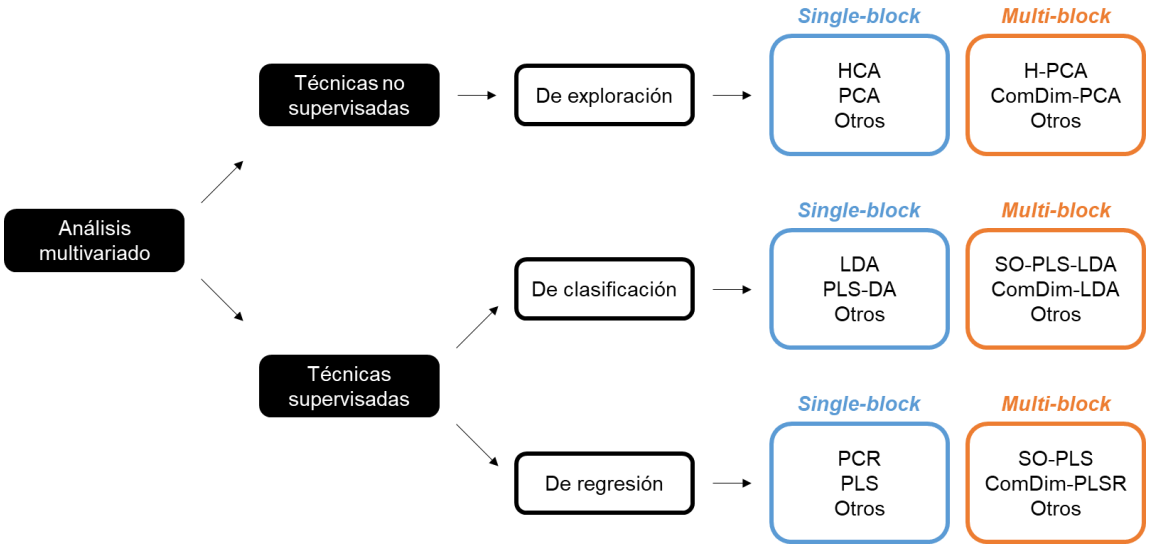


Figura 1-1. Tipos de técnicas multivariadas. *HCA: hierarchical cluster analysis; PCA: principal component analysis; H-PCA: hierarchical PCA; ComDim-PCA: common dimensions PCA; LDA: linear discriminant analysis; PLS-DA: PLS discriminant analysis; SO-PLS-LDA: sequential orthogonalized PLS-LDA; ComDim-LDA: common dimensions LDA; PCR: principal component regression; PLS: partial least squares; SO-PLS: sequential orthogonalized PLS; ComDim-PLSR: common dimensions PLS regression.*

Para estudiar una determinada condición mediante análisis multivariado no supervisado o supervisado, es posible contar con un solo bloque de datos, o bien con múltiples bloques de datos de distinto origen. En el primer caso, es suficiente aplicar técnicas multivariadas convencionales, que permiten realizar análisis del tipo *single-block*. Las técnicas *single-block* presentan diversas ventajas por sobre las estrategias univariadas: permiten mejorar el poder predictivo, lograr una

visualización integral de toda la información disponible, e identificar patrones de variables clave para el desempeño de un modelo [26,27]. Por el contrario, en el segundo caso, es necesario recurrir a técnicas multivariadas más avanzadas, que permiten llevar a cabo análisis del tipo *multi-block*, también conocido como *data fusion*. Las técnicas *multi-block* comparten las principales características de las técnicas *single-block* y, además, permiten identificar la información que es común entre los diferentes bloques, y también la que es distintiva para cada bloque. Lo anterior posibilita una mejor comprensión de la relación entre las variables en estudio y, generalmente, la obtención de modelos de mayor poder predictivo [26,28].

Los métodos multivariados han sido ampliamente utilizados para enfrentar y resolver problemáticas del área biomédica [29,30]. De manera interesante, recientemente se ha planteado que, en el tiempo, el análisis multivariado logrará transformar distintas áreas de la biomedicina, entre ellas el diagnóstico clínico [31]. Sin embargo, antes de su implementación en la realidad clínica, los modelos predictivos deben ser validados. Existen dos tipos de validación, la interna y la externa. La validación interna consiste en evaluar el desempeño del modelo mediante la predicción de la mismas muestras que fueron utilizadas para entrenarlo, mediante técnicas de *resampling*, tales como validación cruzada dejando uno o más fuera (*leave-one-out* y *leave-more-out cross-validation*), validación cruzada simple o doble (*single* y *double cross-validation*),

bootstrapping o permutación [32]. Por otro lado, la validación externa consiste en evaluar el desempeño del modelo por medio de la predicción de muestras que no fueron utilizadas para entrenarlo, por ejemplo, con una nueva cohorte de la misma institución de salud en un periodo temporal distinto (validación externa temporal), o de un centro médico, ciudad o país diferente (validación externa geográfica) [33]. La validación externa puede ser realizada dentro del mismo estudio que desarrolla el modelo, o bien en un estudio independiente. El primer caso generalmente involucra una cohorte similar a la de entrenamiento, en tanto el segundo asegura que la población de prueba es diferente. Ya que un modelo multivariado puede presentar un alto poder predictivo en un grupo particular de sujetos, pero uno bajo en otros, es crucial evaluar su desempeño en diferentes oportunidades y contextos. En efecto, si un modelo predictivo no está validado en distintas poblaciones, no es apropiado para su aplicación clínica [34].

1.3.2. Estrategias multivariadas que predicen diabetes gestacional en etapas tempranas del embarazo en base a datos gineco-obstétricos

Muchos autores han desarrollado modelos multivariados para la predicción de DG antes de las 24-28 semanas de gestación. La mayoría de estas estrategias se basan en variables gineco-obstétricas [35], es decir, en antecedentes clínicos y/o parámetros bioquímicos maternos (**Tabla 1-1**). El término data gineco-obstétrica abarca antecedentes maternos que pueden ser recabados mediante

anamnesis y examen físico en la práctica clínica, e.g. historia médica personal y familiar, factores antropométricos, información demográfica, entre otros; y marcadores bioquímicos que pueden ser identificados o cuantificados en muestras biológicas maternas mediante métodos de laboratorio clínico, e.g. perfil bioquímico, perfil lipídico, hormonas, anticuerpos, variantes genéticas, entre otros. Mientras algunos estudios consideran solo antecedentes clínicos [36,37] o solo parámetros bioquímicos [38,39], la mayor parte de ellos combina ambos tipos de variables [40–46]. De acuerdo a un reciente meta-análisis, las diez variables más comunes en este tipo de modelos son: edad materna, historia familiar de diabetes, índice de masa corporal o IMC, glicemia en ayunas, IMC pregestacional, historia de diabetes, etnicidad, triglicéridos, hemoglobina A1c y presión sistólica; en tanto la técnica multivariada más utilizada es la regresión logística (LR) [47]. Estas estrategias permiten la predicción de DG con desempeños muy variados, existiendo algunas de muy bajo poder predictivo, y otras de muy alto poder predictivo [35]. En promedio, su poder predictivo es moderado. De manera interesante, más de la mitad de los trabajos que proponen modelos basados en datos de origen médico, no utiliza ninguna estrategia de validación [35]. Por lo tanto, su desempeño podría estar sobreestimado. Por otro lado, los pocos modelos de este tipo que han sido validados en estudios independientes, han presentado un limitado poder predictivo [35].

Los métodos multivariados basados en datos gineco-obstétricos se cimentan en gran medida en información que se recopila de manera habitual y sistemática en la primera visita prenatal de una embarazada, lo que hace realista su implementación clínica. Además, dado su bajo costo, estas estrategias tienen el potencial de ser aplicadas a todas las gestantes. En otras palabras, un modelo de este tipo podría ser utilizado de manera universal, lo que cobra gran relevancia en el marco de la detección temprana de DG cuando se considera que entre un 17 y un 48% de las gestantes que desarrollan esta patología no presentan factores de riesgo evidentes [48]. De acuerdo a lo anterior, estos métodos son una excelente alternativa para la predicción de DG en etapas tempranas del embarazo.

Ya que los modelos basados en data gineco-obstétrica que se reportan en literatura suelen presentar una exactitud moderada para la detección de DG antes de las 24-28 semanas del embarazo, existe la oportunidad de mejorar la predicción de esta patología mediante el desarrollo y la validación de nuevos métodos que consideren tanto variables clínicas y/o bioquímicas como herramientas multivariadas que no hayan sido evaluadas en estudios anteriores.

Tabla 1-1. Estudios que reportan el uso de métodos multivariados basados en data gineco-obstétrica para la predicción de DG previo a las 24-28 semanas de embarazo.

Tipo de variables	Técnica multivariada	Desempeño o poder predictivo*				Tipo de validación*	Ref
		Se	Es	Ex	AUC		
Clínicas	LR	61%	71%	-	0,70	Independiente	[36]
	PCA; RF	-	-	78%	-	Interna	[37]
Bioquímicas	LR	-	-	-	0,72	Ninguna	[38]
	BLCM	94%	86%	-	0,92	Ninguna	[39]
Clínicas y bioquímicas	LR; BN; NN; SVM; CHAIDT; CSHM	62%	100%	-	-	Externa	[40]
	LR	80%	44%	-	0,73	Ninguna	[41]
	LR	83%	80%	-	0,90	Ninguna	[42]
	LR	86%	80%	-	0,91	Ninguna	[43]
	LR	69%	80%	-	0,81	Ninguna	[44]
	LR; BI-LR	66%	75%	64%	0,77	Ninguna	[45]
	GB	90%	63%	-	0,85	Externa	[46]

*LR: logistic regression; PCA: principal component analysis; RF: random forest; BLCM: Bayesian latent class modeling; BN: Bayesian network; NN: neural network; SVM: support vector machine; CHAIDT: chi-square automatic interaction detection tree; CSHM: cost-sensitive hybrid method; BI: Bayesian inference; GB: gradient boosting; Se: sensibilidad; Es: especificidad; Ex: exactitud; AUC: área bajo la curva ROC; ROC: receiver operating characteristic; Ref: referencia. * Información relativa al mejor modelo reportado por el estudio en cuestión.*

1.3.3. Estrategias multivariadas que predicen diabetes gestacional en etapas tempranas del embarazo en base a métodos instrumentales de alta complejidad

Otros autores han desarrollado estrategias para predecir DG antes de las 24-28 semanas de gestación en base a datos que se obtienen mediante métodos instrumentales de alta complejidad (**Tabla 1-2**). Este término abarca métodos cuyo uso actual se restringe al campo de la investigación, esto es, que no se aplican en laboratorio clínico. Por ejemplo, se han reportado métodos basados

en micro ácidos ribonucleicos [49], ácidos desoxirribonucleicos libres [50] y géneros de microbiota [51,52], determinados por la técnica instrumental de biología molecular reacción en cadena de la polimerasa. También se han publicado modelos basados metabolitos, estudiados por las técnicas instrumentales de química analítica cromatografía acoplada a espectrometría de masas [53,54] y espectroscopía de resonancia magnética nuclear protónica [55,56]. De igual manera, se han reportado estrategias basadas en péptidos y proteínas, analizados por la técnica instrumental cromatografía acoplada a espectrometría de masas [57,58]. Existen estudios que utilizan solo data instrumental de alta complejidad [49,51,53,55,57], en tanto otros la combinan con información gineco-obstétrica [50,52,54,56,58]. Independiente de lo anterior, las estrategias de este tipo presentan desempeños muy variados, existiendo algunas de muy bajo poder predictivo, y otras de muy alto poder predictivo [35]. En promedio, su poder predictivo es moderado-alto. De manera interesante, la mayoría de los trabajos que reportan modelos para predecir DG en base a métodos instrumentales de alta complejidad, utilizan alguna táctica de validación, siendo la validación interna la más usada [35].

Si bien algunas aproximaciones de este tipo permiten la predicción de DG con gran exactitud y cierto grado de confianza dado su estado de validación, los métodos en que se basan consumen mucho tiempo, ya sea porque requieren un laborioso tratamiento de muestra previo, porque la técnica instrumental

propiamente tal se asocia a largos tiempos de análisis, o porque la interpretación de la información analítica que se genera es compleja y operador-dependiente, lo que dificulta su implementación clínica. De manera interesante, existen métodos basados en técnicas instrumentales más rápidas, por ejemplo, técnicas de espectroscopía vibracional; sin embargo, su potencial utilidad en el marco de la detección temprana de esta patología aún no ha sido evaluada.

Tabla 1-2. Estudios que reportan el uso de métodos multivariados basados en métodos instrumentales de alta complejidad para la predicción de DG previo a las 24-28 semanas de embarazo.

Tipo de variables	Matriz de estudio	Técnica analítica	Técnica multivariada	Desempeño o poder predictivo*					Tipo de validación*	Ref
				Se	Es	Ex	AUC	Q^2		
miRNA	Plasma	PCR	PCA; LR; RF; AB	94%	80%	90%	0,91	-	Interna	[49]
cfDNA	Plasma	PCR	HCA; LR	71%	71%	71%	0,71	-	Externa	[50]
Microbiota	Heces	PCR	LDA; RF	78%	69%	-	0,74	-	Interna	[51]
	Heces	PCR	LDA	-	-	-	0,70	-	Externa	[52]
Metabolitos	Suero y orina	XC-MS	OPLS-DA	-	-	-	-	0,45	Interna	[53]
	Plasma y orina	XC-MS	PCA; PLS-DA; SVM	-	-	-	0,99	-	Interna	[54]
	Plasma y orina	^{1}H-NMR	PCA; PLS-DA; OPLS-DA	-	-	-	-	0,28	Interna	[55]
	Suero	^{1}H-NMR	LR	-	-	-	0,77	-	Interna	[56]
Péptidos y proteínas	Suero	XC-MS	LR	80%	95%	-	0,99	-	Ninguna	[57]
	Suero	XC-MS	LR	75%	64%	-	0,73	-	Externa	[58]

*miRNA: micro ribonucleic acid; cfDNA: cell free deoxyribonucleic acid; PCR: polymerase chain reaction; XC: liquid or gas chromatography; MS: mass spectrometry; NMR: nuclear magnetic resonance; PCA: principal component analysis; LR: logistic regression; RF: random forest; AB: adaptive boosting; HCA: hierarchical cluster analysis; LDA: linear discriminant analysis; OPLS-DA: orthogonal PLS-DA; PLS-DA: partial least squares discriminant analysis; SVM: support vector machine; Se: sensibilidad; Es: especificidad; Ex: exactitud; AUC: área bajo la curva ROC; ROC: receiver operating characteristic; Ref: referencia. * Información relativa al mejor modelo reportado por el estudio en cuestión.*

1.4. Potencial de técnicas de espectroscopía vibracional para la predicción de diabetes gestacional antes de las 24-28 semanas de gestación

1.4.1. Espectroscopía vibracional

La espectroscopía vibracional (EV) es el estudio de la interacción de la radiación electromagnética con las vibraciones moleculares [59]. Ésta involucra distintas técnicas instrumentales, entre ellas la espectroscopía infrarroja y la espectroscopía Raman. La primera mide la absorción, transmisión o reflexión de la radiación en el infrarrojo medio (MIR) o cercano (NIR); mientras que la segunda mide la radiación dispersada de manera inelástica por una muestra tras ser incidida por un haz de luz monocromático en el rango ultravioleta, visible o NIR [60]. Estas técnicas tienen múltiples ventajas: no son invasivas ni destructivas, requieren bajo volumen y mínima preparación de muestra, no precisan reactivos específicos ni generan desechos, son simples, rápidas, de bajo costo y fácilmente automatizables [61].

La EV provee información sobre la composición química de la materia. En efecto, ya que las biomoléculas son activas en espectroscopía infrarroja y Raman [61], estas técnicas permiten la obtención de un patrón espectral característico para una muestra biológica a partir de todos sus constituyentes. Además, al ser combinadas con técnicas multivariadas, éstas permiten la distinción entre

muestras según su composición bioquímica, lo que da cabida a su uso con fines de apoyo diagnóstico en base a la comparación de muestras sanas y patológicas. Es así como se han desarrollado estrategias basadas en EV y análisis multivariado para la detección de cáncer, VIH/SIDA, diabetes, asma, entre otras enfermedades [60].

1.4.2. Uso de técnicas de espectroscopía vibracional en estudios que investigan diabetes gestacional

Pocos trabajos han reportado el uso de EV en el marco de DG (**Tabla 1-3**). La mayoría ha utilizado espectroscopía MIR. Por ejemplo, Si y colaboradores utilizaron espectroscopía MIR para caracterizar el perfil de polisacáridos de diferentes preparados de ajo fresco y negro. El ajo negro fue administrado diariamente por 40 semanas, solo o en combinación con *Lactobacillus bulgaricus*, a embarazadas con DG. La intervención nutricional inició alrededor de las 12 semanas de embarazo. Los autores observaron que, luego de 24-28 semanas de tratamiento, el grupo que recibió el preparado con el probiótico presentó una menor concentración plasmática de marcadores prooxidantes, una mayor concentración plasmática de marcadores antioxidantes, una menor glicemia en ayunas y post carga, y una menor incidencia de complicaciones perinatales, que el grupo control. De acuerdo a lo anterior, los autores concluyeron que *L. bulgaricus* mejora la capacidad antioxidante del ajo negro en pacientes con DG

[62]. Por su parte, Shapira y colaboradores utilizaron espectroscopía MIR para comparar el contenido nutricional de la leche materna de mujeres que cursaron un embarazo con DG respecto al de mujeres que no. A partir de estos experimentos los autores demostraron que las muestras de leche madura (14 días post parto) de las madres con DG tienen un menor contenido graso y energético que aquellas con tolerancia normal a la glucosa (TNG) [63]. Si bien los dos artículos anteriores aplican espectroscopía MIR en investigaciones relativas a DG, ninguno tiene como fin su detección temprana. Otra publicación que reporta el uso de espectroscopía MIR en el contexto de DG es la de Zhou y colaboradores. Los autores utilizaron espectroscopía MIR para comparar tejido placentario de embarazos que cursaron con DG, y de embarazos control. El análisis de los perfiles espectrales con técnicas multivariadas permitió identificar diferencias biomoleculares entre los dos grupos de estudio. Las placentas de embarazos que cursaron con DG presentaban alteraciones en la conformación de estructuras lipídicas, un mayor nivel de oxidación en lípidos, una mayor razón de estructuras proteicas β en relación a α-hélice, además de alteraciones en la composición de carbohidratos, ácidos nucleicos y fosfolípidos, respecto al grupo control. De manera interesante, el análisis multivariado de la región espectral vinculada a lípidos (1800–1480 cm^{-1}) se asoció a un valor de Q^2 de 0,876, lo que da cuenta del poder discriminante que tiene la información espectral MIR para diferenciar las muestras de embarazos con y sin DG [64]. De manera similar, Bernardes-Oliveira y colaboradores utilizaron espectroscopía MIR para comparar

muestras de plasma de mujeres con y sin DG, a las 12-38 (DG) o 9-39 (control) semanas de embarazo. Los autores analizaron el rango espectral de huella dactilar, de 1800 a 900 cm^{-1}, con diferentes técnicas multivariadas. Su mejor modelo permitió diferenciar las muestras de plasma de ambos grupos con un 100% de sensibilidad, especificidad y exactitud [65]. En primera instancia se podría pensar que las estrategias presentadas por Zhou y colaboradores, y Bernardes-Oliveira y colaboradores, podrían ser utilizadas para la detección temprana de DG. Sin embargo, si bien ambas permiten diferenciar muestras biológicas de mujeres con y sin DG con un alto poder discriminante, el momento de toma de muestra que reportan es incompatible con la predicción de esta patología en etapas tempranas del embarazo. El estudio de Zhou y colaboradores compara muestras de placenta, las cuales se obtienen una vez finalizado el embarazo. Los patrones espectrales identificados como relevantes para diferenciar los dos grupos de estudio son representativos del último periodo de la gestación. Lo anterior, junto a la imposibilidad de recolectar muestras de tejido placentario durante un embarazo en curso, implica que esta estrategia no es compatible con la predicción de DG antes de las 24-28 semanas de gestación. Por otro lado, el estudio de Bernardes-Oliveira comparó muestras de plasma recolectadas en un amplio rango de tiempo: 9-39 semanas de embarazo para el grupo control, y 12-38 semanas de embarazo para el grupo con DG. Pese a que esta ventana temporal incluye semanas que son previas al diagnóstico de DG, los autores no detallan cuántas muestras fueron recolectadas antes de ese momento, ni cuántas

fueron recolectadas después. Este es un detalle muy importante, dado que es mucho más probable encontrar alteraciones bioquímicas una vez que la patología ya está establecida, esto es, posterior al diagnóstico. El uso de muestras posteriores a las 24-28 semanas de gestación podría haber sesgado los resultados de este trabajo. Tanto los patrones espectrales identificados como relevantes para diferenciar los dos grupos de estudio, como el desempeño de los modelos multivariados desarrollados, podrían estar fuertemente influenciados por la presencia de muestras recolectadas después del diagnóstico de DG. En consecuencia, aunque esta estrategia da cuenta del gran potencial que tiene la espectroscopia MIR para distinguir pacientes con y sin DG, tampoco es apta para ser utilizada en la detección temprana de esta patología.

También existen publicaciones enfocadas en DG que reportan el uso de espectroscopia NIR. Dipla, Kintiraki, Vounzoulaki y colaboradores utilizaron espectroscopía NIR para evaluar alteraciones en la oxigenación músculo-esquelética y cerebral de embarazadas con DG. Para ello realizaron mediciones continuas del rango espectral NIR en el antebrazo y la corteza prefrontal de las mujeres en estudio. Mediante estas determinaciones, los autores concluyeron que existe una reducción en la oxigenación de los dos órganos tanto durante [66,67] como después [68] del embarazo con DG. Pese a que los tres artículos mencionados utilizan espectroscopía NIR en investigaciones que se enfocan en DG, ninguno tiene como fin su detección temprana.

Por otro lado, existe un solo artículo que reporta el uso de espectroscopía Raman en el marco de DG. Kim y colaboradores desarrollaron un inmunoensayo para cuantificar adiponectina en el suero de mujeres con y sin DG. Los autores utilizaron nanopartículas de oro conjugadas a anticuerpos contra adiponectina, y a una sonda detectable por espectroscopía Raman. El inmunoensayo permitió cuantificar adiponectina sérica a las 9-26 semanas de embarazo con un amplio rango dinámico (10^{-15}-10^{-6} g/mL), un buen ajuste en el rango de concentraciones de relevancia clínica (R^2 = 0,994), un límite de detección en la escala de los femto-gramos por mL ($3{,}0\times10^{-16}$ g/mL), y una excelente selectividad [69]. Este estudio tiene un matiz diagnóstico, dado que la adiponectina es considerada un posible biomarcador de DG. En el método descrito, la espectroscopía Raman se limita a ser un mero método de detección. Ningún estudio ha utilizado el perfil espectral Raman de muestras biológicas para la predicción de DG.

Tabla 1-3. Estudios que reportan el uso de técnicas de espectroscopía vibracional en DG.

Técnica de EV	Técnica multivariada	Matriz de estudio	Momento de análisis	Finalidad de utilizar EV	Ref
MIR	-	Preparado de ajo negro	-	Caracterizar preparados de ajo negro, previo a intervención nutricional en DG	[62]
	-	Calostro y leche	PP	Evaluar alteración del contenido nutricional de leche materna tras DG	[63]
	PCA; OPLS-DA	Placenta	PP	Comparar la composición biomolecular del tejido placentario de embarazos con y sin DG	[64]
	PCA; LDA; QDA; SVM	Plasma	9-39 o 12-38 SG	Diferenciar muestras de plasma de embarazadas con y sin DG	[65]

NIR	-	Antebrazo	26-30 SG	Evaluar alteración de la función microvascular y la oxigenación de músculo esquelético en DG	[66]
	-	Corteza prefrontal	26-32 SG	Evaluar alteración de la oxigenación cerebral en DG	[67]
	-	Antebrazo y corteza prefrontal	27 SG y PP	Evaluar alteración de la oxigenación músculo-esquelética y cerebral en y tras DG	[68]
Raman	-	Suero	9-26 SG	Evaluar inmunoensayo SERS para cuantificar adiponectina en embarazadas con y sin DG	[69]

MIR: mid-infrared; NIR: near-infrared; PP: post parto; SG: semanas de gestación; SERS: surface-enhanced Raman scattering; Ref: referencia.

1.4.3. Espectroscopía vibracional y análisis multivariado para la detección de patologías similares a diabetes gestacional

Si bien no se han desarrollado estrategias multivariadas basadas en EV para la predicción de DG en etapas tempranas del embarazo, sí se han utilizado aproximaciones de este tipo para la detección de enfermedades similares. En el marco de las patologías del embarazo, diferentes publicaciones reportan el uso de métodos basados en espectroscopía MIR [70–73], NIR [74,75] y Raman [76–79] con este fin (**Tabla 1-4**). Por ejemplo, se ha logrado la detección temprana de síndrome de dificultad respiratoria neonatal con un 83% de sensibilidad y un 97% de especificidad, a las 28-41 semanas de gestación [71]; parto prematuro y rotura temprana de membranas con un 74% y un 69% de exactitud, respectivamente, a las 14-25 semanas de embarazo [72]; parto prematuro con un 100% de sensibilidad y un 100% de especificidad, a las 12-20 semanas de embarazo [75];

y aborto espontáneo con un gran poder discriminante, esto es, con una distinción total entre muestras patológicas y control, a las 6-10 semanas de gestación [79]. Si bien el desempeño de los modelos reportados es variable, estos estudios demuestran que es factible aplicar métodos basados en EV y análisis multivariado para la predicción de patologías en etapas tempranas del embarazo.

Uno de los factores que influye en el desempeño de los modelos predictivos es la matriz que se analiza. En este contexto, es curioso que no se hayan desarrollado métodos para la predicción de patologías del embarazo en base al análisis de muestras que proceden de sangre, es decir, sangre total, suero o plasma, que son las muestras biológicas de más común obtención durante la gestación. Un estudio que da cuenta del potencial que tiene el uso de estas matrices para apoyo diagnóstico es el de Mukherjee y colaboradores. Los autores utilizaron espectroscopía MIR para analizar muestras de suero de embarazadas con preeclampsia de inicio temprano (24-34 semanas de gestación) y tardío (>34 semanas de gestación), y evaluaron diferencias metabólicas entre ambas condiciones, y también respecto a embarazadas que no cursaron con la patología. Mediante métodos multivariados, los autores lograron distinguir totalmente las muestras preeclámpticas de las control, y también las muestras de ambos grupos patológicos [73]. El poder discriminante de esta estrategia permite entrever la potencia que tendría un método multivariado basado en data espectral de suero para la detección temprana de patologías del embarazo.

Tabla 1-4. Publicaciones que reportan el uso de estrategias basadas en espectroscopía vibracional y análisis multivariado para la detección de patologías del embarazo.

Técnica de EV	Técnica multivariada	Matriz de estudio	Condición investigada	Desempeño o poder predictivo *				Ref
				Se	Es	Ex	C	
MIR	LDA; PLS	Líquido amniótico	Maduración pulmonar fetal	91%	82%	88%	-	[70]
	PLS	Líquido amniótico	Síndrome de dificultad respiratoria neonatal	83%	97%	-	-	[71]
	PCA; PLS-DA	Líquido amniótico	Malformación fetal	75%	85%	80%	-	[72]
			Parto prematuro	41%	87%	74%	-	
			Rotura temprana de membranas	70%	75%	69%	-	
	LDA; HCA; PCA	Suero	Preeclampsia	-	-	-	ST	[73]
NIR	PLS	Líquido amniótico	Maduración pulmonar fetal	95%	53%	-	-	[74]
	ILS	Líquido amniótico	Parto prematuro	100%	100%	-	-	[75]
Raman	PCA	Calostro	Infección por *Toxoplasma gondii*	-	-	-	ST	[76]
	PCA	Placenta	Preeclampsia	-	-	-	SP	[77]
	PCA; SVM	Líquido amniótico	Infección intra-amniótica y parto prematuro	>94%	>92%	>93%	-	[78]
	PCA; PLS	Orina	Aborto espontáneo	-	-	-	ST	[79]

MIR: mid-infrared; ATR: attenuated total reflectance; FT: Fourier transform; NIR: near-infrared; CM: confocal microscopy; SE: surface-enhanced; LDA: linear discriminant analysis; PLS: partial least squares; PCA: principal component analysis; PLS-DA: partial least squares discriminant analysis; HCA: hierarchical cluster analysis; ILS: inverse least squares; SVM: support vector machine; Se: sensibilidad; Es: especificidad; Ex: exactitud; C: cualitativo; ST: separación total; SP: separación parcial; Ref: referencia.

De la misma manera, se han desarrollado métodos multivariados basados en espectroscopía MIR [80,81], NIR [82,83] y Raman [84–87] para la detección de DM2 (**Tabla 1-5**), enfermedad que involucra alteraciones fisiopatológicas similares a las que ocurren en DG. En general, los modelos reportados permiten la predicción de esta patología con alta sensibilidad y especificidad. Esto podría

deberse, en parte, a que la adquisición espectral se realiza cuando la DM2 ya está establecida, lo cual aumenta la probabilidad de hallar diferencias entre el grupo patológico y el control. Independiente de lo anterior, destacan por su desempeño los modelos que se basan en el análisis espectral de muestras que provienen de sangre. Por ejemplo, Guang y colaboradores utilizaron espectroscopía MIR para analizar muestras de sangre total de personas con y sin DM2. Mediante métodos multivariados, los autores consiguieron detectar esta patología con un 95% de sensibilidad y un 96% de especificidad [81]. Por su parte, González-Solis y colaboradores utilizaron espectroscopía Raman para analizar muestras de suero de personas diabéticas y no diabéticas. A partir del análisis multivariado de la data espectral, los autores lograron la detección de DM2 con un 96% de sensibilidad y un 99% de especificidad [84]. De manera similar, Wang y colaboradores utilizaron espectroscopía Raman para evaluar similitudes y diferencias entre glóbulos rojos de humanos con TNG, humanos con DM2, ratas control y dos modelos murinos de DM2. Mediante herramientas multivariadas, los autores lograron distinguir los cinco tipos de eritrocitos con un 100% de exactitud [88]. En conjunto, estas publicaciones reafirman el potencial que tiene el uso de métodos multivariados basados en EV y muestras que proceden de sangre para la predicción de patologías similares a DG.

Tabla 1-5. Publicaciones que reportan el uso de estrategias basadas en espectroscopía vibracional y análisis multivariado para la detección de DM2.

Técnica de EV	Técnica multivariada	Matriz, sitio o data de estudio	Desempeño o poder predictivo *					Ref
			Se	Es	Ex	AUC	C	
MIR	LDA	Saliva	100%	75%	88%	-	-	[80]
	PCA; XGB	Sangre total	95%	96%	96%	-	-	[81]
NIR	PLS	Antebrazo	78%	70%	-	0,80	-	[82]
	BT; SVM; AB; GL; NN; RF	Imágenes de iris	99%	97%	90%	-	-	[83]
Raman	PCA; LDA	Suero	96%	99%	-	-	-	[84]
	NN; PCA; SVM	Lóbulo de oreja, interior de brazo, uña de pulgar, vena mediana cubital	89-91%	89-91%	89-91%	0,92-0,96	-	[85]
	PCA; SVM	Glóbulos rojos	-	-	100%	-	-	[86]
	CA; PLS	Vesículas extracelulares urinarias	-	-	-	-	ST	[87]

MIR: mid-infrared; FT: Fourier transform; ATR: attenuated total reflectance; NIR: near-infrared; CM: confocal microscopy; LDA: linear discriminant analysis; PCA: principal component analysis; XGB: extreme gradient boosting; PLS: partial least squares; BT: binary tree; SVM: support vector machine; AB: adaptive boosting; GL: generalized linear; NN: neural network; RF: random forest; CA: cluster analysis; Se: sensibilidad; Es: especificidad; Ex: exactitud; AUC: área bajo la curva ROC; ROC: receiver operating characteristic; C: cualitativo; ST: separación total; Ref: referencia.

En suma, los estudios publicados en el marco de la detección de patologías del embarazo y DM2 sugieren que una estrategia basada en análisis multivariado, EV y muestras que proceden de sangre permitiría la predicción de DG en etapas tempranas del embarazo con un alto poder predictivo.

2. PROBLEMA, HIPÓTESIS Y OBJETIVOS

2.1. Problema de investigación

Los métodos que se han desarrollado para la predicción de DG en etapas tempranas del embarazo son deficientes, ya sea por presentar un bajo poder predictivo o por requerir largos tiempos de análisis.

2.2. Hipótesis

Una estrategia basada en el análisis multivariado de datos gineco-obstétricos y/o espectrales de suero del primer trimestre de gestación permitirá mejorar la predicción de DG en etapas tempranas del embarazo, ya sea incrementando el poder predictivo, o reduciendo los tiempos de análisis.

2.3. Objetivo general

Evaluar métodos basados en el análisis multivariado de datos gineco-obstétricos y/o espectrales de suero para la predicción de DG en el primer trimestre del embarazo.

2.4. Objetivos específicos

Objetivo específico 1: Construir un modelo multivariado de clasificación a partir de datos gineco-obstétricos del primer trimestre de gestación de embarazadas con TNG y DG.

Objetivo específico 2: Implementar un método basado en técnicas de EV y análisis multivariado para la distinción entre muestras de suero del primer trimestre de gestación de embarazadas con TNG y DG.

Objetivo específico 3: Desarrollar un modelo multivariado basado en la combinación de datos gineco-obstétricos y espectrales de suero para la predicción de DG en el primer trimestre del embarazo.

3. ESTRATEGIA ANALÍTICA

Esta investigación se llevó a cabo de la siguiente manera:

3.1. Profundización en el estado del arte

Se buscaron publicaciones que reportaran el uso de métodos multivariados en el contexto de patologías y complicaciones del embarazo. Se encontraron trabajos que utilizaron análisis multivariado para estudiar DG, preeclampsia, muerte perinatal, aborto espontáneo, parto pretérmino, cesárea y malformación fetal. Para cada artículo se registró el objetivo asociado al hecho de aplicar herramientas de esta índole, el tipo de variables analizadas, la técnica multivariada utilizada, y el resultado de haber aplicado tal técnica. Esta revisión dio origen a una publicación en la revista *Frontiers in Endocrinology* [30], la cual se presenta en la sección 4.1.

Por otro lado, se buscaron publicaciones que reportaran métodos multivariados para predecir DG antes de las 24-28 semanas de embarazo. Se encontraron 109 artículos originales, y se categorizaron de acuerdo al tipo de variables que analizaron: médicas (que incluyen parámetros clínicos y/o bioquímicos), alternativas (que consideran metabolitos, péptidos o proteínas, variantes genéticas, géneros de microbiota, entre otras) y mixtas (que reúnen data médica

y alternativa). Para cada trabajo se registraron las variables utilizadas, el poder predictivo del mejor modelo, y su estado de validación. Esta revisión dio origen a una publicación en la revista *Artificial Intelligence in Medicine* [35], la cual se presenta en la sección 4.2.

Los hallazgos de esta revisión fueron utilizados para establecer los siguientes umbrales, con el fin de aceptar o rechazar la hipótesis planteada:

3.1.1. Umbral de poder predictivo

- Para modelo basado en data gineco-obstétrica: un área bajo la curva ROC (AUC) de 0,798 en validación interna, que es la media de estudios similares en literatura (i.e. artículos revisados en [35] que se basan en variables médicas, variantes genéticas, o ambas).

- Para modelo basado en data espectral de suero: un AUC de 0,820 en validación interna, que es la media de estudios similares en literatura (i.e. artículos revisados en [35] que se basan en variables alternativas, excepto variantes genéticas).

- Para modelo basado en data gineco-obstétrica y espectral de suero: un AUC de 0,804 en validación interna, que es la media de estudios similares en

literatura (i.e. artículos revisados en [35] que se basan en variables mixtas, excepto variables médicas + variantes genéticas).

3.1.2. Umbral de tiempo de análisis

- Para modelo basado en data espectral de suero: un tiempo de procesamiento de muestra y adquisición de data instrumental de 360 minutos, que es la media de estudios similares en literatura (i.e. artículos revisados en [35] que se basan en variables alternativas, excepto variantes genéticas).

- Para modelo basado en data gineco-obstétrica y espectral de suero: un tiempo de procesamiento de muestra y adquisición de data instrumental de 360 minutos, que es la media de estudios similares en literatura (i.e. artículos revisados en [35] que se basan en variables mixtas, excepto variables médicas + variantes genéticas).

3.2. Desarrollo del primer objetivo

Se tabularon 63 variables gineco-obstétricas del primer trimestre de gestación (53 parámetros clínicos y 10 bioquímicos) de 66 embarazadas. De acuerdo a la glicemia post carga (PTGO, 75 g, 2 h) del segundo trimestre de gestación y el criterio de referencia del Ministerio de Salud de Chile [3], de las 66 mujeres, 12

tenían DG y 54 TNG. Los dos grupos fueron descritos y comparados a través de técnicas clásicas de análisis univariado, tales como el test t de Student, el test de Mann-Whitney y el test exacto de Fisher. Luego, los datos fueron preprocesados por autoescalado, y analizados mediante diferentes técnicas multivariadas: la técnica no-supervisada *principal component analysis* (PCA); y las técnicas supervisadas de clasificación LR, *linear support vector machine* (L-SVM), *partial least squares discriminant analysis* (PLS-DA), *classification and regression tree* (CART) y *extreme gradient boosting* (XGB). Todos los modelos fueron evaluados según su sensibilidad, especificidad, tasa de no-error y AUC; tanto en calibración como en validación dejando uno fuera. Estos resultados dieron origen a una publicación en la revista *PLOS ONE* [89], la cual se presenta en la sección 4.3.

3.3. Desarrollo del segundo objetivo

Se reunieron muestras de suero del primer trimestre de gestación de 82 embarazadas. De acuerdo a la glicemia post carga (PTGO, 75 g, 2 h) del segundo trimestre de gestación y el criterio de referencia del Ministerio de Salud de Chile [3], de las 82 mujeres, 15 tenían DG y 67 TNG. Se adquirieron y promediaron cinco espectros NIR (modo de adquisición, transflectancia; rango, 10500-4000 cm^{-1}; resolución, 4 cm^{-1}; número de scans, 32) por cada muestra de suero (10 µl, secada a 37ºC por 30 minutos). Para el análisis multivariado, se evaluaron cuatro rangos espectrales, y 80 combinaciones distintas de transformaciones,

incluyendo suavizado, primera y segunda derivada con tamaño de ventana variable; corrección de dispersión por *standard normal variate* (SNV); corrección de línea base por *weighted least squares* (WLS); y normalización por norma Euclidiana. Los datos fueron preprocesados por centrado a la media, y analizados mediante la técnica supervisada de clasificación PLS-LDA. Todos los modelos fueron evaluados de acuerdo a su sensibilidad, especificidad, tasa de no-error y AUC en validación cruzada doble. Estos resultados dieron origen a un artículo, en preparación para ser enviado a la revista *Analytical and Bioanalytical Chemistry*, que se presenta en la sección 4.4.

3.4. Desarrollo del tercer objetivo

Se consideró la misma cohorte del segundo objetivo. Se tabularon 28 variables gineco-obstétricas del primer trimestre de gestación (28 parámetros clínicos). En paralelo, se seleccionó el rango y el pretratamiento espectral de mayor poder predictivo. Los datos gineco-obstétricos y espectrales de suero fueron preprocesados por autoescalado y centrado a la media, respectivamente; y posteriormente analizados mediante la técnica supervisada *multi-block* de clasificación *sequential and orthogonalized partial least squares linear discriminant analysis* (SO-PLS-LDA). Todos los modelos fueron evaluados según su sensibilidad, especificidad, tasa de no-error y AUC en validación cruzada doble. Estos resultados dieron origen a un artículo, en preparación para ser

enviado a la revista *Analytical and Bioanalytical Chemistry*, que se presenta en la sección 4.4.

La estrategia analítica de este trabajo se resume a la forma de flujograma de decisiones (**Figura 3-1**).

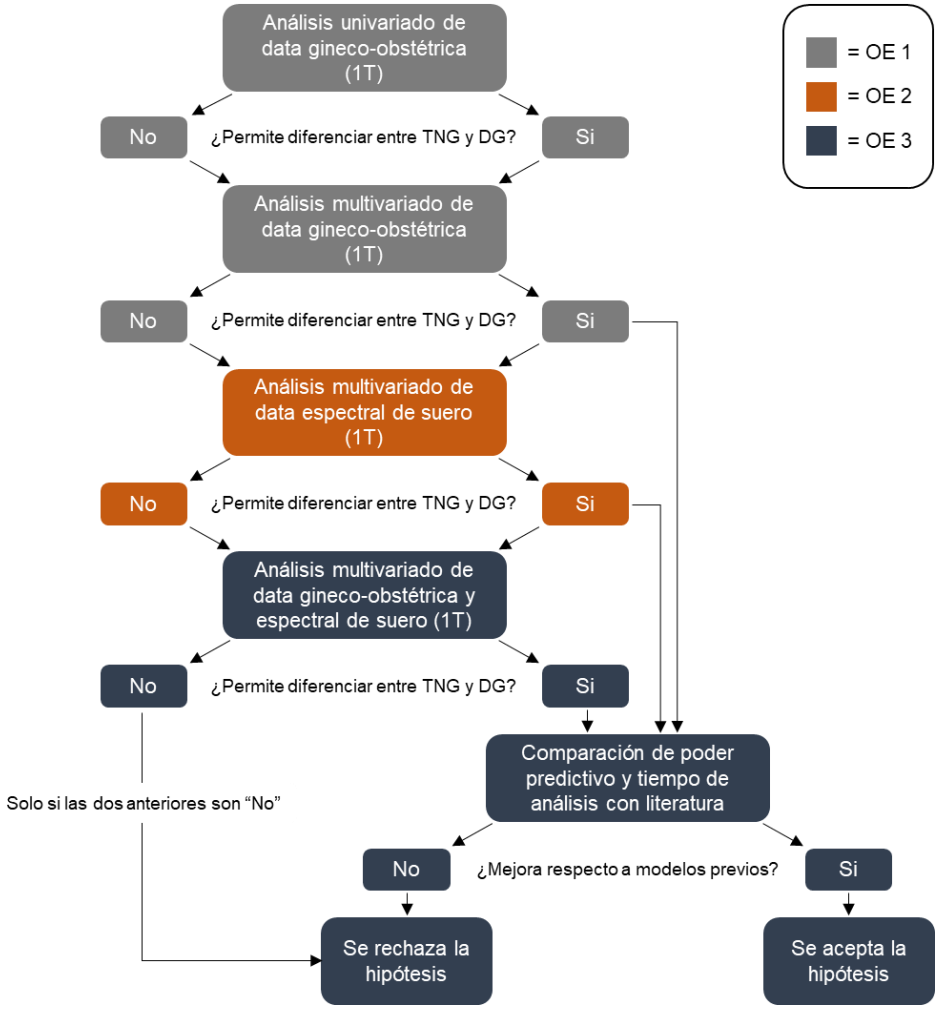


Figura 3-1. Estrategia analítica. *OE: objetivo específico; 1T: primer trimestre de gestación; TNG: tolerancia normal a la glucosa; DG: diabetes gestacional.*

4. RESULTADOS Y DISCUSIÓN

El presente capítulo expone los artículos científicos que son producto de esta tesis doctoral. Este se divide en cuatro secciones, cada una correspondiente a un manuscrito.

Las dos primeras secciones de este capítulo (4.1 y 4.2) corresponden a dos revisiones bibliográficas, que permitieron profundizar en el estado del arte y establecer los umbrales de poder predictivo y tiempo de análisis necesarios para aceptar o rechazar la hipótesis planteada. La tercera sección (4.3) es una publicación que involucra el trabajo realizado para cumplir el primer objetivo específico propuesto. La cuarta y última sección de este capítulo (4.4) es un artículo científico que comprende el trabajo realizado para cumplir el segundo y el tercer objetivo específico propuesto.

Los cuatro manuscritos presentados en este capítulo exhiben una estructura interna similar. En términos generales, esta consta de los siguientes apartados: resumen, introducción, metodología, resultados, discusión, conclusión, referencias y material suplementario.

4.1. Machine learning applied in maternal and fetal health: A narrative review focused on pregnancy diseases and complications

En esta sección se presentan y discuten los resultados que dieron origen al artículo “Machine learning applied in maternal and fetal health: A narrative review focused on pregnancy diseases and complications”, publicado en la revista *Frontiers in Endocrinology* (https://doi.org/10.3389/fendo.2023.1130139).

Machine learning applied in maternal and fetal health: A narrative review focused on pregnancy diseases and complications

Daniela Mennickent[1,2,3], Andrés Rodriguez[3,4], Cecilia Opazo[5], Claudia Riedel[6], Erica Castro[7], Alma Eriz-Salinas[8], Javiera Appel-Rubio[1], Claudio Aguayo[1], Alicia Damiano[9], Enrique Guzmán-Gutiérrez[1,3,*] and Juan Araya[2,3,*]

[1] Departamento de Bioquímica Clínica e Inmunología, Facultad de Farmacia, Universidad de Concepción, Chile. [2] Departamento de Análisis Instrumental, Facultad de Farmacia, Universidad de Concepción, Chile. [3] Machine Learning Applied in Biomedicine (MLAB), Chile. [4] Departamento de Ciencias Básicas, Facultad de Ciencias Básicas, Universidad del Bio-Bio, Chillán, Chile. [5] Instituto de Ciencias Naturales, Facultad de Medicina Veterinaria y Agronomía, Universidad de Las Américas, Santiago, Chile. [6] Departamento de Ciencias Biológicas, Facultad de Ciencias de la Vida, Universidad Andrés Bello, Santiago, Chile. [7] Departamento de Obstetricia y Puericultura, Facultad de Ciencias de la Salud, Universidad de Atacama, Copiapó, Chile. [8] Departamento de Obstetricia y Puericultura, Facultad de Medicina, Universidad de Concepción, Chile. [9] Cátedra de Biología Celular y Molecular, Departamento de Ciencias Biológicas, Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Argentina.

*** Correspondence:** eguzman@udec.cl (EG-G); jarayaq@udec.cl (JA)

Abstract

Introduction: Machine learning (ML) corresponds to a wide variety of methods that use mathematics, statistics and computational science to learn from multiple variables simultaneously. By means of pattern recognition, ML methods are able to find hidden correlations and accomplish accurate predictions regarding different conditions. ML has been successfully used to solve varied problems in different areas of science, such as psychology, economics, biology and chemistry. Therefore, we wondered how far it has penetrated into the field of obstetrics and gynecology. **Aim:** To describe the state of art regarding the use of ML in the context of pregnancy diseases and complications. **Methodology:** Publications were searched in PubMed, Web of Science and Google Scholar. Seven subjects of interest were considered: gestational diabetes mellitus, preeclampsia, perinatal death, spontaneous abortion, preterm birth, cesarean section, and fetal malformations. **Current state:** ML has been widely applied in all the included subjects. Its uses are varied, the most common being the prediction of perinatal disorders. Other ML applications include (but are not restricted to) biomarker discovery, risk estimation, correlation assessment, pharmacological treatment prediction, drug screening, data acquisition and data extraction. Most of the reviewed articles were published in the last five years. The most employed ML methods in the field are non-linear. Except for logistic regression, linear methods are rarely used. **Future challenges:** To improve data recording, storage and

update in medical nd research settings from different realities. To develop more accurate and understandable ML models using data from cutting-edge instruments. To carry out validation and impact analysis studies of currently existing high-accuracy ML models. **Conclusion:** The use of ML in pregnancy diseases and complications is quite recent, and has increased over the last few years. The applications are varied and point not only to the diagnosis, but also to the management, treatment, and pathophysiological understanding of perinatal alterations. Facing the challenges that come with working with different types of data, the handling of increasingly large amounts of information, the development of emerging technologies, and the need of translational studies, it is expected that the use of ML continue growing in the field of obstetrics and gynecology.

4.1.1. Introduction

Pregnancy is a physiological process that provides all conditions for normal fetus growth and subsequent birth. Due to certain circumstances, a seemingly normal pregnant woman starts with physiological disorders that can trigger pregnancy diseases (e.g. gestational diabetes mellitus and preeclampsia) or other perinatal complications (e.g. stillbirth, cesarian section, macrosomia and respiratory distress). The search for new strategies for early diagnosis, screening and risk determination could reduce the severity of these alterations and also the negative impact in both mother's and offspring's health. Interestingly, in recent years, machine learning (ML) has been used to find solutions for these problems.

ML corresponds to a wide variety of methods that use mathematics, statistics and computational science to learn from multivariate data. By means of pattern recognition performed on various measured variables, different algorithms are able to find correlations, often hidden to the human eye, and perform accurate predictions about different conditions, such as the belonging of an individual to a certain group or class, or the concentration of a particular biomarker in a sample of interest.

Multivariate methods (i.e. those employed to analyze the behavior of multiple variables simultaneously) have been used for several decades to solve problems

in different areas of knowledge, such as psychology, economics, biology, chemistry, etc. However, in the clinical field these tools have begun to penetrate only recently. Remarkably, the use of these tools has received different names throughout history depending on the area of application, i.e. psychometrics in psychology, biometrics in biology, chemometrics in chemistry, etc. In the last years it has become popular to address to these methods as artificial intelligence, ML, data mining, or in a more general sense, data science. The boundaries between the scopes of these different terms are still a subject of debate, and several different opinions and definitions can be found in specialized literature [1,2]. However, unconcerned of this debate, it seems that ML has been the preferred name used in healthcare-related studies, therefore that will be the term used in this manuscript.

One of the most common applications of ML in biomedicine is the detection or prediction of particular pathological conditions [3]. It seems logical that in pregnancy the focus has also been in diagnostics [4,5]. However, as it has been evidenced in different disciplines, ML can also be used for other purposes, such as identification of important variables in a system or process, correlation analysis, data management and extraction, noise removal, dimensionality reduction, among others [6,7]. Given the success ML has had in other areas of science, we wondered how far it has penetrated into the field of obstetrics and gynecology. In this review we propose to describe the state of art regarding the use of ML in the

context of pregnancy diseases and complications, including its capability for early diagnosis, screening and risk determination, and also other applications of this versatile tool.

4.1.2. Methodology

4.1.2.1. Type of study and search strategy

This is a narrative review. Publications regarding the use of ML in maternal and fetal health were searched in different databases, including PubMed, Web of Science and Google Scholar. Seven subjects of interest were considered as representative conditions of the vast domain of obstetrics and gynecology, due to their prevalence and clinical relevance [8]: gestational diabetes mellitus, preeclampsia, perinatal death, spontaneous abortion, preterm birth, cesarean section, and fetal malformations.

4.1.2.2. Information synthesis

The papers main results were summarized in tables, comprising input, ML technique and output. Tables should be understood as follows: each table is associated with a specific pregnancy disease or complication, as stated on the table's title. For every table, each row refers to a particular study. For each

reference (first column), the input, the ML technique and the output (third, fourth and fifth columns, respectively) are directly linked to the ML application (second column) of that study. Most of the tabulated information is written and further extended in the text related to each table.

4.1.2.3. Manuscript organization

This manuscript is organized as follows: section 4.1.3 gives a general overview on ML-related definitions and concepts, section 4.1.4 describes different ML applications in the context of pregnancy diseases and complications, addressed from the highest to the lowest prevalence, section 4.1.5 discusses the current state and future challenges in the field, and section 4.1.6 rounds off with a brief conclusion.

4.1.3. ML: definitions and concepts

ML models can have varied purposes. The most typical one is early detection, but they can also be used for alternative screening, risk estimation, correlation assessment, biomarker discovery, among other possible applications.

In very simple terms, the development of a ML model requires three main parts: the input, the ML technique and the output.

The input is the data that is used to build the ML model. It consists of samples (usually in the biomedical field, the subjects) and variables, which can be very diverse. There is discrete data, e.g. the information retrieved from questionaries; the clinical and biochemical data found in physical and electronic health records (EHR); and the metabolites, peptides/proteins, transcripts or genes identified as relevant in omics studies. Likewise, there is continuous data, e.g. the traces obtained by Doppler ultrasonography, electrohysterography (EHG) or cardiotocography (CTG); and the images recorded by ultrasonography, computed tomography (CT) or echocardiography. The type of data determines what kind of pretreatment has to be performed prior to ML analysis, an aspect that is described in detail elsewhere [9,10].

The selection of the ML technique depends on the purpose of the study. Non-supervised techniques are used to explore the data, i.e. to assess if there is any spontaneous clustering or correlation between samples and/or variables. Typical examples of non-supervised techniques are principal component analysis (PCA) and K-means. In contrast, supervised techniques are used to predict a property. In the ML field, the word “prediction” refers to the forecast of future behaviors or unobserved outcomes [11]. In particular, classification ML techniques allow to predict a class or category, e.g. healthy or diseased; whereas regression ML techniques allow to predict a continuous quantity, e.g. the concentration of a specific biomarker. Moreover, supervised techniques can be linear or non-linear,

depending on the nature of the mathematical function that underlies the classification or regression task. The most common linear classifiers are logistic regression (LR) and linear discriminant analysis (LDA), while some examples of linear regression techniques are linear regression and partial least squares (PLS). On the other hand, random forest (RF), support vector machines (SVM) and neural networks (NN) are classical examples of non-linear ML techniques that allow to perform both classification and regression analyses.

The output is the result of having applied the ML model. The most common outputs are those that account for the model predictive performance. In classification, the performance is typically expressed using parameters such as sensitivity, specificity, accuracy and area under the receiver operating characteristic curve (AUC). In regression, other parameters, such as mean absolute error and root mean squared error (RMSE), are used. These and other performance metrics are well described in literature [12,13]. It is important to mention that the aforementioned metrics can be calculated in different stages of the model's development: training, internal validation and external validation. The ideal situation is that the model is tested in all the three stages, to ensure it will be accurate and useful in different populations. This idea has been discussed in greater depth by other authors [14,15]. Another very common output is variable importance. This information allows to identify the variables that contribute the most to predict the property under study, which is useful to identify new

biomarkers for a certain condition. There are other possible outputs, depending on the ML application. They are addressed and discussed throughout section 4.1.4.

4.1.4. ML in pregnancy diseases and complications: applications

Pregnancy diseases

4.1.4.1. Gestational Diabetes Mellitus

The American Diabetes Association defines gestational diabetes mellitus (GDM) as a “diabetes diagnosed in the second or third trimester of pregnancy that was not clearly overt diabetes prior to gestation” [16]. This disease has been related to several negative outcomes on maternal and fetal health. In the short-term, it increases the risk of pre-eclampsia, preterm delivery, macrosomia and clinical neonatal hypoglycemia; and in the long-term, of maternal prediabetes, maternal diabetes, offspring obesity and offspring impaired fasting glucose [17].

ML has been applied in GDM research, for diverse purposes (**Table 4.1-1**).

4.1.4.1.1. ML for GDM prediction

Numerous studies that have applied ML in the context of GDM, have used it to predict this disease at early stages of pregnancy [18]. Some of them have based their predictive models on a small number of variables. For example, Xiong et al. assessed hepatic, renal and coagulation function biochemical data to predict GDM at 10-19 gestational weeks [19]. Univariate analysis showed that coagulation parameters differed between GDM and control women, so they combined two of them, patient prothrombin time and reference activated partial thromboplastin time, to build different ML predictive models. They achieved AUCs of 99.83% and 99.74% by light gradient boosting and SVM, respectively. Likewise, Zheng et al. used known GDM clinical and biochemical risk factors to predict it at 8-20 gestational weeks [20]. By Bayesian adaptive sampling, they selected four maternal variables, maternal age, pre-pregnancy body mass index (BMI), fasting plasma glucose and triglycerides, and then used them to generate a multivariate Bayesian logistic regression model. They got an AUC of 0.766. In contrast, some articles have based their predictive models on a large number of variables. For instance, Wu et al. assessed 73 maternal clinical and biochemical variables and different ML techniques for GDM prediction before 12 gestational weeks [21]. Their deep neural networks (DNN) model achieved an AUC of 0.80. Furthermore, they built a simpler model in order to facilitate clinical application. By using seven sequential feature selection chosen variables and LR they got an

AUC of 0.77. Similarly, Artzi et al. used 2355 variables from EHR and gradient boosting (GB) to predict GDM before 20 gestational weeks, and obtained an AUC of 0.85 [22]. They also built a simpler model to ease clinical implementation. Their nine questions based model yielded an AUC of 0.80. Interestingly, both the full and the simplified models outperformed a baseline score, which involved seven GDM known risk factors and got an AUC of 0.68.

It is worth mentioning that some papers that have sought GDM prediction, have also revealed GDM novel risk factors. That is the case of Artzi et al. study, in which the most important predictor of their full model was the prior pregnancy glucose challenge test result, a previously unreported risk factor for GDM [22]. Likewise, Balani et al. used clinical data and different ML techniques to predict GDM in obese pregnant women at 14-17 gestational weeks [23]. Their RF model achieved an accuracy of 77.53%, and showed that the most relevant predictor was visceral fat mass, a previously unknown risk factor for GDM.

In addition, it is interesting to notice that all the aforementioned studies reported models that allow to predict GDM, but that are restricted to do so under a particular diagnostic criteria. Recently, Mennickent et al. reported a novel strategy that overcomes that limitation [24]. The authors used first trimester clinical and biochemical data and PLS to predict the very post load glycemia value that pregnant women would have at 24-28 gestational weeks. Since the predicted

value can be interpreted as control or GDM with any diagnostic criteria, the prediction of GDM is no longer restricted to a particular criteria. Their best model allowed to predict the second trimester post load glycemia with a RMSE of 23.1 and a relative error of 20.7% in cross-validation analyses.

4.1.4.1.2. ML for GDM biomarker discovery

Several studies have applied ML to search new biological markers for GDM. This has been typically done by means of omics techniques. For example, Scott et al. used ^{1}H-NMR metabolomics and 14-27 gestational weeks urine samples to find novel biomarkers for GDM [25]. Their statistically significant metabolites, identified through variable importance analysis based on random variable combination, were tested for classification by orthogonal partial least squares discriminant analysis, and achieved an AUC of 0.803. The top three metabolic markers for that model were formic acid, dimethylamine and galactose, which were downregulated in GDM. Similarly, Yoffe et al. applied a targeted transcriptomics approach and 9-11 gestational weeks plasma samples to identify miRNAs that could serve as early biomarkers for GDM [26]. Based on multiplex expression assays and RT-qPCR data and DESeq2 analyses, they found two differentially expressed miRNAs, miR-223 and miR-23a, which were upregulated in GDM. These miRNA markers were combined and assessed for classification by LR, and reached an AUC of 0.91. Another case is the study of Guo et al., who used a genomics strategy and 18 or

less gestational weeks plasma samples to find cfDNA biosignatures that could be useful for GDM detection at early stages of pregnancy [27]. Based on whole-genome sequencing and qPCR promoter profiling data, they identified 800 differentially expressed genes between GDM and control women. Eleven of those genes, CC2D2B, NAT10, SIPA1, ZNF565, ZNF552, WDR35, MICALL1, CTNNB1, CLOCK, BCKDHB and TGIF2LY, were selected by a step-wise feature selection method, and then combined and tested for classification by LR. The eleven marker based model yielded an overall accuracy of 72.1%. Likewise, Liu et al. applied an epigenomics approach to identify CpG markers for GDM [28]. They used DNA methylation data from two previous studies, in which placenta samples from GDM and control mothers, and blood samples from children born in GDM and control pregnancies were analyzed. By an overlapped CpGassoc epigenome-wide association study they identified nine differentially methylated CpGs between GDM and control subjects. The LR model built with five of them revealed that the most important CpGs for GDM and control samples differentiation were cg11169102, cg21179618 and cg21620107. The combination of those three biomarkers was assessed for classification by the same ML technique, and achieved an AUC of 0.8519.

4.1.4.1.3. Other ML applications in GDM research

Some GDM studies have used ML for other purposes, such as risk estimation, screening, correlation assessment and management. For instance, Ehrlich et al. aimed to evaluate the effect of exercise during the first trimester of pregnancy on the risk of GDM [29]. Data from a pregnancy physical activity questionnaire, effected at 10-13 gestational weeks, were analyzed by different ML techniques. Their targeted maximum likelihood estimation (TMLE) and SuperLearner (SL) method with extra learners model showed that meeting or exceeding the cohort's 75th percentile of moderate to vigorous intensity exercise reduced the risk of GDM by 2.1 fewer cases per 100 women. Another example is Bernardes-Oliveira et al. study. They intended to develop a fast and low-cost screening tool for GDM, using 9-39 gestational weeks plasma samples, attenuated total reflection Fourier-transform infrared spectroscopy and ML techniques [30]. Their genetic algorithm with LDA model, which comprised ten wavenumbers mainly from lipids and proteins spectral regions, achieved an accuracy of 100%. A different case is the study of Araya et al., who meant to determine whether there was a correlation between the maternal thyroid profile and GDM [31]. Using clinical and biochemical data registered at 10-14 and 24-28 gestational weeks, and PCA, they demonstrated that maternal thyroid-related hormones from the first and the second trimesters of pregnancy were strongly correlated with GDM. Finally, Velardo et al. aimed to develop a ML tool capable to improve the timeliness of

GDM management [32]. They used mobile health real-time collected data and different ML techniques to automatically evaluate the switch from diet-based management to pharmacological treatment. Data included blood glucose levels measured at different time points, maternal age, BMI and other GDM clinical risk factors. Their lasso feature selection LR model allowed to predict the timing of initiation of pharmacotherapy with an AUC of 0.8.

Table 4.1-1. ML applications in GDM research.

Reference	ML application	Input	ML technique	Main output
[19]	Disease prediction	Biochemical markers	Light GB	AUC = 99.83%, Se = 92.5% and Sp = 99.2%
[20]	Disease prediction	Clinical and biochemical factors	Bayesian LR	AUC = 0.766, Ac = 0.64, Se = 0.66 and Sp = 0.75
[21]	Disease prediction	Electronic health records	DNN	AUC = 0.80, Se = 63%, Sp = 82% and YI = 0.45
[22]	Disease prediction	Electronic health records	GB	AUC = 0.850 and AUPR = 0.324
[23]	Disease prediction	Clinical parameters	RF	Ac = 77.53%
[24]	Disease prediction, independent of diagnostic criteria	Clinical and biochemical factors	PLS	RMSE = 23.1, RE = 20.7% and r = 0.259 for post load glycemia prediction
[25]	Biomarker discovery	Metabolomics data	OPLS-DA	Formic acid, dimethylamine and galactose as novel biomarkers
[26]	Biomarker discovery	Transcriptomics data	LR	miR-223 and miR-23a as novel biomarkers
[27]	Biomarker discovery	Genomics data	LR	CC2D2B, NAT10, SIPA1, ZNF565, ZNF552, WDR35, MICALL1, CTNNB1, CLOCK, BCKDHB and TGIF2LY as novel biomarkers

[28]	Biomarker discovery	Epigenomics data	LR	cg11169102, cg21179618 and cg21620107 as novel biomarkers
[29]	Risk estimation	Physical activity questionnaire data	SL-EL	2.1 fewer cases of GDM per 100 women for moderate to vigorous intensity exercise
[30]	Disease screening	Spectrochemical data	LDA	Ac = 100%, Se = 100% and Sp = 100%
[31]	Correlation assessment	Clinical and biochemical factors	PCA	Strong correlation between maternal thyroid profile and GDM
[32]	Pharmacological treatment prediction	Mobile real-time collected data	LR	AUC = 0.8

ML: machine learning; GDM: gestational diabetes mellitus; GB: gradient boosting; LR: logistic regression; DNN: deep neural networks; RF: random forest; PLS: partial least squares; OPLS-DA: orthogonal PLS discriminant analysis; SL-EL: SuperLearner with extra learners; LDA: linear discriminant analysis; PCA: principal component analysis; AUC: area under the receiver operating characteristic curve; Se: sensitivity; Sp: specificity; Ac: accuracy; YI: Youden index; AUPR: area under the precision-recall curve; RMSE: root mean squared error; RE: relative error.

4.1.4.2. Preeclampsia

Preeclampsia (PE) is a pregnancy syndrome that presents two different clinical scenarios, both characterized by the development of maternal hypertension from the 20th week of gestation, an alteration that persists throughout pregnancy. The first of the scenarios is characterized by a moderate form of PE, which symptoms become evident late, from 34 weeks of gestation. It is characterized by blood pressure ≥140/90 mmHg, and other symptoms that indicate liver or renal damage, thrombocytopenia or proteinuria ≥3g/24h, and by not inducing alterations on fetal growth. The second of the scenarios correspond to a severe form of PE, which

symptoms become evident before 34 weeks of gestation. It is characterized by blood pressure ≥160/110 mmHg, multisystemic damage and/or proteinuria ≥5g/24h, and for being generally associated with intrauterine growth retardation (IUGR) [33]. These conditions can also lead to more serious situations than PE alone, such as HELLP syndrome and eclampsia, which is a severe form of PE accompanied by seizures [34]. Severe forms of PE are associated with at least two times the risk of IUGR, and fetal and neonatal death [35]. The origin of PE is still unknown, however, the most accepted hypothesis indicates that the placenta does not form properly. The latter would not allow a correct flow of maternal blood towards the placenta, triggering a compensatory response that would increase blood pressure to meet the metabolic requirements of the fetus in gestation. This process would begin during the first trimester of pregnancy, producing serious effects on the mother, and affecting the fetus during the second and third trimesters of pregnancy [36]. Thus, the early detection of PE, i.e. before the appearance of adverse symptoms in the mother, is necessary.

The early detection of PE has been assessed through the determination of the levels of human chorionic gonadotropin [37], anti-Müllerian hormone [38], sFlt-1 [39], the soluble form of Endoglin [40], among others, with sensitivities between 20 and 80%, and specificities between 40 and 90%. Interestingly, algorithms mediated by ML have been proposed as new strategies to predict this pathology earlier (**Table 4.1-2**). Various ML models have been developed for PE prediction

using different types of variables, such as metabolites [41], proteins [42], plasma DNA [27] and circular RNA [43], but by far the most common approaches are based on maternal medical data [44–51].

Some ML-based studies have aimed to predict PE before 20 weeks of pregnancy. For example, Marić et al. used clinical and biochemical maternal data and different ML techniques to predict this pregnancy complication before 16 gestational weeks. Their elastic net (EN) model achieved an AUC of 0.79 for all cases of PE, and an AUC of 0.89 for early-onset PE, showing that ML approaches can become a powerful early prediction tool for this obstetric disorder [48]. Sandstrom et al. also used clinical and biochemical maternal variables and different ML techniques to predict PE, but before 15 weeks of gestation. Their LR model with 12 pre-specified variables yielded AUCs of 0.68, 0.68 and 0.67 for PE with delivery <34, <37 and ≥37 weeks of pregnancy, respectively [46]. A different example is the study of Gupta et al., who aimed to predict hypertensive disorders of pregnancy, including PE, with placenta ultrasound images from the first trimester of gestation. The analysis of abnormal placental image texture with deep convolutional neural networks (CNN) achieved a sensitivity of 70.6% and a specificity of 76.6% [51]. In contrast to the aforementioned articles, other ML-based studies have intended to predict PE from 20 weeks of gestation onwards. For instance, Han et al. measured 25 parameters of maternal clinical chemistry before PE clinical diagnosis, and combined them to predict this pregnancy disorder. Their back-propagation neural

networks (BPNN) model, which strongest predictors were ALB, MPV, BUN, LDH and TG, displayed an accuracy of 79.8% [47]. Likewise, Jhee et al. retrieved maternal data (collected between 14 and 34 weeks of pregnancy) from EHR and tested them to predict late-onset PE. Their ML models, based on decision trees (DT), naïve Bayes (NB), SVM, RF, stochastic GB, and LR reached AUCs of 0.857, 0.776, 0.573, 0.894, 0.924 and 0.806, respectively [45].

Other PE-related studies have applied ML in additional contexts, such as biomarker identification, risk estimation and drug screening. For example, Liu et al. analyzed microarray data to identify hub genes as diagnostic biomarkers of PE. Their bioinformatics approach revealed 17 differentially expressed hub genes between PE and control subjects: IL7R, IL18, CCL2, HLA-DRA, CD247, ITK, CD2, IRF8, CD48, GZMK, CCR7, HLA-DPA1, LEP, IL1B, CD8A, CD3D and GZMA. Those hub genes were combined and assessed for classification by SVM. Their model reached an AUC of 0.958 in the training set, and an AUC of 0.834 in the test set [52]. Similarly, Guo et al. screened placental mRNA data to identify PE biomarkers. Their ML-based approach allowed them to select a subset of 13 mRNA features: HTRA4, PROCR, MYCN, ERO1A, EAF1, PPP1R16B, CRH, FLNB, PIK3CB, PLAAT3, FBN2, RFLNB, and TKT, which were combined and tested for PE and control subjects classification by ML. Their model, which fused three ML classifiers, C4.5, AdaBoost and multilayer perceptron, yielded an accuracy of 82.2% [53]. A different case is the study of Bodnar et al., who aimed

to assess the effect of fruit and vegetable intake and dietary synergy on the risk of various adverse pregnancy outcomes. Their SL with TMLE ML model revealed that high fruit and vegetable densities were associated with 3.2 and 4.0 fewer cases of PE per 100 births, respectively [54]. A final example is the article of Tejera et al., who developed a ML-based strategy to identify currently existing drugs that could be repurposed for PE management. Their approach was built on pharmacological targets of drugs under clinical trial for PE, and was designed to exclude those that have shown negative effects in pregnancy. Their ML-based virtual screening identified estradiol, estriol, vitamins E and D, lynestrenol, mifrepristone, simvastatin, ambroxol, and some antibiotics and antiparasitics as potential drugs for PE treatment [55].

Table 4.1-2. ML applications in PE research.

Reference	ML application	Input	ML technique	Main output
[41]	Disease prediction	Metabolomics data	LR	AUC = 0.868, Se = 75.1% and Sp = 83.0%
[42]	Disease prediction	Proteomics data	LDA	AUC = 0.96, Se = 0.90 and Sp = 0.90 for early-onset cases with maternal vascular malperfusion
[27]	Disease prediction	Genomics data	LR	AUC = 0.825, Ac = 83.0%, Se = 81.7% and Sp = 83.3%
[43]	Disease prediction	Transcriptomics data and biochemical markers	LR	AUC = 0.940, Se = 86.67% and Sp = 96.67%
[47]	Disease prediction	Biochemical markers	BPNN	Ac = 79.8%
[45]	Disease prediction	Electronic health records	Stochastic GB	AUC = 0.924, Ac = 0.973, Se = 0.603, Sp = 0.991 and DR = 0.771 for late-onset cases

[48]	Disease prediction	Electronic health records	EN	AUC = 0.89, Se = 72.3% and Sp = 91.2% for early-onset cases
[44]	Disease prediction	Clinical and biochemical factors	LR	AUC = 0.962, Se = 79.3%, Sp = 97.7%, PPV = 92% and NPV = 93.4%
[46]	Disease prediction	Clinical and biochemical factors	LR	AUC = 0.68, Se = 30.6% and Sp = 90% for early-onset cases
[49]	Disease prediction	Clinical and biochemical factors	RF	AUC = 0.976, AUPR = 0.958, Ac = 92.6%, Se = 91% and Sp = 93% for placental dysfunction-related disorders
[50]	Disease prediction	Clinical parameters	RF	AUC = 0.90, Se = 0.70, Sp = 0.89 and Pr = 0.88
[51]	Disease prediction	Ultrasound images	CNN	Se = 70.6% and Sp = 76.6% for hypertension disorders of pregnancy
[52]	Biomarker discovery	Genomics data	SVM	IL7R, IL18, CCL2, HLA-DRA, CD247, ITK, CD2, IRF8, CD48, GZMK, CCR7, HLA-DPA1, LEP, IL1B, CD8A, CD3D and GZMA as novel biomarkers
[53]	Biomarker discovery	Transcriptomics data	C4.5, AB and MLP	HTRA4, PROCR, MYCN, ERO1A, EAF1, PPP1R16B, CRH, FLNB, PIK3CB, PLAAT3, FBN2, RFLNB, and TKT as novel biomarkers
[54]	Risk estimation	Food frequency questionnaire data	SL	3.2 and 4.0 fewer cases of PE per 100 births for high density fruit and vegetable intake
[55]	Drug screening	Drug databases information	TPOTC	Estradiol, estriol, vitamins E and D, lynestrenol, mifrepristone, simvastatin, ambroxol, and some antibiotics and antiparasitics as potential drugs for PE

ML: machine learning; PE: preeclampsia; LR: logistic regression; LDA: linear discriminant analysis; BPNN: back-propagation neural networks; GB: gradient boosting; EN: elastic net; RF: random forest; CNN: convolutional neural networks; SVM: support vector machines; AB: adaptative boosting; MLP: multilayer perceptron; SL: SuperLearner; TPOTC: tree-based pipeline optimization tool classifier; AUC: area under the receiver operating characteristic curve; Se: sensitivity; Sp: specificity; Ac: accuracy; DR: detection rate; PPV: positive predictive value; NPV: negative predictive value; AUPR: area under the precision-recall curve; Pr: precision.

4.1.4.3. Perinatal death

The World Health Organization (WHO) defines perinatal deaths as those that occur from late stillbirth, i.e. after 28 weeks of gestation, up to 28 days of extra-uterine life, including late neonatal deaths [56]. Worldwide, more than 5 million perinatal deaths happen every year [57]. Progress in reducing the high numbers of stillbirths and neonatal deaths has been slow. Even though the rate of perinatal deaths has been lowered in developed countries, its reduction in low- and middle-income countries has been insufficient. Indeed, low- and middle-income countries present the highest rates and the slowest reduction [58,59]. The Sustainable Development Goals set by the United Nations General Assembly include to put an end to the avoidable deaths of newborns by 2030 [60], however, during 2019 there were approximately 7000 newborns deaths each day [61]. These numbers highlight the necessity to implement new methods and techniques to identify high-risk pregnancies, early enough to be able to provide them personalized attention so as to improve prevention, or reduce risk and perinatal death.

4.1.4.3.1. Stillbirth

Studies to predict pregnancies with high risk of perinatal death have been difficult due to small sample size [62]. This, along with the difficulty posed by a relatively high percentage of missing data, forces researchers to look for strategies to impute missing data or lose variables to avoid biased results [63]. Routinely collected perinatal records have a great potential to improve the risk assessment of perinatal death, by providing massive databases that are available for researchers to develop and test ML-based models (**Table 4.1-3**). These records are commonly composed of maternal demographic and medical history information, which can be used as predictors. The high amount of data available in these records also allows to have appropriate validation sets to assess the quality of the prediction. Koivu et al. used publicly available data obtained from the US Centers for Disease Control and Prevention, to build ML-based risk prediction models for early stillbirth, late stillbirth and preterm birth (PTB) pregnancies [64]. Using only maternal demographic and medical history data (pregnancy and sexual transmitted diseases) from almost 16 million pregnancies, of which 92,753 were infant deaths, they achieved AUCs of 0.76 for early stillbirth, 0.63 for late stillbirth and 0.64 for PTB. Those results were obtained using an algorithm based on self-normalizing neural networks. An important highlight of this study is that model validation was performed using an external set from a different population, which is the strictest and most reliable type of validation, often

resulting in lower performances compared to other more permissive validation methods (such as resampling methods), which are prone to overfitting.

Using a similar approach, Malacova et al. developed stillbirth risk prediction models using different ML algorithms [62]. The study population was a cohort from Western Australia, consisting in almost 1 million births, of which 5,788 were stillbirths. The variables used to build the models were a combination of maternal socio-demographic characteristics, medical history, congenital anomalies and, more importantly, current pregnancy complications, which helped to achieve the greatest sensitivity. Different models were built, since not all subjects had the same amount of information available. For all models AUC varied from 0.59 to 0.84, which suggests the importance of variable selection to achieve better performances. The best results were obtained using XGBoost, resulting in a correct prediction of 45% of all stillbirths.

Shukla et al. also performed ML-based predictive modelling for perinatal mortality, but in a wider population, a cohort of near half million pregnancies in low- and middle-income countries located in South Asia, Africa and Central America [65]. They developed different models using prenatal and post-delivery variables up to two days after birth, to predict outcomes from intrapartum stillbirth and neonatal death at different time frames. The variables used included maternal, socio-demographic, and medical information along with delivery and neonatal variables

(the last two for neonatal death prediction only). They observed that the prediction of perinatal deaths using just prenatal and predelivery information reached AUC values of 0.72 or less, and that the predictive accuracy of the model improved as more post-delivery variables were included. Indeed, their best results were obtained with post-delivery data, which allowed to predict neonatal deaths with an AUC value of 0.87 by LR.

Mboya et al. studied a cohort of 42,319 singleton deliveries in Tanzanian population [66] and build ML models to predict both stillbirth and neonatal death (defined as death of live births within 7 days of life) using data available in the birth registry, i.e. mainly sociodemographic characteristics. The best results were achieved using RF, NB and Boosting with an AUC of 0.79. Khatibi et al., used a two-step ensemble classifier ML-based method (including DT, GB, LR, RF and SVM) to predict both stillbirth before delivery and stillbirth during labor occurred in Iran in a population of almost 1,5 million births [67]. They used a combination of maternal socio-demographic features, labor descriptors, delivery properties and clinical history of the mother and fetus, and achieved an average AUC of 0.9. Although this value is much higher than the previously discussed studies, the aim of the authors was not early prediction, but to predict stillbirth at labor-delivery instead, therefore they used variables that are not available in early prediction studies.

A common result in these studies is that gestational age and fetal height are the two most important features to discriminate livebirth from stillbirth [65–67]. Some authors suggest that risk prediction models that only use demographic and medical history could be further improved with the addition of biochemical and/or biophysical variables, however to the date these approaches are yet to be explored.

4.1.4.3.2. Neonatal Death

Regarding neonatal death prediction, a recent work published in early 2021 made a systematic review on ML models used to predict neonatal mortality [68]. They focused on works with a high amount of subjects (n>500 individuals) that analyzed both perinatal and neonatal factors, and excluded studies using exclusively antenatal factors, and in which neonatal mortality was not the primary outcome of study. They found eleven publications that met their criteria, among which the AUC value varied from 0.58 to 0.97. The most used ML methods were artificial neural networks (ANN), RF and LR, although the best overall model was obtained using LDA. Interestingly, from all studies reviewed in that work, only two conducted an external validation, which ensures a higher reliability. This fact also stresses the necessity of appropriate analytical methodologies and validations in future studies to ease their application by health care providers.

Other research groups, not covered in the aforementioned systematic review, have reported the prediction of neonatal death using ML-based models, with relatively high success (AUC of 95.99% for the best results) [69–71]. In a different study conducted in Iran, different ML-based models were built to predict neonatal deaths in neonatal intensive care units [72]. This work stands out since its models were prospectively applied and evaluated in a new cohort of neonates. Seventeen variables considered important in neonatal mortality prediction were used and different ML methods were tested, such as ANN, DT, SVM, Bayesian network and ensemble classifier. The highest AUC was achieved by the RF, SVM and ensemble models with a value of 0.98, however, when they prospectively applied the models for mortality prediction in new neonates, the best overall performance was obtained using ANN, with an AUC of 0.92, whereas the highest precision and specificity were obtained using DTs (0.97 and 0.87 respectively).

Table 4.1-3. ML applications in perinatal death research.

Reference	ML application	Input	ML technique	Main output
[62]	Complication prediction	Clinical parameters	Extreme GB	AUC = 0.842, Ac = 94.71%, Se = 45.3%, Sp = 95%, PPV = 4.81%, NPV = 99.68%, +LR = 9.03 and -LR = 0.58 for stillbirth
[63]	Complication prediction	Clinical parameters	LR	AUC = 0.82 for stillbirth
[64]	Complication prediction	Clinical parameters	SNNN	AUC = 0.76, Se = 38% and Sp = 90% for early stillbirth
[65]	Complication prediction	Clinical parameters	LR	AUC = 0.872 for neonatal death

[66]	Complication prediction	Clinical parameters	RF	AUC = 0.79, Ac = 0.87, Se = 0.54, Sp = 0.88, PPV = 0.15 and NPV = 0.98
[67]	Complication prediction	Clinical parameters	DT, GB, LR, RF and SVM	AUC = 90.00%, Ac = 90.56%, Se = 91.37%, Sp = 88.10%, Pr = 88.02% and F1 = 90.58% for stillbirth before delivery and during labor
[71]	Complication prediction	Clinical parameters	MLP	AUC = 95.99%, Ac = 96.79%, Se = 86.20%, Sp = 98.37%, RMSE = 0.1702 and RRSE = 47.47% for neonatal death
[69]	Complication prediction	Clinical parameters	SL	AUC = 0.89 and U = -0.0003 for neonatal death
[70]	Complication prediction	Clinical parameters	RF	AUC = 0.922, Ac = 0.903, Se = 0.674, Sp = 0.919, PPV = 0.377, F1 = 0.477 and mean F1 = 0.712 for neonatal death
[72]	Complication prediction	Clinical parameters	ANN	AUC = 0.92, Ac = 0.86, Se = 0.86, Sp = 0.83, Pr = 0.96 and F1 = 0.91 for neonatal death

ML: machine learning; GB: gradient boosting; LR: logistic regression; SNNN: self-normalizing neural networks; RF: random forest; DT: decision tree; SVM: support vector machines; MLP: multilayer perceptron; SL: SuperLearner; ANN: artificial neural networks; AUC: area under the receiver operating characteristic curve; Ac: accuracy; Se: sensitivity; Sp: specificity; PPV: positive predictive value; NPV: negative predictive value; +LR: positive likelihood ratio; -LR: negative likelihood ratio; Pr: precision; RMSE: root mean squared error; RRSE: root relative squared error.

4.1.4.4. Spontaneous abortion

Spontaneous abortion (SA) is defined as the loss of pregnancy before the 20th week of gestation [73]. It is often referred also as miscarriage, but according to literature, miscarriage is considered to occur before the 24th week of gestation [74]. Both situations imply a common and serious pregnancy complication that has a significant psychological impact on the mother and the family. For this

reason, and due to its complicated etiology [75], SA has become a hot topic in scientific research and gynecology.

Recent advances in technology, particularly in the artificial intelligence field, have allowed the use of the increasing amount of data that can be obtained in biomedical studies to improve patients' outcomes. This is consistent with the notion of precision medicine, that is, the need of a more personalized medicine to improve or predict the medical outcome [76], in this case, of a pregnant woman.

Interestingly, ML has been applied in the context of SA and miscarriage (**Table 4.1-4**). In 2013, Bottomley at al., developed a score based on demographic data, symptom variables and ultrasound data to predict the likelihood of a woman to have a successful pregnancy by performing a retrospective study [74]. The ML method used was LR. Interestingly, the authors found that the combination of all the factors was able to provide a more accurate prediction of pregnancy viability than the obtained by analyzing the factors in a separated way, with an AUC of 0.924. This score model worked, but at that time it was not proven if it would be able to prevent miscarriage and, as the authors pointed out, the psychological morbidity associated with pregnancy loss should be integrated to the analyses. A distinct approach was made in 2019 using next generation sequencing to analyze 200 DNA samples of 100 couples presenting recurrent miscarriages (RM) [77]. This work aimed to develop an algorithm based on the genetic analysis of the HLA

protein codifying genes, considering the relationship of the HLA antigen sharing between couples and SA [78] in the context of immune interactions as a possible cause of SA and RM. It has been described that when the mother and the father share HLA antigens, the mother and the fetus will be homozygous for several of these loci. This issue alters the mother immunologic protection to the fetus inducing immunologic rejection and consequently SA [79]. The SVM-based algorithm was able to correctly classify 67% of the total subjects, with an AUC of 0.71 and a false positive rate of 57%, which negatively affected the algorithm performance. Interestingly, this study is one of the first to predict RM probabilities in a case-by case basis, having a potential use in couple genetic counseling before conception. A different example is the study of Wu et al., who aimed to predict recurrent SA with prethrombotic state (PTS) serum biomarkers. PTS is known as one of the possible causes of SA. Wu et al. work was based on the analysis of different PTS-related proteins using multiplex array technology [80]. They were able to distinguish control and affected individuals with high accuracy and precision using IL-24, exotoxin-3 and epidermal growth factor. Indeed, their DT model got an AUC of 1.000. Despite this excellent result, the cohort used for this study needs to be incremented to evaluate the real diagnostic power of this promising model.

In vitro fertilization embryo transfer (IVF-ET) is nowadays an alternative for couples with difficulties to conceive. This procedure implies high risks of

miscarriage, being psychologically stressful for couples. Therefore, it becomes necessary to find a way or system that allows the prediction of the transfer outcome, and the early detection of possible problems [81]. Recently, Liu et al. developed a ML-based model with historic data obtained by transvaginal ultrasonography from females that underwent IVF-ET. The study only considered women with viable singleton and 6-12 weeks of pregnancy [82]. The authors were able to predict embryonic development after transfer using six different ML-classifiers, with AUCs ranging from 0.91 to 0.97 when fetal heart rate (FHR) was included among the predictors. The most accurate prediction was obtained by RF at the 10th week after embryo transfer, with an AUC of 0.99. Other example is the article of Huang et al., who used deep learning to predict pregnancy outcomes in patients with recurrent reproductive failure (RRF), including recurrent pregnancy loss (RPL) and recurrent implantation failure (RIF). The study defined RPL as two or more SA before 20 weeks of pregnancy, and RIF as couples unable to conceive after multiple IVF-ET cycles. The authors analyzed EHR data with sparse coding, and predicted four pregnancy outcomes: biochemical pregnancy, clinical pregnancy, ongoing pregnancy and live birth. They got testing accuracies that ranged between 54.2% and 89.7% for the different pregnancy outcomes. Notably, the best model for the prediction of biochemical pregnancy was obtained with a panel of 10 endometrial immunological markers, while the best models for the other three outcomes, were obtained with a panel of 15 autoantibodies. The

authors discussed that this knowledge could help clinicians to plan a more personalized diagnosis and treatment for patients with RRF [83].

Table 4.1-4. ML applications in SA research.

Reference	ML application	Input	ML technique	Main output
[74]	Complication prediction	Clinical parameters	LR	AUC = 0.924, Se = 0.922, Sp = 0.733, PPV = 84.7% and NPV = 85.4%
[77]	Complication prediction	Genomics data	SVM	AUC = 0.71, Ac = 67%, Se = 86% and Sp = 43%
[80]	Complication prediction	Proteomics data	DT	AUC = 1, Ac = 100%, Se = 100%, Sp = 100%, Kappa = 1, PPV = 1 and NPV = 1 for recurrent SA with prethrombotic state
[82]	Complication prediction	Clinical parameters	RF	AUC = 0.99, Ac = 0.99, Pr = 0.99, Re = 0.99, F1 = 0.99
[83]	Complication prediction	Electronic health records	SC	AUC = 0.909 and Ac = 89.7% for live birth

ML: machine learning; SA: spontaneous abortion; LR: logistic regression; SVM: support vector machines; DT: decision tree; RF: random forest; SC: sparse coding; AUC: area under the receiver operating characteristic curve; Se: sensitivity; Sp: specificity; PPV: positive predictive value; NPV: negative predictive value; Ac: accuracy; Pr: precision; Re: recall.

4.1.4.5. Preterm birth

The WHO defines PTB as the delivery of alive babies before 37 weeks of pregnancy are completed [56]. Based on gestational age, it can be sub-categorized as: extremely preterm, before 28 weeks; very preterm, between 28

and 32 weeks; and moderate to late preterm, between 32 and 37 weeks. Most of preterm deliveries are spontaneous, although some are provider-initiated [56].

PTB is the main cause of death in children under 5 years of age worldwide. Furthermore, it has short and long-term consequences on newborns' health, which imply a significant psychological and economic burden to families and health systems [84]. The development of PTB predictive tests could be useful to identify high risk pregnancies, which could guide the healthcare personnel to offer prophylactic interventions and make antenatal management decisions [85].

ML has already been applied to develop predictive models for PTB (**Table 4.1-5**). For instance, Khatibi et al. aimed to predict spontaneous and provider-initiated PTB with data from the Iranian Maternal and Neonatal registry, which includes information of more than 1,400,0000 deliveries and 112 features. The authors used different big data ML algorithms to classify pregnant women in two steps. In the first step, all subjects were classified into term or PTB; and in the second step, the subjects classified as PTB in the first step, were then sub-classified as spontaneous or provider-initiated. Their best model, an ensemble of DT, SVM and RF, achieved a weighted average accuracy of 81%, and an AUC of 68% [86]. Similarly, Belaghi et al. used first and second trimester information from the Ontario's Better Outcomes Registry and Network database, and different ML methods to predict overall and spontaneous PTB. The investigation considered

112,963 pregnancies. For overall cases, the best models were obtained by ANN, and reached AUCs of 60.3% and 79.8% in the validation cohort at the first and second trimester, respectively. For spontaneous cases, the best results were obtained by LR, and got validation AUCs of 59.4% and 64.5% at the first and second trimester, respectively [87]. A different approach was followed by Gao et al., who used EHR text data and deep learning ML methods to predict extreme PTB. Their dataset involved 10 years of EHR information from 25,689 deliveries at the Vanderbilt University Medical Center. The long short-term memory (LSTM) recurrent neural networks (RNN) ensemble model allowed to predict extreme PTB with an AUC of 0.744 in the validation cohort, greater than the obtained by LR, SVM and GB [88]. This is an interesting result, although this work didn't differentiate spontaneous from provider-initiated cases. Likewise, Zhang et al. aimed to predict PTB with continuous EHR data and LSTM. Their dataset included first and second trimester medical parameters from more than 25,000 pregnant women who received antenatal care and had vaginal delivery at the Hangzhou Women's Hospital. Notably, the time-series deep learning technique LSTM achieved a better predictive performance than the traditional cross-sectional ML technique XGBoost, with cross-validation AUCs of 0.651 and 0.516-0.601, respectively [89].

All the aforementioned studies based their predictive models on clinical and biochemical maternal information available in databases. However, other articles

have assessed alternative types of data to predict PTB. Such studies are very useful to find novel biomarkers for PTB, and to propose informed hypotheses about its causes and underlying mechanisms, which are not fully understood [84,85]. For instance, Aung et al. measured an extensive set of 65 urine and plasma biomarkers, and combined them with ML to predict PTB at 26 weeks of gestation. They tested three ML methods: LR, adaptive EN and RF. The best validation results were obtained with the latter. The combination of all the biomarkers with RF yielded AUCs of 0.85 and 0.79 for overall and spontaneous PTB, respectively. Then, the authors divided the biomarkers into five groups, i.e. DNA damage markers, angiogenic factors, protein damage markers, inflammatory markers and lipid damage markers. The best predictive performances were obtained with lipid damage markers and RF, with AUCs of 0.84 and 0.79 for overall and spontaneous cases, respectively. Furthermore, the study identified the enzymatic pathway that contributed the most to that prediction: the eicosanoid lipoxygenase pathway. The combination of 15 lipoxygenase metabolites with RF got AUCs of 0.83 and 0.82 for overall and spontaneous PTB, respectively [90]. Another example is the study of Chen et al., who applied untargeted LC-MS plasma metabolomics to identify metabolites that could be related to PTB, at 24-28 gestational weeks. The authors identified 17 and 16 biomarkers for overall and spontaneous cases, respectively, and tested their predictive performance with seven ML classifiers. The best results were obtained by RF, with AUCs of 0.92 and 0.89 in the testing dataset. Interestingly, most of

the identified biomarkers were fatty acids, which suggests their involvement in the pathogenesis of PTB [91]. Similarly, Jehan et al. performed an early pregnancy multiomics characterization of PTB. The authors applied untargeted transcriptomics and targeted proteomics on plasma samples, and untargeted metabolomics on urine specimens. They used a 2-step ML algorithm, in which a model was first trained for each omics dataset, and then combined into a final model. The integrated model achieved a cross-validation AUC of 0.83, higher than the obtained for the different omics datasets alone. The work also identified the features that were more associated with PTB: a proteomics inflammatory module, including IL-6, IL-1RA, G-CSF, RARRES2 and CCL3; and an urine metabolomic module, enriched for glutamine and glutamate metabolism, and valine, leucine and isoleucine biosynthesis pathways [92].

Some less common approaches have also been applied in the context of PTB prediction. For example, Despotovic et al. tested EHG recordings to predict PTB. They built ML models using k-nearest neighbors (KNN), SVM, RF, RF with synthetic minority oversampling technique (SMOTE), and RF with adaptative synthetic (ADASYN) sampling. Their RF-ADASYN model allowed to predict PTB at 22-25 weeks of pregnancy, with an accuracy of 99.23% and an AUC of 0.999 in cross-validation [93]. A different case is the work of Rawashdeh et al., who combined 19 clinical maternal parameters with ML methods to predict PTB in a high risk cohort. They developed two different strategies to analyze their data.

The first one aimed to predict whether the pregnancy would continue beyond 26 gestational weeks (the lower limit for PTB in this study) and the potential value of performing cervical cerclage to prolong the pregnancy. For this first aim, the authors tested four different classification ML methods, DT, RF, KNN and NN; solo and with SMOTE. The highest testing AUC was obtained by the KNN-SMOTE model, with a value of 1.000. The second strategy of the authors aimed to predict the timing of spontaneous delivery after cervical cerclage, an approach that wasn't assessed in any of the previously discussed articles. For this second aim, they tested five different regression ML methods, linear regression, Gaussian process, RF, K-star and locally weighted learning. The best correlation with the actual gestational age at delivery was obtained by the RF model, with a value of 0.752 in the testing dataset. Such a regression ML model could help physicians to define prophylactic interventions timely, and reduce PTB-related perinatal morbidity and mortality [94].

Table 4.1-5. ML applications in PTB research.

Reference	ML application	Input	ML technique	Main output
[86]	Complication prediction	Clinical and biochemical factors	DT, SVM and RF	AUC = 68% and Ac = 81% for spontaneous and provider-initiated cases
[87]	Complication prediction	Clinical and biochemical factors	ANN	AUC = 79.8%, Se = 62.7%, Sp = 84.6%, PPV = 23.2% and NPV = 97.0%
[88]	Complication prediction	Electronic health records	LSTM	AUC = 0.744, Se = 0.682, Sp = 0.743 and PPV = 0.028 for extreme cases

[89]	Complication prediction	Electronic health records	LSTM	AUC = 0.651, Ac = 0.739, Se = 0.407 and Sp = 0.982
[90]	Biomarker discovery	Biochemical markers	RF	PGA2, 15DO12,14-PGJ2, BCPGE2, 13,14DHK-PGF2a, RVD1, LTE4, LTB4, linolenic acid and IL-10 as novel biomarkers
[91]	Biomarker discovery	Metabolomics data	RF	FA(17:1), FA(24:6), FA(14:2), CAR(18:2), hexanoylcarnitine, FA(14:0(Ke)), FA(26:1), raffinose, PC(18:0/16:3), FA(16:3), glycocholic acid, PC(33:4), FA(22:5), FA(14:1(Ke)), heptadecanoic acid, FA(19:1) and FA(14:1) as novel biomarkers
[92]	Biomarker discovery	Metabolomics, proteomics and transcriptomics data	RF	IL-6, IL-1RA, G-CSF, RARRES2, CCL3, ANGPTL4, PAD12, TfR, and metabolites from glutamine/glutamate metabolism, and valine/leucine/isoleucine biosynthesis pathways as novel biomarkers
[93]	Complication prediction	Electrohysterography recordings	RF	AUC = 0.999, Ac = 99.23%, Se = 98.40%, Sp = 99.76% and Pr = 95.86%
[94]	Complication prediction	Clinical parameters	KNN; RF	AUC = 1.00, Ac = 0.95, Se = 0.67, Sp = 1.00, G-means = 0.82 for PTB and the potential value of performing cervical cerclage to prolong the pregnancy; MAE = 3.521, MSE = 4.560 and R = 0.752 for timing of spontaneous delivery

ML: machine learning; PTB: preterm birth; DT: decision tree; SVM: support vector machines; RF: random forest; ANN: artificial neural networks; LSTM: long short-term memory; KNN: k-nearest neighbors; AUC: area under the receiver operating characteristic curve; Ac: accuracy; Se: sensitivity; Sp: specificity; PPV: positive predictive value; NPV: negative predictive value; Pr: precision; MAE: mean absolute error; MSE: mean squared error.

4.1.4.6. Cesarean section

Cesarean section is an effective mean to solve medical and surgical complications during dystocia and severe pregnancy disorders, and has an irreplaceable role [95]. Delivery through cesarean section reduces the risk of maternal-fetal morbidity and mortality, when is medically indicated [96]. Emergency cesarean section (EMCS) can be a procedure that saves lives if pregnant women experience abnormal conditions during vaginal delivery, such as fetal suffering, eclampsia or severe preeclampsia [97]. Deciding to perform an EMCS is a complicated process, occurring only in specific obstetric conditions, and requires awareness and rapid assessment of the risk of the situation [98]. Failure to perform EMCS on time can lead to postpartum mental disorders and other severe adverse maternal and fetal outcomes [99,100]. Recognizing an acute situation during pregnancy, labor or delivery, that can be life threatening and that could require an EMCS, is considered one of the most challenging tasks in obstetrics [101].

Visual inspection of CTG traces by obstetricians and midwives is the gold standard for monitoring the wellbeing of the fetus during antenatal care [102]. One of the areas in which mathematical and computational tools for data analysis, such as ML methods, excel is in the analysis of instrumental continuous signals (**Table 4.1-6**). Several output data from instruments used in clinical diagnosis or

monitoring are composed of this type of signals, in which between any two points there can be a large amount of data points, as large as allowed by the signal resolution, or even an infinite amount in the case of analog instruments. CTG traces are a great example of this type of data in obstetrics. The problem with this type of data is that its visual interpretation is highly dependent on the observer's experience and can be strongly subjective. Most importantly, clinical decisions such as pregnancy intervention through cesarean section are made using visual inspection of CTG traces. It has been reported that the positive predictive value produced by obstetricians to anticipate negative outcomes that require cesarean section deliveries is only 30% [103]. However, although human eye may fail to provide a reliable and objective interpretation, mathematical tools for pattern recognition are not subjected to the observer's bias.

Two different articles published by the same group in Liverpool have addressed the observer variability of CTG traces using a ML approach [102,104]. The authors applied signal processing techniques to extract relevant features from CTG traces and modeled the data using different ML methods, such as DNN, LDA, RF, SVM and ensemble classifiers. They were able to classify cesarean section and vaginal deliveries from CTG traces with cross-validation AUC values of 96-99%. Other study performed by an Italian research group used a similar methodology and obtained consistent results, that is, a cross-validation AUC value of 96.7% by RF [105]. Likewise, a Chinese study that proposed a comparable strategy to classify

normal and abnormal CTG traces reported an AUC of 0.95 by CNN in cross-validation [106]. Their results demonstrate that ML methods significantly improve the prediction efficiency of necessary cesarean sections, and that their use provide a valuable decision support tool to minimize subjective interpretations of CTG traces from medical practitioners.

Besides CTG traces analysis, ML methods have been applied on EHR information to predict cesarean section and identify important variables, as well as to understand the interaction between those variables. The model developed by Clark et al. using a classification and regression tree had an AUC value of 0.7, which was considered acceptable [107]. The three features that contributed the most to that model were hospital type, maternal BMI and intrapartum oxytocin dose.

Other uses of ML have been tested in the context of cesarean sections. For example, a decision-support ML-based model for assessing intrathecal hyperbaric bupivacaine dose using physical variables during cesarean section was developed, providing the anesthesiologists a new tool that gives new insights into the potential impact of controversial parameters [108]. The least absolute shrinkage and selection operator regression model got a mean squared error of 0.0087. ML has also been applied to predict surgical site infection in cesarean section wounds, which is a leading cause of mortality and an important health

concern in low-resource countries [109]. The best model was obtained with mobile device images and LR, and achieved an AUC of 1.0. Prediction of likelihood of a successful vaginal birth after former cesarean deliveries has also been addressed using ML, which may help as a decision-making tool that could contribute to a reduction in cesarean deliveries rates [110]. The EHR-based RF model reached an AUC of 0.69, better than the obtained by DT and LR.

Table 4.1-6. ML applications in cesarean section research.

Reference	ML application	Input	ML technique	Main output
[102]	Complication prediction	Cardiotocography traces	DNN	AUC = 99%, Se = 94%, Sp = 91%, F1 = 100% and MSE = 1%
[104]	Complication prediction	Cardiotocography traces	LDA, RF and SVM	AUC = 96%, Se = 87%, Sp = 90% and MSE = 9%
[105]	Complication prediction	Cardiotocography traces	RF	AUC = 96.7%, Ac = 91.1%, Se = 90.0%, Sp = 92.2% and Pr = 92.1%
[106]	Complication prediction	Cardiotocography traces	CNN	AUC = 0.95, Ac = 94.70%, Pr = 94.71% and Re = 94.68%
[107]	Complication prediction	Electronic health records	CART	AUC = 0.7
[108]	Anesthesia dose prediction	Clinical parameters	LASSO	MSE = 0.0087 and R^2 = 0.8070
[109]	Surgical site infection prediction	Clinical parameters and mobile images	LR	AUC = 1.0, Ac = 100%, Se = 1.0 and Sp = 1.0
[110]	Later vaginal birth prediction	Electronic health records	RF	AUC = 0.69, Ac = 70.0%, Se = 97.9% and Sp = 6.9%

ML: machine learning; DNN: deep neural networks; LDA: linear discriminant analysis; RF: random forest; SVM: support vector machines; CNN: convolutional neural networks; CART: classification and regression tree; LASSO: least absolute shrinkage and selection operator; LR: logistic

regression; AUC: area under the receiver operating characteristic curve; Se: sensitivity; Sp: specificity; MSE: mean squared error; Ac: accuracy; Pr: precision; Re: recall.

4.1.4.7. Fetal Malformations

4.1.4.7.1. General Congenital Diseases

Congenital anomalies are seen in 1–3% of the population, and approximately 60–70% of the anomalies can be diagnosed via ultrasonography, while the remaining 30–40% can be diagnosed after childbirth. An e-Health android application was developed by comparing the performance of nine binary ML classification models (averaged perceptron, boosted DT, Bayes point machine, decision forest, decision jungle, locally-deep SVM, LR, NN, SVM) (**Table 4.1-7**). The models were trained with the clinical dataset of 96 pregnant women and used to predict fetal anomaly status based on maternal clinical data. The decision forest model reached the best performance, with 89.5% of accuracy, 75% of F1-Score and 95% of AUC. An external validation testing with 16 users, showed that the classification algorithm accuracy was 87.5%. This estimate is enough to give a general overview of fetal health before the patient visits the physician [111].

4.1.4.7.2. Craniosynostosis

Craniosynostosis is a congenital condition characterized by a premature fusion of the fetal cranial sutures, which induces one or more cranial bones in a fetal skull to join too early. Since this happens before the fetal brain is fully formed, as the brain grows, the skull can become deformed. Craniosynostosis is a common cause of pediatric skull deformities, affecting 1 of every 2000 to 2500 live births worldwide. This birth defect occurs in a predictable pattern because of localized fusions and the compensatory expansion of the cranial vault [112]. It is usually detected early in life, both due to its cosmetic manifestations and functional consequences, as it can result in limited brain growth, elevated intra-cranial pressure, and respiratory and visual impairment. Early diagnosis is crucial for management, prevention of complications, and consideration for early surgical correction [113]. In parallel with the growing understanding of the pathophysiology of craniosynostosis, new advances include the improvement of existing technologies such as ultrasound, and the introduction of new technologies such as ML and augmented reality [114].

Various algorithms and mathematical models have been developed to allow the computer to reliably and accurately predict specific outcomes, based on premature fusion suture input data. Using data from CT-derived measurements of cranial suture fusion, cranial deformation and curvature discrepancy, different

ML methods (RF, LDA and SVM) were tested to determine the presence or absence of craniosynostosis. The best classification performance was obtained by the LDA model, with 92.7% of sensitivity, 98.9% of specificity and the probability of correctly classifying a new subject of 95.7% [113]. In a different study, SVM and RF were used on ultrasound images in order to decrease the user error involved in the interpretation of craniosynostosis diagnostic imaging. They got a diagnostic accuracy of 88.63% and an AUC of 0.89 by SVM [115]. Finally, PCA has proven effective in differentiating between healthy controls, scaphocephalic, and trigonocephalic patients, when applied on images obtained via stereophotogrammetry [116].

4.1.4.7.3. Congenital Heart Disease

The incidence of congenital heart disease (CHD) has been estimated between 0.6% and 1.2% among live births [117]; however, it has been reported an increased incidence of 8.3% when stillborn infants of ≥26 weeks of gestation are included [118]. There could be an even higher incidence in early gestation, given spontaneous and elective pregnancy termination. A multitude of factors are associated with an increased risk of identifying CHD in the fetus, which are related to familial, maternal, or fetal conditions. The leading reason of referral for fetal cardiac evaluation is the suspicion of a structural heart abnormality on obstetric ultrasound, which results in a diagnosis of CHD in 40% to 50% of the referred

fetuses. In general, subjects with risk levels exceeding ≥2% should have a detailed fetal echocardiogram by a trained examiner.

Fetal echocardiology has evolved from the description of cardiac anatomical abnormalities toward the quantitative assessment of cardiac dimensions, shape, and function. It has been demonstrated to be useful in the diagnosis and monitoring of fetuses with a compromised cardiovascular system, which may be related to several fetal conditions, such as IUGR, twin-to-twin transfusion syndrome, and CHD [119,120]. Different ultrasound approaches are currently used to evaluate fetal cardiac structure and function, including conventional 2D imaging, and M-mode and tissue Doppler imaging, among others [121]. However, assessing fetal cardiac function is still challenging due to fetus involuntary movements, the small size of the heart, the high heart rate, the limited access to the fetus, and the lack of expertise in fetal echocardiography of some sonographers. After having obtained an optimal image, various measurements must be performed to extract relevant cardiac features related to remodeling and functional status. Therefore, the use of new technologies to improve the primary acquired images, or to help extract and standardize measurements is of great importance for optimal assessment of the fetal heart. ML techniques can help to optimize three different aspects of fetal echocardiology: acquisition, quantification and features extraction, and fetal diagnosis.

4.1.4.7.3.1. Acquisition

ML-powered methods can speed up the acquisition process, decreasing the learning curve, standardizing the resulting images and increasing data quality. In such case, standardization occurs with minimal human intervention. In this regard, Bridge et al. implemented a framework for tracking key features from healthy fetal heart ultrasound videos through RF [122]; and Yu et al. and Muduli et al. used independent component analysis along with a DT [123] and a stacked denoising autoencoder neural networks-based deep learning approach [124] to reconstruct fetal electrocardiography (ECG) signals from abdominal ECG recordings.

4.1.4.7.3.2. Quantification and Feature Extraction

The vast majority of the research in this field focuses on automatically measuring the heartbeat. Some examples are the detection of fetal cardiac activity from maternal abdomen ultrasound videos using SVM [125], the extraction of FHR features from CTG recordings applying empirical mode decomposition (EMD) [126], the extraction of FHR from fetal ECG signals employing a combination of CNN and LSTM RNN [127,128], and the detection of fetal heart beats from continuous Doppler ultrasound signals by EMD [129].

4.1.4.7.3.3. Fetal Diagnosis

One of the subfields in which ML has been extensively applied is the improvement of the diagnosis of fetal hypoxia or acidemia, based on CTG analyses. For example, Zhao et al. used CNN and got an AUC value of 97.82% for fetal acidemia caused by hypoxia [130]. There have also been some attempts to translate these methods into clinical practice via the development of software that could provide additional support in the interpretation of CTG signals and, therefore, improve the assessment of fetal status. Some examples are Infant [131], PeriCALM [132] and Foetos [133].

ML has also been assessed to improve the diagnosis of IUGR, a pathology that affects about 10% of pregnancies and that has been associated with cardiac remodeling in utero [134]. IUGR early detection models have been developed using ultrasound biometric measurements and NN [135], CTG data and SVM [136], and 2D ultrasound images and ANN [137]. Such strategies got classification accuracies of 95%, 78% and 91-94%, respectively.

Finally, ML has been recently applied to improve heart diseases prenatal diagnosis. Yeo et al. presented an intelligent ML navigation method called FINE, to automatically obtain different echocardiography anatomical views of the fetal heart and identify abnormalities within the cardiac anatomy [138]. Their method

allowed to predict CHD with a sensitivity of 98% and a specificity of 93%. Moreover, Han et al. used an artificial intelligence algorithm based on a compound network to segment echocardiography images, and then screen for fetal CHD during pregnancy. Their method achieved an accuracy of 99.0% [139]. Likewise, Arnaout et al. applied a CNN deep learning supervised method on echocardiography images to identify the five most important views of the fetal heart, to segment and to measure the cardiac structures, and to distinguish between normal hearts, tetralogy of Fallot and hypoplastic left heart syndrome [140]. The best results were obtained in the diagnosis of hypoplastic left heart syndrome, with a sensitivity and a specificity of 100% and 90%, respectively.

4.1.4.7.4. Fetal Alcohol Spectrum Disorder (FASD)

Gestational alcohol exposure is the most important known cause of neurodevelopmental disability, affecting nearly 5% of children in the US. It leads to complex epigenetic and transcriptomic modifications, which subsequently impair signaling pathways in neural and morphological development [141]. In this regard, identifying transcriptomic mechanisms that regulate alcohol's teratogenicity during embryonic development is crucial to understand different phenotypic outcomes, and may allow future therapeutic interventions that could mediate alcohol's effects. In order to understand transcriptomic changes in FASD, spanning gene, exon and splicing variants, ML approaches can be used to

corroborate traditional statistical methods, and to robust genomic functional studies. For example, Al-Shaer. applied PCA and K-means clustering on transcriptome sequencing (RNA-Seq) data. They identified 6857 differentially expressed exons, which represented 1251 gene IDs that deviated from baseline expression, and 18 miRNAs with significantly different expression profiles in response to alcohol. Several of those exons regulate focal adhesion, FoxO signaling, insulin signaling and Wnt signaling [142].

4.1.4.7.5. Macrosomia

Fetal macrosomia is diagnosed when fetal growth is beyond a specific threshold, regardless of the gestational age. In developed countries, the most used threshold is a weight above 4,000 g [143]. Macrosomia is associated with an increased risk of several maternal and newborn delivery complications, like shoulder dystocia, brachial plexus injury, asphyxia, prolonged labor, postpartum hemorrhage, and laceration of the anal sphincter [144]. Predicting macrosomia is important for making decisions about induction or cesarean delivery before the start of labor. For example, Shigemi et al. developed LR and RF ML models to predict macrosomia using maternal clinical parameters. The generated LR risk scoring system allowed to determine the association of each predictor with macrosomia, and achieved an AUC value of 0.880 [145]. Likewise, Tao et al. tested different ML techniques to predict fetal birthweight from EHR data. They considered three

categorical outcomes: small for gestational age (SGA), appropriate for gestational age (AGA) and large for gestational age (LGA). SGA was defined as birthweight lower than 2,500 g; AGA as birthweight between 2,500 and 4,000 g; and LGA as birthweight greater than 4,000 g. Remarkably, the time-series deep learning technique LSTM achieved a classification accuracy of 93.3%, outperforming the traditional cross-sectional ML techniques LR, BPNN, CNN and RF [146].

4.1.4.7.6. Teratogenicity

Teratogenicity is the most serious manifestation of iatrogenic fetal toxicity. Developing fetuses are especially sensitive to chemical exposures. Teratogens lead to fetal malformation and are implicated in lifelong physical and/or mental disabilities [147]. Teratogenicity scoring for small molecules is unsystematic, and is performed outside the clinical environment [148]. Moreover, prescribing behavior for gravid patients is based on limited human data and conflicting cases of adverse outcomes, due to the exclusion of pregnant populations from randomized controlled trials [149]. Using unsupervised t-distributed stochastic neighbor embedding and supervised GB ML methods, Challa et al. demonstrated that small molecule drug structure is a good predictor of teratogenicity. The application of such methods also allowed to discover relationships between chemical functionalities within drugs prescribable in pregnancy and existing teratogenicity information. Three chemical functionalities that predispose a drug

towards increased teratogenicity and two moieties with potentially protective effects were discovered. The ML algorithm predicted three clinically relevant classes of teratogenicity with an AUC of 0.8, and nearly double the predictive accuracy of a blind control for the same task, suggesting a successful modeling [147].

Table 4.1-7. ML applications in fetal malformations research.

Reference	ML application	Input	ML technique	Main output
[111]	Complication prediction	Mobile collected data	DF	Ac = 87.5%
[113]	Complication prediction	Computed tomography images	LDA	Ac = 95.7%, Se = 92.7% and Sp = 98.9% for craniosynostosis
[115]	Complication prediction	Ultrasound images	SVM	AUC = 0.89, Ac = 88.63%, Se = 95%, Sp = 82% and +LR = 5.25 for craniosynostosis
[116]	Complication differentiation	Stereophotogrammetry images	PCA	Clear differentiation between craniosynostosis and control patients
[122]	Data acquisition	Ultrasound videos	RF	Estimation of heart position, orientation, viewing plane and cardiac phase
[123]	Data acquisition	Electrocardiography recordings	ICA and DT	Reconstruction of fetal electrocardiogram
[124]	Data acquisition	Electrocardiography recordings	SDAE	Reconstruction of fetal electrocardiogram
[125]	Data extraction	Ultrasound videos	SVM	Detection of fetal presentation and heartbeat
[126]	Data extraction	Cardiotocography recordings	EMD	Extraction of fetal heart rate
[128]	Data extraction	Electrocardiography recordings	CNN and LSTM	Extraction of fetal heart rate
[127]	Data extraction	Electrocardiography recordings	CNN and LSTM	Extraction of fetal heart rate

[129]	Data extraction	Doppler ultrasound recordings	EMD	Extraction of fetal heart rate
[130]	Complication prediction	Cardiotocography recordings	CNN	AUC = 97.82%, Ac = 98.34%, Se = 98.22%, Sp = 94.87% and QI = 96.53% for fetal acidemia caused by hypoxia
[131]	Decision making support	Cardiotocography recordings	Infant software	Identification of fetal status
[132]	Decision making support	Cardiotocography recordings	PeriCALM software	Identification of fetal status
[133]	Decision making support	Cardiotocography recordings and ultrasound measurements	Foetos software	Identification of fetal status
[135]	Complication prediction	Ultrasound measurements	NN	Ac = 95% for intrauterine growth restriction
[136]	Complication prediction	Cardiotocography recordings	SVM	Ac = 78,26%, Se = 0.78 and Sp = 0.79 for intrauterine growth restriction
[137]	Complication prediction	Ultrasound images	ANN	Ac = 91-94% for intrauterine growth restriction
[138]	Complication prediction	Echocardiography images	FINE software	Se = 98%, Sp = 93%, +LR = 14 and -LR: 0.02 for congenital heart disease
[139]	Complication prediction	Echocardiography images	CON	Ac = 99.0%, Se = 75%, Sp = 99.6%, PPV = 99% and NPV = 88.5% for congenital heart disease
[140]	Complication prediction	Echocardiography images	CNN	Se = 100% and Sp = 90% for hypoplastic left heart syndrome
[142]	Biomarker discovery	Transcriptomics data	PCA and K-means	miR-1647, miR-3064, mirR-3533, miR-6544, miR-6590, miR-6593, miR-6602, miR-6604, miR-6639, miR-6667, miR-6706, miR-6710, miR-1650, miR-1665, miR-6542, miR-6565, miR-6619 and miR-6706 as novel biomarkers for fetal

				alcohol spectrum disorder
[145]	Complication prediction	Clinical parameters	LR	AUC = 0.880, Se = 1.00, Sp = 0.49, PPV = 0.03 and NPV = 1.00 for macrosomia
[146]	Complication prediction	Electronic health records	LSTM	Ac = 93.3% for small, appropriate and large for gestational age
[147]	Drug teratogenicity prediction	Drug databases information	t-SNE and GB	AUC = 0.8

ML: machine learning; DF: decision forest; LDA: linear discriminant analysis; SVM: support vector machines; PCA: principal component analysis; RF: random forest; ICA: independent component analysis; DT: decision tree; SDAE: stacked denoising autoencoder; EMD: empirical mode decomposition; CNN: convolutional neural networks; LSTM: long short-term memory; NN: neural networks; ANN: artificial neural networks; CON: compound network; LR: logistic regression; t-SNE: t-distributed stochastic neighbor embedding; GB: gradient boosting; Ac: accuracy; Se: sensitivity; Sp: specificity, AUC: area under the receiver operating characteristic curve; +LR: positive likelihood ratio; QI: quality index; -LR: negative likelihood ratio; PPV: positive predictive value; NPV: negative predictive value.

4.1.5. ML in pregnancy diseases and complications: current state and future challenges

4.1.5.1. Current state

ML has been widely applied in all the seven subjects considered in this review: gestational diabetes mellitus, preeclampsia, perinatal death, spontaneous abortion, preterm birth, cesarean section, and fetal malformations. The applications are varied, including early detection, alternative screening, biomarker discovery, risk estimation, correlation assessment, pharmacological treatment prediction, drug screening, data acquisition, data extraction, among others. We

observed that the most common ML use is the prediction of perinatal diseases or complications. This is in line with what was described in two recent reviews on ML and pregnancy care. The scoping review of Abuelezz et al. explored the contribution of artificial intelligence in pregnancy, and categorized the applications in “prediction of pregnancy disorders/complications”, “treatment and management” and “assist with patients’ safety outcome”. 75% of the reviewed studies fell into the first category [4]. Likewise, the systematic review of Islam et al. dug into the use of ML to predict pregnancy outcomes. They categorized the reviewed articles according to their scope: “predicting pregnancy risks/complications”, “exploring pregnancy factors”, “predicting mode of delivery”, “predicting outcome of IVF treatment”, “predicting labor outcome” and “comparing two birth weight groups”. The most common was the first category, with a frequency of 35% [5]. Furthermore, we noted that the number of studies employing ML in pregnancy has increased over time, with most of the reviewed articles being published in the last five years. This tendency was also identified by previous reviews in the field [4,5,150].

Depending on the type of data available, different ML methods are preferred for studying pregnancy-related alterations. When the data available come from medical records, the information available is rich in socio-demographic characteristics, medical history variables and anthropometric measurements. We observed that when this is the case, the researchers usually have a massive

amount of data (patients) available, obtained from the aforementioned medical records, to train the ML model. In this scenario, the most used ML methods correspond to non-linear methods, such as SVM, NN, DT, ensemble methods, etc. This could be explained by the fact that correlations between this type of data and the diseases or complications we focused on in this review are complex, not directly or linearly correlated. Non-linear and non-parametrical methods seem to be more suitable in such scenario, in which data is affected by a higher amount of variability and uncertainty. This is especially true when data from medical questionnaires and other surveys are used, in which the answers and values obtained thereof are highly dependent on the patient's perception. Appropriate variable selection and validation of the models is perhaps even more important in those cases. In several studies reviewed in this work, the authors used some level of validation to test their models, and therefore, the accuracies they reported demonstrate a certain relationship between the data used and the pathology studied, even though that relationship is not necessarily linear. Therefore, it is possible to obtain adequate ML models to study adverse perinatal outcomes from data already available. This adds value to currently existing medical records databases.

A fundamental precept in data science is that, in order to predict a property (e.g. a pathology, or the concentration of a particular biomarker) the data must contain information related to that property, and the stronger the correlation, the better the

performance of the model. In this regard, it has been suggested that prediction models could be improved when using biochemical or biophysical variables [64]. This type of data is less affected by human bias and is more directly related to the physiology of an individual, or the pathophysiology of a disease. Most variables of this type correspond to biochemical analytes or ultrasonography parameters. In this scenario, the type of variables used are not too different from the data used in chemical, environmental or pharmaceutical sciences. Analytical chemists have been successfully using chemometrics (i.e. ML applied in chemistry) for several decades to extract relevant information from chemical data, to find correlations or predict a sample property. In essence, the exercise to identify the origin of certain wine from its metal profile, an example of a common application of chemometrics in analytical chemistry, would be no different than predicting a pathology based on the characteristic multivariate pattern of a blood biochemical profile. Likewise, biophysical variables such as the continual recording of FHR through CTG are very similar to the graphical outputs obtained from the analytical instruments used in chemistry (e.g. chromatogram or spectrogram), in the sense that an analytical signal is continuously recorded from an instrument. Therefore, the robust chemometrical platform used in analytical chemistry for the analysis of this type of data could also be exploited in biomedical science. In chemometrics, the most used methods are linear, i.e. are based on linear combinations of the original variables, with which they find hidden correlations that can be used to predict a particular property. Methods such as PCA, partial least squares regression, soft

independent modeling of class analogies, discriminant analysis, or variations of them, are among the most used in chemistry [6,7]. These methods are more intuitive than the non-linear methods mentioned before. Furthermore, they usually provide valuable information about the importance or weight of the variables on the prediction of a certain property, as well as variable-variable and variable-sample relationships, which are some of the reasons they are preferred in chemical analysis. Curiously, in this review we observed that these methods are not common in pregnancy-related applications, where non-linear methods are the trend and LR seems to be almost the only linear method chosen. This observation is consistent with the systematic review and meta-analysis of Sufriyana et al., who found that the most common ML techniques in prognostic prediction studies for pregnancy care are LR (64.8%) and ANN (14.1%) [151]. As clinical chemistry can be considered as a type of analytical chemistry, a more widespread application in biomedicine of the linear ML methods used in chemistry could be highly beneficial, whenever biochemical data is available.

4.1.5.2. Future challenges

It is difficult to think of a field of knowledge in which ML has not been applied. Consequently, it is quite challenging to be innovative regarding the use of ML in the context of pregnancy diseases and complications. An aspect that could be improved is data management, for example by automating their recording,

storage and update in both medical and research settings. The later could ease data extraction, analysis and posterior interpretation. Even though EHR are common in developed countries, they are not frequent in low- or middle- income countries [152,153]. Therefore, the spread of EHR and their adaptation to different realities is an important task for the scientific community in the near future. Moreover, it is necessary to adapt ML applications to the emerging technologies in biomedical sciences, with which novel and more complex types of data can be obtained [154]. This could be employed not only to develop more accurate predictive models, but also to find new biomarkers that could help to better understand the pathophysiology of a particular disease or complication, which could in turn lead to an improvement in its prevention, management or treatment. Furthermore, although there are a lot of published ML models aiming to improve maternal and fetal health, many of them have never been validated, nor subjected to impact analysis. This translational issue was also identified by a recent systematic review on ML-based clinical decision support systems in the context of pregnancy care [155]. To be considered suitable for clinical implementation, ML models have to exhibit a good predictive performance in both internal and external validation, and also prove to foster positive changes in medical settings (e.g. improve patients outcomes, reduce management costs, etc.) without impairing care quality and patient satisfaction [14]. Hence, besides developing new ML tools in the field of pregnancy alterations, it is necessary to carry out studies to test the already published models in different populations and

healthcare facilities. This would allow to know if it is really worthwhile to implement them in clinical practice. To perform such studies is a demanding task, since the recruitment and follow-up of large cohorts of subjects require a very well-coordinated multidisciplinary work, and both time and financial resources are spent. However, it is the only way to lead ML models closer to real medical applications.

Pregnancy lasts only nine months, and the first three have been proposed as the ideal time frame for the early detection, treatment and management of gestational alterations [156–158]. This window of time is narrow, but represents a great opportunity to exploit all the advantages that are associated with the use of ML, i.e. finishing complex assignments rapidly, dealing with multiple tasks efficiently, and predicting short- and long- term outcomes accurately [159,160]. Indeed, this review widely demonstrates that ML methods have a great potential to be applied in such a context, and to contribute to reducing the impact of pregnancy diseases and complications on maternal and fetal health.

4.1.5.3. Strengths of this review

This review is not restricted to a particular ML application on pregnancy diseases and complications. There are a couple of recent systematic reviews that are similar to our work, however they focus on a specific ML application in the field of

pregnancy care, such as the screening of adverse perinatal outcomes [161] and the prediction of perinatal complications [150]. In contrast, this review covers the wide variety of applications that ML may have on maternal and fetal health, including not only the screening or prediction of perinatal alterations, but also biomarker discovery, risk estimation, correlation assessment, pharmacological treatment prediction, drug screening, data acquisition, data extraction, among others, in the context of such alterations. Moreover, this review has a marked clinical focus. There are some recent narrative and systematic reviews that describe different pregnancy-related ML applications, but their emphases are on the applications themselves, and not on specific perinatal pathologies [4,5,162,163]. On the contrary, this review focuses on particular diseases and complications, and gives a broad overview of ML applications for each, which allows to visualize how much ML has penetrated into specific areas of the field of obstetrics and gynecology. Finally, this review covers a considerable body of literature. Most of the reviews found in literature regarding ML and perinatal care include a small number of references [4,5,150,161,162]. Contrarily, this review comprises an important number of scientific articles, which ensures giving a comprehensive overview of the state of art regarding the use of ML in the context of pregnancy diseases and complications.

4.1.5.4. Limitations of this review

Due to the narrative nature of this review, the search and selection of articles was not performed by means of a systematic protocol. Hence, it could be subjected to bias. In addition, this review was restricted to seven selected pregnancy diseases and complications, and English-written articles. Hence, we may have missed some promising ML applications in the field of maternal and fetal health.

4.1.6. Conclusion

The use of ML methods in the context of pregnancy diseases and complications is fairly recent, and is becoming increasingly popular. The applications are varied, and go beyond diagnosis. Indeed, ML has been used to improve the management, treatment, and also the understanding of the pathophysiological mechanisms underlying different perinatal alterations. Facing the challenges that come with working with different types of data, the handling of increasingly large amounts of information, the development of emerging technologies, and the need of translational studies, it is expected that the use of ML methods continue growing in the field of obstetrics and gynecology in the coming years.

Author Contributions: DM, AR, CO, CR, EC, AE-S, JA-R, EG-G and JA tabulated literature information; DM and JA wrote the manuscript; DM, CA, AD, EG-G and JA improved the manuscript.

Funding: This work was financially supported by ANID through PhD scholarship 21190736 (DM), FONDECYT 11181153 (JA) and FOVI210057 (EG-G); and UdeC/MINEDUC via Proyecto UCO 1866 (DM).

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

4.1.7. References

1. Helm JM, Swiergosz AM, Haeberle HS, Karnuta JM, Schaffer JL, Krebs VE, et al. Machine Learning and Artificial Intelligence: Definitions, Applications, and Future Directions. Curr Rev Musculoskelet Med. 2020;13: 69–76. doi:10.1007/s12178-020-09600-8

2. Amigo JM. Data Mining, Machine Learning, Deep Learning, Chemometrics. Definitions, common points and Trends (Spoiler Alert: VALIDATE your models!). Brazilian J Anal Chem. 2021;8: 45–61. doi:10.30744/brjac.2179-3425.AR-38-2021

3. Goecks J, Jalili V, Heiser LM, Gray JW. How Machine Learning Will Transform Biomedicine. Cell. 2020;181: 92–101. doi:10.1016/j.cell.2020.03.022

4. Abuelezz I, Hassan A, Jaber BA, Sharique M, Abd-Alrazaq A, Househ M, et al. Contribution of Artificial Intelligence in Pregnancy: A Scoping Review. Stud Health Technol Inform. 2022;289: 333–336. doi:10.3233/SHTI210927

5. Islam MN, Mustafina SN, Mahmud T, Khan NI. Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda. BMC Pregnancy Childbirth. 2022;22: 1–19. doi:10.1186/s12884-022-04594-2

6. Brereton RG. Applied Chemometrics for Scientists. Wiley; 2007. doi:10.1002/9780470057780

7. Otto M. Chemometrics. Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA; 2016. doi:10.1002/9783527699377

8. Gibbs RS, Karlan BY, Haney AF, Nygaard IE. Danforth's Obstetrics and Gynecology. 10th ed. Baltimore, MD: Lippincott Williams & Wilkins; 2008

9. Daszykowski M, Kaczmarek K, Vander Heyden Y, Walczak B. Robust statistics in data analysis — A review. Chemom Intell Lab Syst. 2007;85: 203–219. doi:10.1016/j.chemolab.2006.06.016

10. Engel J, Gerretzen J, Szymańska E, Jansen JJ, Downey G, Blanchet L, et al. Breaking with trends in pre-processing? TrAC Trends Anal Chem. 2013;50: 96–106. doi:10.1016/j.trac.2013.04.015

11. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2018;15: 233–234. doi:10.1038/nmeth.4642

12. Ballabio D, Grisoni F, Todeschini R. Multivariate comparison of classification performance measures. Chemom Intell Lab Syst. 2018;174: 33–44. doi:10.1016/j.chemolab.2017.12.004

13. Botchkarev A. A New Typology Design of Performance Metrics to Measure Errors in Machine Learning Regression Algorithms. Interdiscip J Information, Knowledge, Manag. 2019;14: 045–076. doi:10.28945/4184

14. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS, et al. Users' Guides to the Medical Literature. JAMA. 2000;284: 79. doi:10.1001/jama.284.1.79

15. Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagnostic Progn Res. 2019;3: 16. doi:10.1186/s41512-019-0060-y

16. ADA. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2021. Diabetes Care. 2021;44: S15–S33. doi:10.2337/dc21-S002

17. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. Nat Rev Dis Prim. 2019;5: 47. doi:10.1038/s41572-019-0098-8

18. Mennickent D, Rodríguez A, Farías-Jofré M, Araya J, Guzmán-Gutiérrez E. Machine learning-based models for gestational diabetes mellitus prediction before 24–28 weeks of pregnancy: A review. Artif Intell Med. 2022;132: 102378. doi:10.1016/j.artmed.2022.102378

19. Xiong Y, Lin L, Chen Y, Salerno S, Li Y, Zeng X, et al. Prediction of gestational diabetes mellitus in the first 19 weeks of pregnancy using machine learning techniques. J Matern Neonatal Med. 2022;35: 2457–2463. doi:10.1080/14767058.2020.1786517

20. Zheng T, Ye W, Wang X, Li X, Zhang J, Little J, et al. A simple model to predict risk of gestational diabetes mellitus from 8 to 20 weeks of gestation in Chinese women. BMC Pregnancy Childbirth. 2019;19: 252. doi:10.1186/s12884-019-2374-8

21. Wu Y-TT, Zhang C-JJ, Mol BW, Kawai A, Li C, Chen L, et al. Early Prediction of Gestational Diabetes Mellitus in the Chinese Population via Advanced Machine Learning. J Clin Endocrinol Metab. 2021;106: e1191–e1205. doi:10.1210/clinem/dgaa899

22. Artzi NS, Shilo S, Hadar E, Rossman H, Barbash-Hazan S, Ben-Haroush A, et al. Prediction of gestational diabetes based on nationwide electronic health records. Nat Med. 2020;26: 71–76. doi:10.1038/s41591-019-0724-8

23. Balani J, Hyer S, Shehata H, Mohareb F. Visceral fat mass as a novel risk factor for predicting gestational diabetes in obese pregnant women. Obstet Med. 2018;11: 121–125. doi:10.1177/1753495X17754149

24. Mennickent D, Ortega-Contreras B, Gutiérrez-Vega S, Castro E, Rodríguez A, Araya J, et al. Evaluation of first and second trimester maternal thyroid profile on the prediction of gestational diabetes mellitus and post load glycemia. PLoS One. 2023;18: e0280513. doi:10.1371/journal.pone.0280513

25. Scott HD, Buchan M, Chadwick C, Field CJ, Letourneau N, Montina T, et al. Metabolic dysfunction in pregnancy: Fingerprinting the maternal metabolome using proton nuclear magnetic resonance spectroscopy. Endocrinol Diabetes Metab. 2021;4: 1–10. doi:10.1002/edm2.201

26. Yoffe L, Polsky A, Gilam A, Raff C, Mecacci F, Ognibene A, et al. Early diagnosis of gestational diabetes mellitus using circulating microRNAs. Eur J Endocrinol. 2019;181: 565–577. doi:10.1530/EJE-19-0206

27. Guo Z, Yang F, Zhang J, Zhang Z, Li K, Tian Q, et al. Whole-Genome Promoter Profiling of Plasma DNA Exhibits Diagnostic Value for Placenta-Origin Pregnancy Complications. Adv Sci. 2020;7: 1901819. doi:10.1002/advs.201901819

28. Liu Y, Wang Z, Zhao L. Identification of diagnostic cytosine-phosphate-guanine biomarkers in patients with gestational diabetes mellitus via epigenome-wide association study and machine learning. Gynecol Endocrinol. 2021;37: 857–862. doi:10.1080/09513590.2021.1937101

29. Ehrlich SF, Ferrara A, Hedderson MM, Feng J, Neugebauer R. Exercise During the First Trimester of Pregnancy and the Risks of Abnormal Screening and Gestational Diabetes Mellitus. Diabetes Care. 2021;44: 425–432. doi:10.2337/dc20-1475

30. Bernardes-Oliveira E, de Freitas DLD, de Morais C de LM, Cornetta M da C de M, Camargo JD de AS, de Lima KMG, et al. Spectrochemical differentiation in gestational diabetes mellitus based on attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy and multivariate analysis. Sci Rep. 2020;10: 19259. doi:10.1038/s41598-020-75539-y

31. Araya J, Rodriguez A, Lagos-SanMartin K, Mennickent D, Gutiérrez-Vega S, Ortega-Contreras B, et al. Maternal thyroid profile in first and second trimester of pregnancy is correlated with gestational diabetes mellitus through machine learning. Placenta. 2021;103: 82–85. doi:10.1016/j.placenta.2020.10.015

32. Velardo C, Clifton D, Hamblin S, Khan R, Tarassenko L, Mackillop L. Toward a Multivariate Prediction Model of Pharmacological Treatment for Women With Gestational Diabetes Mellitus: Algorithm Development and Validation. J Med Internet Res. 2021;23: e21435. doi:10.2196/21435

33. Ababneh M. Management of pre-eclampsia/eclampsia. Middle East J Anaesthesiol. 2004;17: 939–50. Available: http://www.ncbi.nlm.nih.gov/pubmed/15449750

34. ACOG. Hypertension in Pregnancy. Obstet Gynecol. 2013;122: 1122–1131. doi:10.1097/01.AOG.0000437382.03963.88

35. Mureșan D, Rotar IC, Stamatian F. The usefulness of fetal Doppler evaluation in early versus late onset intrauterine growth restriction. Review of the literature. Med Ultrason. 2016;18: 103. doi:10.11152/mu.2013.2066.181.dop

36. Gyselaers W, Mullens W, Tomsin K, Mesens T, Peeters L. Role of dysfunctional maternal venous hemodynamics in the pathophysiology of pre-eclampsia: a review. Ultrasound Obstet Gynecol. 2011;38: 123–129. doi:10.1002/uog.9061

37. Keikkala E, Vuorela P, Laivuori H, Romppanen J, Heinonen S, Stenman U-H. First trimester hyperglycosylated human chorionic gonadotrophin in serum – A marker of early-onset preeclampsia. Placenta. 2013;34: 1059–1065. doi:10.1016/j.placenta.2013.08.006

38. Birdir C, Fryze J, Vasiliadis H, Nicolaides KH, Poon LC. Maternal serum anti-Müllerian hormone at 11–13 weeks' gestation in the prediction of preeclampsia. J Matern Neonatal Med. 2015;28: 865–868. doi:10.3109/14767058.2014.937418

39. Stepan H, Unversucht A, Wessel N, Faber R. Predictive Value of Maternal Angiogenic Factors in Second Trimester Pregnancies With Abnormal Uterine Perfusion. Hypertension. 2007;49: 818–824. doi:10.1161/01.HYP.0000258404.21552.a3

40. Stepan H, Geipel A, Schwarz F, Krämer T, Wessel N, Faber R. Circulatory soluble endoglin and its predictive value for preeclampsia in second-trimester pregnancies with abnormal uterine perfusion. Am J Obstet Gynecol. 2008;198: 175.e1-175.e6. doi:10.1016/j.ajog.2007.08.052

41. Lee SM, Kang Y, Lee EM, Jung YM, Hong S, Park SJ, et al. Metabolomic biomarkers in midtrimester maternal plasma can accurately predict the development of preeclampsia. Sci Rep. 2020;10: 16142. doi:10.1038/s41598-020-72852-4

42. Tarca AL, Romero R, Benshalom-Tirosh N, Than NG, Gudicha DW, Done B, et al. The prediction of early preeclampsia: Results from a longitudinal

proteomics study. PLoS One. 2019;14: e0217273. doi:10.1371/journal.pone.0217273

43. Jiang M, Lash GE, Zhao X, Long Y, Guo C, Yang H. CircRNA-0004904, CircRNA-0001855, and PAPP-A: Potential Novel Biomarkers for the Prediction of Preeclampsia. Cell Physiol Biochem. 2018;46: 2576–2586. doi:10.1159/000489685

44. Mello G, Parretti E, Ognibene A, Mecacci F, Cioni R, Scarselli G, et al. Prediction of the Development of Pregnancy-Induced Hypertensive Disorders in High-Risk Pregnant Women by Artificial Neural Networks. Clin Chem Lab Med. 2001;39: 801–805. doi:10.1515/CCLM.2001.132

45. Jhee JH, Lee S, Park Y, Lee SE, Kim YA, Kang S-W, et al. Prediction model development of late-onset preeclampsia using machine learning-based methods. PLoS One. 2019;14: e0221202. doi:10.1371/journal.pone.0221202

46. Sandström A, Snowden JM, Höijer J, Bottai M, Wikström A-K. Clinical risk assessment in early pregnancy for preeclampsia in nulliparous women: A population based cohort study. PLoS One. 2019;14: e0225716. doi:10.1371/journal.pone.0225716

47. Han Q, Zheng W, Guo XD, Zhang D, Liu HF, Yu L, et al. A new predicting model of preeclampsia based on peripheral blood test values. Eur Rev Med Pharmacol Sci. 2020;24: 7222–7229. doi:10.26355/eurrev_202007_21874

48. Marić I, Tsur A, Aghaeepour N, Montanari A, Stevenson DK, Shaw GM, et al. Early prediction of preeclampsia via machine learning. Am J Obstet Gynecol MFM. 2020;2: 100100. doi:10.1016/j.ajogmf.2020.100100

49. Sufriyana H, Wu Y-W, Su EC-Y. Prediction of Preeclampsia and Intrauterine Growth Restriction: Development of Machine Learning Models on a Prospective Cohort. JMIR Med Informatics. 2020;8: e15411. doi:10.2196/15411

50. Sufriyana H, Wu Y-W, Su EC-Y. Artificial intelligence-assisted prediction of preeclampsia: Development and external validation of a nationwide health insurance dataset of the BPJS Kesehatan in Indonesia. EBioMedicine. 2020;54: 102710. doi:10.1016/j.ebiom.2020.102710

51. Gupta K, Balyan K, Lamba B, Puri M, Sengupta D, Kumar M. Ultrasound placental image texture analysis using artificial intelligence to predict

hypertension in pregnancy. J Matern Neonatal Med. 2022;35: 5587–5594. doi:10.1080/14767058.2021.1887847

52. Liu K, Fu Q, Liu Y, Wang C. An integrative bioinformatics analysis of microarray data for identifying hub genes as diagnostic biomarkers of preeclampsia. Biosci Rep. 2019;39: 866–875. doi:10.1042/BSR20190187

53. Guo R, Teng Z, Wang Y, Zhou X, Xu H, Liu D. Integrated Learning: Screening Optimal Biomarkers for Identifying Preeclampsia in Placental mRNA Samples. Comput Math Methods Med. 2021;2021: 1–7. doi:10.1155/2021/6691096

54. Bodnar LM, Cartus AR, Kirkpatrick SI, Himes KP, Kennedy EH, Simhan HN, et al. Machine learning as a strategy to account for dietary synergy: an illustration based on dietary intake and adverse pregnancy outcomes. Am J Clin Nutr. 2020;111: 1235–1243. doi:10.1093/ajcn/nqaa027

55. Tejera E, Pérez-Castillo Y, Chamorro A, Cabrera-Andrade A, Sanchez ME. A Multi-Objective Approach for Drug Repurposing in Preeclampsia. Molecules. 2021;26: 777. doi:10.3390/molecules26040777

56. WHO. Making Every Baby Count : audit and review of stillbirths and neonatal deaths. World Health Organization. 2016.

57. WHO. The WHO application of ICD-10 to deaths during the perinatal period: ICD-PM. World Health Organization. 2016.

58. Liu L, Oza S, Hogan D, Perin J, Rudan I, Lawn JE, et al. Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: an updated systematic analysis. Lancet. 2015;385: 430–440. doi:10.1016/S0140-6736(14)61698-6

59. Blencowe H, Cousens S, Jassir FB, Say L, Chou D, Mathers C, et al. National, regional, and worldwide estimates of stillbirth rates in 2015, with trends from 2000: a systematic analysis. Lancet Glob Heal. 2016;4: e98–e108. doi:10.1016/S2214-109X(15)00275-2

60. Willcox ML, Price J, Scott S, Nicholson BD, Stuart B, Roberts NW, et al. Death audits and reviews for reducing maternal, perinatal and child mortality. Cochrane Database Syst Rev. 2020;2020. doi:10.1002/14651858.CD012982.pub2

61. Abebe H, Adane D, Shitu S. Essential newborn care practice and its associated factors in Southwest Ethiopia. Arch Public Heal. 2021;79: 42. doi:10.1186/s13690-021-00568-6

62. Malacova E, Tippaya S, Bailey HD, Chai K, Farrant BM, Gebremedhin AT, et al. Stillbirth risk prediction using machine learning for a large cohort of births from Western Australia, 1980–2015. Sci Rep. 2020;10: 5354. doi:10.1038/s41598-020-62210-9

63. Kayode GA, Grobbee DE, Amoakoh-Coleman M, Adeleke IT, Ansah E, de Groot JAH, et al. Predicting stillbirth in a low resource setting. BMC Pregnancy Childbirth. 2016;16: 274. doi:10.1186/s12884-016-1061-2

64. Koivu A, Sairanen M. Predicting risk of stillbirth and preterm pregnancies with machine learning. Heal Inf Sci Syst. 2020;8: 14. doi:10.1007/s13755-020-00105-9

65. Shukla V V., Eggleston B, Ambalavanan N, McClure EM, Mwenechanya M, Chomba E, et al. Predictive Modeling for Perinatal Mortality in Resource-Limited Settings. JAMA Netw Open. 2020;3: e2026750. doi:10.1001/jamanetworkopen.2020.26750

66. Mboya IB, Mahande MJ, Mohammed M, Obure J, Mwambi HG. Prediction of perinatal death using machine learning models: a birth registry-based cohort study in northern Tanzania. BMJ Open. 2020;10: e040132. doi:10.1136/bmjopen-2020-040132

67. Khatibi T, Hanifi E, Sepehri MM, Allahqoli L. Proposing a machine-learning based method to predict stillbirth before and during delivery and ranking the features: nationwide retrospective cross-sectional study. BMC Pregnancy Childbirth. 2021;21: 202. doi:10.1186/s12884-021-03658-z

68. Mangold C, Zoretic S, Thallapureddy K, Moreira A, Chorath K, Moreira A. Machine Learning Models for Predicting Neonatal Mortality: A Systematic Review. Neonatology. 2021;118: 394–405. doi:10.1159/000516891

69. Cooper JN, Minneci PC, Deans KJ. Postoperative neonatal mortality prediction using superlearning. J Surg Res. 2018;221: 311–319. doi:10.1016/j.jss.2017.09.002

70. Jaskari J, Myllarinen J, Leskinen M, Rad AB, Hollmen J, Andersson S, et al. Machine Learning Methods for Neonatal Mortality and Morbidity Classification. IEEE Access. 2020;8: 123347–123358. doi:10.1109/ACCESS.2020.3006710

71. Rezaeian A, Rezaeian M, Khatami SF, Khorashadizadeh F, Moghaddam FP. Prediction of mortality of premature neonates using neural network and logistic regression. J Ambient Intell Humaniz Comput. 2022;13: 1269–1277. doi:10.1007/s12652-020-02562-2

72. Sheikhtaheri A, Zarkesh MR, Moradi R, Kermani F. Prediction of neonatal deaths in NICUs: development and validation of machine learning models. BMC Med Inform Decis Mak. 2021;21: 131. doi:10.1186/s12911-021-01497-8

73. Larsen EC, Christiansen OB, Kolte AM, Macklon N. New insights into mechanisms behind miscarriage. BMC Med. 2013;11: 154. doi:10.1186/1741-7015-11-154

74. Bottomley C, Van Belle V, Kirk E, Van Huffel S, Timmerman D, Bourne T. Accurate prediction of pregnancy viability by means of a simple scoring system. Hum Reprod. 2013;28: 68–76. doi:10.1093/humrep/des352

75. Mu Y, Yuan Y, Han W, Pian P, Li L, Wang D, et al. Exploring the Molecular Mechanism and Biomarker of Recurrent Spontaneous Abortion Based on RNA Sequencing Analysis. Clin Lab. 2020;66: 2055–2061. doi:10.7754/Clin.Lab.2020.200222

76. Toh C, P. Brody J. Applications of Machine Learning in Healthcare. Smart Manufacturing - When Artificial Intelligence Meets the Internet of Things. IntechOpen; 2021. doi:10.5772/intechopen.92297

77. Mora-Sánchez A, Aguilar-Salvador D-I, Nowak I. Towards a gamete matching platform: using immunogenetics and artificial intelligence to predict recurrent miscarriage. npj Digit Med. 2019;2: 12. doi:10.1038/s41746-019-0089-x

78. Meuleman T, Lashley LELO, Dekkers OM, van Lith JMM, Claas FHJ, Bloemenkamp KWM. HLA associations and HLA sharing in recurrent miscarriage: A systematic review and meta-analysis. Hum Immunol. 2015;76: 362–373. doi:10.1016/j.humimm.2015.02.004

79. Berger DS, Hogge WA, Barmada MM, Ferrell RE. Comprehensive Analysis of HLA-G: Implications for Recurrent Spontaneous Abortion. Reprod Sci. 2010;17: 331–338. doi:10.1177/1933719109356802

80. Wu Y, Xin M, Han Q, Wang J, Yin X, He J, et al. Extensive serum biomarker analysis in the prethrombotic state of recurrent spontaneous abortion. J Cell Mol Med. 2021;25: 6679–6694. doi:10.1111/jcmm.16671

81. Medenica S, Zivanovic D, Batkoska L, Marinelli S, Basile G, Perino A, et al. The Future Is Coming: Artificial Intelligence in the Treatment of Infertility Could Improve Assisted Reproduction Outcomes—The Value of Regulatory Frameworks. Diagnostics. 2022;12. doi:10.3390/diagnostics12122979

82. Liu L, Jiao Y, Li X, Ouyang Y, Shi D. Machine learning algorithms to predict early pregnancy loss after in vitro fertilization-embryo transfer with fetal heart rate as a strong predictor. Comput Methods Programs Biomed. 2020;196: 105624. doi:10.1016/j.cmpb.2020.105624

83. Huang C, Xiang Z, Zhang Y, Tan DS, Yip CK, Liu Z, et al. Using Deep Learning in a Monocentric Study to Characterize Maternal Immune Environment for Predicting Pregnancy Outcomes in the Recurrent Reproductive Failure Patients. Front Immunol. 2021;12: 1–13. doi:10.3389/fimmu.2021.642167

84. Chawanpaiboon S, Vogel JP, Moller A-B, Lumbiganon P, Petzold M, Hogan D, et al. Global, regional, and national estimates of levels of preterm birth in 2014: a systematic review and modelling analysis. Lancet Glob Heal. 2019;7: e37–e46. doi:10.1016/S2214-109X(18)30451-0

85. Suff N, Story L, Shennan A. The prediction of preterm delivery: What is new? Semin Fetal Neonatal Med. 2019;24: 27–32. doi:10.1016/j.siny.2018.09.006

86. Khatibi T, Kheyrikoochaksarayee N, Sepehri MM. Analysis of big data for prediction of provider-initiated preterm birth and spontaneous premature deliveries and ranking the predictive features. Arch Gynecol Obstet. 2019;300: 1565–1582. doi:10.1007/s00404-019-05325-3

87. Arabi Belaghi R, Beyene J, McDonald SD. Prediction of preterm birth in nulliparous women using logistic regression and machine learning. PLoS One. 2021;16: e0252025. doi:10.1371/journal.pone.0252025

88. Gao C, Osmundson S, Velez Edwards DR, Jackson GP, Malin BA, Chen Y. Deep learning predicts extreme preterm birth from electronic health records. J Biomed Inform. 2019;100: 103334. doi:10.1016/j.jbi.2019.103334

89. Zhang Y, Lu S, Wu Y, Hu W, Yuan Z. The Prediction of Preterm Birth Using Time-Series Technology-Based Machine Learning: Retrospective Cohort Study. JMIR Med Informatics. 2022;10: e33835. doi:10.2196/33835

90. Aung MT, Yu Y, Ferguson KK, Cantonwine DE, Zeng L, McElrath TF, et al. Prediction and associations of preterm birth and its subtypes with eicosanoid enzymatic pathways and inflammatory markers. Sci Rep. 2019;9: 17049. doi:10.1038/s41598-019-53448-z

91. Chen Y, He B, Liu Y, Aung MT, Rosario-Pabón Z, Vélez-Vega CM, et al. Maternal plasma lipids are involved in the pathogenesis of preterm birth. Gigascience. 2022;11: 1–11. doi:10.1093/gigascience/giac004

92. Jehan F, Sazawal S, Baqui AH, Nisar MI, Dhingra U, Khanam R, et al. Multiomics Characterization of Preterm Birth in Low- and Middle-Income Countries. JAMA Netw Open. 2020;3: e2029655. doi:10.1001/jamanetworkopen.2020.29655

93. Despotovic D, Zec A, Mladenovic K, Radin N, Turukalo TL. A Machine Learning Approach for an Early Prediction of Preterm Delivery. 2018 IEEE 16th Int Symp Intell Syst Informatics. 2018; 265–270. doi:10.1109/SISY.2018.8524818

94. Rawashdeh H, Awawdeh S, Shannag F, Henawi E, Faris H, Obeid N, et al. Intelligent system based on data mining techniques for prediction of preterm birth for women with cervical cerclage. Comput Biol Chem. 2020;85: 107233. doi:10.1016/j.compbiolchem.2020.107233

95. Ye J, Zhang J, Mikolajczyk R, Torloni M, Gülmezoglu A, Betran A. Association between rates of caesarean section and maternal and neonatal mortality in the 21st century: a worldwide population-based ecological study with longitudinal data. BJOG An Int J Obstet Gynaecol. 2016;123: 745–753. doi:10.1111/1471-0528.13592

96. Martinez LD, Glynn LM, Sandman CA, Wing DA, Davis EP. Cesarean delivery and infant cortisol regulation. Psychoneuroendocrinology. 2020;122: 104862. doi:10.1016/j.psyneuen.2020.104862

97. Gosset M, Ilenko A, Bouyou J, Renevier B. Emergency caesarean section. J Visc Surg. 2017;154: 47–50. doi:10.1016/j.jviscsurg.2016.09.012

98. David M, Scherer K, Henrich W, Breckenkamp J. Does an Immigrant Background Affect the Indication, Incidence or Outcome of Emergency Cesarean Section? Results of the Prospective Data Collection of 111 Births. Geburtshilfe Frauenheilkd. 2018;78: 167–172. doi:10.1055/s-0044-100147

99. Ryding EL, Wijma K, Wijma B. Psychological impact of emergency Cesarean section in comparison with elective Cesarean section, instrumental and normal vaginal delivery. J Psychosom Obstet Gynecol. 1998;19: 135–144. doi:10.3109/01674829809025691

100. Yang X-J, Sun S-S. Comparison of maternal and fetal complications in elective and emergency cesarean section: a systematic review and meta-analysis. Arch Gynecol Obstet. 2017;296: 503–512. doi:10.1007/s00404-017-4445-2

101. Wangel A-M, Molin J, Östman M, Jernström H. Emergency cesarean sections can be predicted by markers for stress, worry and sleep disturbances in first-time mothers. Acta Obstet Gynecol Scand. 2011;90: no-no. doi:10.1111/j.1600-0412.2010.01056.x

102. Fergus P, Hussain A, Al-Jumeily D, Huang D-S, Bouguila N. Classification of caesarean section and normal vaginal deliveries using foetal heart rate signals and advanced machine learning algorithms. Biomed Eng Online. 2017;16: 89. doi:10.1186/s12938-017-0378-z

103. Pinas A, Chandraharan E. Continuous cardiotocography during labour: Analysis, classification and management. Best Pract Res Clin Obstet Gynaecol. 2016;30: 33–47. doi:10.1016/j.bpobgyn.2015.03.022

104. Fergus P, Selvaraj M, Chalmers C. Machine learning ensemble modelling to classify caesarean section and vaginal delivery types using Cardiotocography traces. Comput Biol Med. 2018;93: 7–16. doi:10.1016/j.compbiomed.2017.12.002

105. Ricciardi C, Improta G, Amato F, Cesarelli G, Romano M. Classifying the type of delivery from cardiotocographic signals: A machine learning approach. Comput Methods Programs Biomed. 2020;196: 105712. doi:10.1016/j.cmpb.2020.105712

106. Tang H, Wang T, Li M, Yang X. The Design and Implementation of Cardiotocography Signals Classification Algorithm Based on Neural Network. Comput Math Methods Med. 2018;2018: 1–12. doi:10.1155/2018/8568617

107. Clark RRS, Hou J. Three machine learning algorithms and their utility in exploring risk factors associated with primary cesarean section in low-risk women: A methods paper. Res Nurs Health. 2021;44: 559–570. doi:10.1002/nur.22122

108. Wei C, Wang L, Chang X, Zhou Q. A prediction model using machine-learning algorithm for assessing intrathecal hyperbaric bupivacaine dose during cesarean section. BMC Anesthesiol. 2021;21: 116. doi:10.1186/s12871-021-01331-8

109. Fletcher RR, Olubeko O, Sonthalia H, Kateera F, Nkurunziza T, Ashby JL, et al. Application of Machine Learning to Prediction of Surgical Site Infection. 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2019. pp. 2234–2237. doi:10.1109/EMBC.2019.8857942

110. Lipschuetz M, Guedalia J, Rottenstreich A, Novoselsky Persky M, Cohen SM, Kabiri D, et al. Prediction of vaginal birth after cesarean deliveries using machine learning. Am J Obstet Gynecol. 2020;222: 613.e1-613.e12. doi:10.1016/j.ajog.2019.12.267

111. Akbulut A, Ertugrul E, Topcu V. Fetal health status prediction based on maternal clinical history using machine learning techniques. Comput Methods Programs Biomed. 2018;163: 87–100. doi:10.1016/j.cmpb.2018.06.010

112. Guimarães-Ferreira J, Miguéns J, Lauritzen C. Advances in Craniosynostosis Research and Management. Advances and technical standards in neurosurgery. 2004. pp. 23–83. doi:10.1007/978-3-7091-0558-0_2

113. Mendoza CS, Safdar N, Okada K, Myers E, Rogers GF, Linguraru MG. Personalized assessment of craniosynostosis via statistical shape modeling. Med Image Anal. 2014;18: 635–646. doi:10.1016/j.media.2014.02.008

114. Safran T, Gorsky K, Viezel-Mathieu A, Kanevsky J, Gilardino MS. The role of ultrasound technology in plastic surgery. J Plast Reconstr Aesthetic Surg. 2018;71: 416–424. doi:10.1016/j.bjps.2017.08.031

115. Wu JY, Tuomi A, Beland MD, Konrad J, Glidden D, Grand D, et al. Quantitative analysis of ultrasound images for computer-aided diagnosis. J Med Imaging. 2016;3: 014501. doi:10.1117/1.JMI.3.1.014501

116. Meulstee JW, Verhamme LM, Borstlap WA, Van der Heijden F, De Jong GA, Xi T, et al. A new method for three-dimensional evaluation of the cranial shape and the automatic identification of craniosynostosis using 3D stereophotogrammetry. Int J Oral Maxillofac Surg. 2017;46: 819–826. doi:10.1016/j.ijom.2017.03.017

117. Wren C. Temporal variability in birth prevalence of cardiovascular malformations. Heart. 2000;83: 414–419. doi:10.1136/heart.83.4.414

118. Moons P, Sluysmans T, De Wolf D, Massin M, Suys B, Benatar A, et al. Congenital heart disease in 111 225 births in Belgium: birth prevalence, treatment and survival in the 21st century. Acta Paediatr. 2009;98: 472–477. doi:10.1111/j.1651-2227.2008.01152.x

119. Rychik J, Tian Z, Bebbington M, Xu F, McCann M, Mann S, et al. The twin-twin transfusion syndrome: spectrum of cardiovascular abnormality and development of a cardiovascular score to assess severity of disease. Am J Obstet Gynecol. 2007;197: 392.e1-392.e8. doi:10.1016/j.ajog.2007.06.055

120. Garcia-Canadilla P, Sanchez-Martinez S, Crispi F, Bijnens B. Machine Learning in Fetal Cardiology: What to Expect. Fetal Diagn Ther. 2020;47: 363–372. doi:10.1159/000505021

121. Crispi F, Valenzuela-Alcaraz B, Cruz-Lemini M, Gratacós E. Ultrasound assessment of fetal cardiac function. Australas J Ultrasound Med. 2013;16: 158–167. doi:10.1002/j.2205-0140.2013.tb00242.x

122. Bridge CP, Ioannou C, Noble JA. Automated annotation and quantitative description of ultrasound videos of the fetal heart. Med Image Anal. 2017;36: 147–161. doi:10.1016/j.media.2016.11.006

123. Yu Q, Yan H, Song L, Guo W, Liu H, Si J, et al. Automatic identifying of maternal ECG source when applying ICA in fetal ECG extraction. Biocybern Biomed Eng. 2018;38: 448–455. doi:10.1016/j.bbe.2018.03.003

124. Muduli PR, Gunukula RR, Mukherjee A. A deep learning approach to fetal-ECG signal reconstruction. 2016 Twenty Second National Conference on Communication (NCC). IEEE; 2016. pp. 1–6. doi:10.1109/NCC.2016.7561206

125. Maraci MA, Bridge CP, Napolitano R, Papageorghiou A, Noble JA. A framework for analysis of linear ultrasound videos to detect fetal presentation and heartbeat. Med Image Anal. 2017;37: 22–36. doi:10.1016/j.media.2017.01.003

126. Krupa N, MA M, Zahedi E, Ahmed S, Hassan FM. Antepartum fetal heart rate feature extraction and classification using empirical mode decomposition and support vector machine. Biomed Eng Online. 2011;10: 6. doi:10.1186/1475-925X-10-6

127. Fotiadou E, Xu M, van Erp B, van Sloun RJG, Vullings R. Deep Convolutional Long Short-Term Memory Network for Fetal Heart Rate Extraction. 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE; 2020. pp. 1–4. doi:10.1109/EMBC44109.2020.9175442

128. Fotiadou E, van Sloun RJG, van Laar JOEH, Vullings R. A dilated inception CNN-LSTM network for fetal heart rate estimation. Physiol Meas. 2021;42: 045007. doi:10.1088/1361-6579/abf7db

129. Al-Angari HM, Kimura Y, Hadjileontiadis LJ, Khandoker AH. A Hybrid EMD-Kurtosis Method for Estimating Fetal Heart Rate from Continuous Doppler Signals. Front Physiol. 2017;8: 1–8. doi:10.3389/fphys.2017.00641

130. Zhao Z, Deng Y, Zhang Y, Zhang Y, Zhang X, Shao L. DeepFHR: intelligent prediction of fetal Acidemia using fetal heart rate signals based on convolutional neural network. BMC Med Inform Decis Mak. 2019;19: 286. doi:10.1186/s12911-019-1007-5

131. Brocklehurst P. A study of an intelligent system to support decision making in the management of labour using the cardiotocograph – the INFANT study protocol. BMC Pregnancy Childbirth. 2016;16: 10. doi:10.1186/s12884-015-0780-0

132. Parer JT, Hamilton EF. Comparison of 5 experts and computer analysis in rule-based fetal heart rate interpretation. Am J Obstet Gynecol. 2010;203: 451.e1-451.e7. doi:10.1016/j.ajog.2010.05.037

133. Alonso-Betanzos A, Moret-Bonillo V, Hernandez-Sande C. Foetos: an expert system for fetal assessment. IEEE Trans Biomed Eng. 1991;38: 199–211. doi:10.1109/10.76387

134. Cruz-Lemini M, Crispi F, Valenzuela-Alcaraz B, Figueras F, Sitges M, Bijnens B, et al. Fetal cardiovascular remodeling persists at 6 months in infants with intrauterine growth restriction. Ultrasound Obstet Gynecol. 2016;48: 349–356. doi:10.1002/uog.15767

135. Gurgen F, Onal E, Varol FG. IUGR detection by ultrasonographic examinations using neural networks. IEEE Eng Med Biol Mag. 1997;16: 55–58. doi:10.1109/51.585518

136. Magenes G, Pedrinazzi L, Signorini MG. Identification of fetal sufferance antepartum through a multiparametric analysis and a support vector machine. The 26th Annual International Conference of the IEEE

Engineering in Medicine and Biology Society. IEEE; 2004. pp. 462–465. doi:10.1109/IEMBS.2004.1403194

137. Gadagkar A V., Shreedhara KS. Features Based IUGR Diagnosis Using Variational Level Set Method and Classification Using Artificial Neural Networks. 2014 Fifth International Conference on Signal and Image Processing. IEEE; 2014. pp. 303–309. doi:10.1109/ICSIP.2014.54

138. Yeo L, Romero R. Fetal Intelligent Navigation Echocardiography (FINE): a novel method for rapid, simple, and automatic examination of the fetal heart. Ultrasound Obstet Gynecol. 2013;42: 268–284. doi:10.1002/uog.12563

139. Han G, Jin T, Zhang L, Guo C, Gui H, Na R, et al. Adoption of Compound Echocardiography under Artificial Intelligence Algorithm in Fetal Congenial Heart Disease Screening during Gestation. Appl Bionics Biomech. 2022;2022: 1–8. doi:10.1155/2022/6410103

140. Arnaout R, Curran L, Chinn E, Zhao Y, Moon-Grady A. Deep-learning models improve on community-level diagnosis for common congenital heart disease lesions. 2018; 1–17. Available: https://arxiv.org/abs/1809.06993v1

141. Resendiz M, Mason S, Lo C-L, Zhou FC. Epigenetic regulation of the neural transcriptome and alcohol interference during development. Front Genet. 2014;5: 1–15. doi:10.3389/fgene.2014.00285

142. Al-Shaer AE, Flentke GR, Berres ME, Garic A, Smith SM. Exon level machine learning analyses elucidate novel candidate miRNA targets in an avian model of fetal alcohol spectrum disorder. PLOS Comput Biol. 2019;15: e1006937. doi:10.1371/journal.pcbi.1006937

143. Langer O, Berkus MD, Huff RW, Samueloff A. Shoulder dystocia: Should the fetus weighing ≥4000 grams be delivered by cesarean section? Am J Obstet Gynecol. 1991;165: 831–837. doi:10.1016/0002-9378(91)90424-P

144. Oral E, Cağdaş A, Gezer A, Kaleli S, Aydinli K, Öçer F. Perinatal and maternal outcomes of fetal macrosomia. Eur J Obstet Gynecol Reprod Biol. 2001;99: 167–171. doi:10.1016/S0301-2115(01)00416-X

145. Shigemi D, Yamaguchi S, Aso S, Yasunaga H. Predictive model for macrosomia using maternal parameters without sonography information. J Matern Neonatal Med. 2019;32: 3859–3863. doi:10.1080/14767058.2018.1484090

146. Tao J, Yuan Z, Sun L, Yu K, Zhang Z. Fetal birthweight prediction with measured data by a temporal machine learning method. BMC Med Inform Decis Mak. 2021;21: 26. doi:10.1186/s12911-021-01388-y

147. Challa AP, Beam AL, Shen M, Peryea T, Lavieri RR, Lippmann ES, et al. Machine learning on drug-specific data to predict small molecule teratogenicity. Reprod Toxicol. 2020;95: 148–158. doi:10.1016/j.reprotox.2020.05.004

148. Garg RC, Bracken WM, M. Hoberman A. Reproductive and developmental safety evaluation of new pharmaceutical compounds. Reproductive and Developmental Toxicology. Elsevier; 2011. pp. 89–109. doi:10.1016/B978-0-12-382032-7.10008-6

149. Kim JH, Scialli AR. Thalidomide: The Tragedy of Birth Defects and the Effective Treatment of Disease. Toxicol Sci. 2011;122: 1–6. doi:10.1093/toxsci/kfr088

150. Bertini A, Salas R, Chabert S, Sobrevia L, Pardo F. Using Machine Learning to Predict Complications in Pregnancy: A Systematic Review. Front Bioeng Biotechnol. 2022;9: 1–16. doi:10.3389/fbioe.2021.780389

151. Sufriyana H, Husnayain A, Chen YL, Kuo CY, Singh O, Yeh TY, et al. Comparison of multivariable logistic regression and other machine learning algorithms for prognostic prediction studies in pregnancy care: Systematic review and meta-analysis. JMIR Med Informatics. 2020;8: 1–26. doi:10.2196/16503

152. Kumar M, Mostafa J. Research evidence on strategies enabling integration of electronic health records in the health care systems of low- and middle-income countries: A literature review. Int J Health Plann Manage. 2019;34: e1016–e1025. doi:10.1002/hpm.2754

153. Shapiro LM, Kamal RN. Implementation of Electronic Health Records During Global Outreach: A Necessary Next Step in Measuring and Improving Quality of Care. J Hand Surg Am. 2022;47: 279–283. doi:10.1016/j.jhsa.2021.09.016

154. Sobti RC, Rai J, Prakash A. Introduction to Emerging Technologies in Biomedical Sciences. Biomedical Translational Research. Singapore: Springer Nature Singapore; 2022. pp. 1–22. doi:10.1007/978-981-16-4345-3_1

155. Du Y, McNestry C, Wei L, Antoniadi AM, McAuliffe FM, Mooney C. Machine learning-based clinical decision support systems for pregnancy care: A systematic review. Int J Med Inform. 2023;173: 105040. doi:10.1016/j.ijmedinf.2023.105040

156. Sweeting A, Park F, Hyett J. The first trimester: Prediction and prevention of the great obstetrical syndromes. Best Pract Res Clin Obstet Gynaecol. 2015;29: 183–193. doi:10.1016/j.bpobgyn.2014.09.006

157. Poon LC, McIntyre HD, Hyett JA, da Fonseca EB, Hod M. The first-trimester of pregnancy – A window of opportunity for prediction and prevention of pregnancy complications and future life. Diabetes Res Clin Pract. 2018;145: 20–30. doi:10.1016/j.diabres.2018.05.002

158. Bouariu A, Panaitescu AM, Nicolaides KH. First Trimester Prediction of Adverse Pregnancy Outcomes—Identifying Pregnancies at Risk from as Early as 11–13 Weeks. Medicina (B Aires). 2022;58: 332. doi:10.3390/medicina58030332

159. Khanzode KCA, Sarode RD. Advantages and Disadvantages of Artificial Intelligence and Machine Learning: A Literature Review. Int J Libr Inf Sci. 2020;9: 30–36. doi:10.17605/OSF.IO/GV5T4

160. Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. Lancet Oncol. 2019;20: e262–e273. doi:10.1016/S1470-2045(19)30149-4

161. Feduniw S, Golik D, Kajdy A, Pruc M, Modzelewski J, Sys D, et al. Application of Artificial Intelligence in Screening for Adverse Perinatal Outcomes—A Systematic Review. Healthc. 2022;10: 1–20. doi:10.3390/healthcare10112164

162. Ramakrishnan R, Rao S, He JR. Perinatal health predictors using artificial intelligence: A review. Women's Heal. 2021;17. doi:10.1177/17455065211046132

163. Davidson L, Boland MR. Towards deep phenotyping pregnancy: A systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes. Brief Bioinform. 2021;22: 1–29. doi:10.1093/bib/bbaa369

4.2. Machine learning-based models for gestational diabetes mellitus prediction before 24-28 weeks of pregnancy: a review

En esta sección se presentan y discuten los resultados que dieron origen al artículo “Machine learning-based models for gestational diabetes mellitus prediction before 24-28 weeks of pregnancy: a review”, publicado en la revista *Artificial Intelligence in Medicine* (https://doi.org/10.1016/j.artmed.2022.102378).

Machine learning-based models for gestational diabetes mellitus prediction before 24-28 weeks of pregnancy: a review

Daniela Mennickent [1,2,3], Andrés Rodríguez [3,4,5], Marcelo Farías-Jofré [6], Juan Araya [2,3,*], Enrique Guzmán-Gutiérrez [1,3,5,*]

[1] Departamento de Bioquímica Clínica e Inmunología, Facultad de Farmacia, Universidad de Concepción, Chile. [2] Departamento de Análisis Instrumental, Facultad de Farmacia, Universidad de Concepción, Chile. [3] Machine Learning Applied in Biomedicine (MLAB), Concepción, Chile. [4] Departamento de Ciencias Básicas, Facultad de Ciencias, Universidad del Bío-Bío, Chile. [5] Group of Research and Innovation in Vascular Health (GRIVAS-Health), Chile. [6] Departamento de Obstetricia, Facultad de Medicina, Pontificia Universidad Católica de Chile, Chile.

*** Correspondence:** eguzman@udec.cl (EG-G); jarayaq@udec.cl (JA)

Declaration of interest: None.

Abstract

Gestational Diabetes Mellitus (GDM) is a hyperglycemia state that impairs maternal and offspring health, short and long-term. It is usually diagnosed at 24-28 weeks of pregnancy (WP), but at that time the fetal phenotype is already altered. Machine learning (ML)-based models have emerged as an auspicious alternative to predict this pathology earlier, however, they must be validated in different populations before their implementation in routine clinical practice. This review aims to give an overview of the ML-based models that have been proposed to predict GDM before 24-28 WP, with special emphasis on their current validation state and predictive performance. Articles were searched in PubMed. Manuscripts written in English and published before January 1, 2022, were considered. 109 original research studies were selected, and categorized according to the type of variables that their models involved: medical, i.e. clinical and/or biochemical parameters; alternative, i.e. metabolites, peptides or proteins, micro-ribonucleic acid molecules, microbiota genera, or other variables that did not fit into the first category; or mixed, i.e. both medical and alternative data. Only 8.3% of the reviewed models have had validation in independent studies, with low or moderate performance for GDM prediction. In contrast, several models that lack of independent validation have shown a very high predictive power. The evaluation of these promising models in future independent validation studies would allow to assess their performance on different populations, and continue

their way towards clinical implementation. Once settled, ML-based models would help to predict GDM earlier, initiate its treatment timely and prevent its negative consequences on maternal and offspring health.

4.2.1. Introduction

Gestational Diabetes Mellitus (GDM) is a hyperglycemia state that is first diagnosed during pregnancy, with lower blood glucose levels than those considered diagnostic for overt diabetes outside of pregnancy [1]. It has short and long-term negative health consequences for mother and offspring. For example, GDM mothers have a higher risk of presenting preeclampsia in the short term [2], and prediabetes or type 2 diabetes in the long term [3]. Likewise, GDM offspring have an increased risk of presenting macrosomia or neonatal hypoglycemia in the short term [2], and obesity [3] or an impaired glucose tolerance in the long term [4]. Notably, GDM has been related to many other adverse outcomes [5].

The diagnosis of GDM is highly dependent on the guidelines used. Their recommendations mainly differ in the methods employed to screen this pregnancy disorder: whether it should be targeted or universal, by a one-step or a two-step protocol, or even whether it should be performed at the first trimester or at 24-28 weeks of pregnancy (WP). Nevertheless, the most commonly suggested time window to perform GDM diagnosis is between 24 and 28 WP [6–8]. Then, it is

treated with diet management, physical activity and pharmacological therapy if needed. Remarkably, some evidence shows that the fetal phenotype is already altered at the time of GDM diagnosis [9–11], which suggests that the mechanisms that lead to its adverse outcomes start before its clinical manifestation. Therefore, it is necessary to implement methodologies to predict GDM before 24-28 WP. This would allow to initiate its treatment opportunely, to prevent its negative consequences on mother and offspring health, and to reduce the costs that those complications carry. Just in terms of prenatal and postpartum care, GDM is associated with an incremental cost per case equivalent to US$318-US$16,530 [12].

Multiple strategies have been proposed to predict GDM at early pregnancy stages; most of them are univariate, i.e. use a single biomarker or variable. They have been reviewed elsewhere [13–16]. The most common disadvantage of univariate approaches is their modest predictive accuracy. Hence, strategies that combine multiple variables, i.e. multivariate or machine learning (ML)-based approaches, have been proposed to predict GDM with superior accuracy. Indeed, different multivariate models outperformed a single risk factor approach for GDM prediction before 14 WP in a recent prospective multicenter study [17].

It has been said that ML-based models will transform clinical diagnostics [18]. However, before their implementation in clinical practice, they must go through a

process called “validation”. Predictive models typically perform better in the dataset that was used to build them, than when applied to a separate set of samples. This accuracy overestimation, called overfitting, can be corrected by the use of distinct validation methods with different degrees of reliance. Internal validation consists of assessing the model performance by predicting the same samples that were employed to construct it [19] (a.k.a. the training, derivation, modelling or calibration set) through resampling techniques like cross-validation, bootstrapping or permutation. On the other hand, external validation consists of assessing the model performance by predicting samples in a dataset that was not used to train it [19] (a.k.a the testing, prediction or validation set), whether in the same or in an independent study. External validation within the same article is usually performed with a cohort that is similar to the training one, while independent validation ensures that the testing population is different. Since a predictive model may perform well in a particular group of subjects, but poorly in others, it is crucial to assess its accuracy in different populations. In fact, if a predictive model is not validated in different populations and clinical settings, it is not suitable for clinical application [20].

This review aims to give an overview of the ML-based models that have been proposed in literature to predict GDM before 24-28 WP, with special emphasis on their current validation state and predictive performance.

4.2.2. Methodology

4.2.2.1. Search strategy

Articles were searched in PubMed using different combinations of the following terms: “gestational diabetes”, “early” or “first trimester”, “detection” or “diagnosis” or “prediction” or “screening”, “machine learning” or “multivariate analysis” or “multivariable analysis” or “multimarker analysis”. Studies written in English and published before January 1, 2022, were considered. Manuscripts were excluded if did not assess the predictive power of their models by at least one of the following parameters: area under ROC curve (AUC), accuracy, specificity, sensitivity, or the predictive squared correlation coefficient Q^2. Their definition can be found elsewhere [21,22]. Review and independent validation articles references were checked to find any other relevant publication.

4.2.2.2. Information synthesis

Original research articles were divided into three categories, according to the type of variables involved in their models: medical, alternative and mixed. Medical variables-based models included parameters that are typically registered or measured on healthcare centers, and therefore are easy to obtain in regular pregnancy controls. These models comprised clinical (i.e. maternal personal data,

medical history, anthropometric features, demographic information and ultrasonographic measurements) and biochemical data (i.e. markers that are quantified in maternal biological fluids, in the context of routine clinical practice). Alternative variables-based models covered parameters that are not usually registered or measured on healthcare centers, and therefore are not easy to obtain in regular pregnancy controls. These models encompassed metabolites, peptides or proteins, microbiota genera, or any other kind of variable that did not belong to the first category. Mixed variables-based models combined both medical and alternative data.

To summarize the information, a table for each category was constructed as follows: for each article, the ML-based model with the highest predictive power, i.e. the best model, was selected. For each model, specific kind, number and name of variables; prediction time, predictive power and validation type were registered. Three validation types were considered, from the lowest to the highest quality: internal, external and independent. Internal validation was considered as the evaluation of the model performance within the original article by predicting the same samples that were employed to construct the model, through resampling techniques. External validation, as the assessment of the model performance within the same manuscript by predicting samples that were not used to build the model. Independent validation, as the evaluation of the model performance in an independent study by predicting samples that were not used to construct the

model. For tabulation, external validation predictive power parameters were prioritized, followed by the internal validation ones. Thus, if an article presented unvalidated, internal validation and external validation AUCs, we only tabulated the last value. If they reported unvalidated and internal validation AUCs, we only tabulated the last one. If they only informed unvalidated AUC, we tabulated that value. Independent validation parameters were not tabulated.

To give a semiquantitative interpretation of predictive power, we used the following notation: “very low” for AUC 0.5-0.6, accuracy 50%-60% and $Q^2 < 0.5$; “low” for AUC 0.6-0.7; “moderate” for AUC 0.7-0.8; “high” for AUC 0.8-0.9; and “very high” for AUC 0.9-1.0 and accuracy 90%-100%.

4.2.3. Results

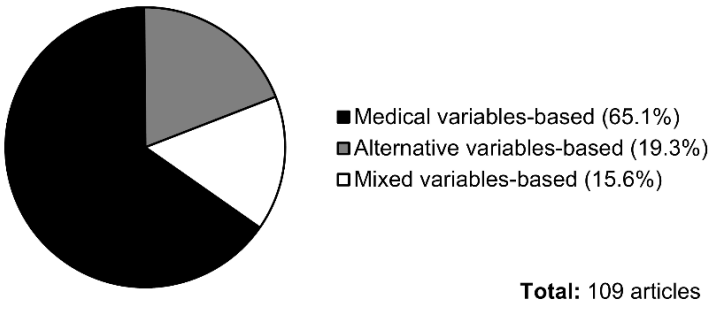
4.2.3.1. Selected articles description

116 articles were selected, including 109 original research manuscripts, five independent validation studies and two review papers. **Figure 4.2-1** presents the distribution of the original research articles by the type of variables their models used, by their publication year and by the higher quality validation type that their best models had. As shown in **Figure 4.2-1A**, of all manuscripts, 71 (65.1%) based their models on medical variables, 21 (19.3%) on alternative parameters

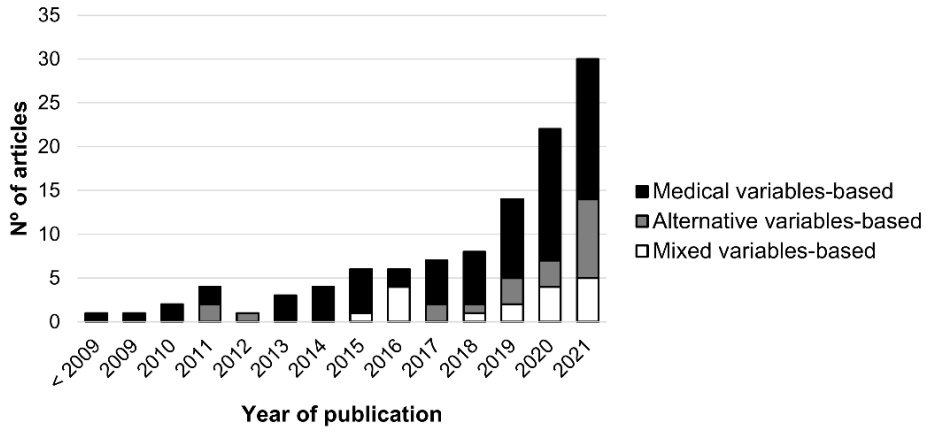
and 17 (15.6%) on mixed data. According to **Figure 4.2-1B**, the first models in the field were based on medical variables, while the last to appear were the ones built on mixed data. The number of total articles has increased over time. However, **Figure 4.2-1C** shows that more than half of the manuscripts, 61 out of 109, did not validate their best GDM predictive models, while 19 validated them only by resampling techniques. Moreover, even though 20 articles performed external validation on their best models, they have had no independent validation. This means that approximately one-third of the reviewed models, 39 out of 109, have only been validated within their original manuscripts. Furthermore, only nine articles reported models which were then independently validated. This was done in at least one of the five independent validation articles that we found [21–25]. Interestingly, all the models that have had independent validation are based on medical variables. The two review articles that we found are also related to this type of variable. One of them focuses on clinical data-based multivariate models for GDM prediction and risk estimation, and considers 14 articles published before December 2014 [26]; while the other one, issued in 2018, contains a table that summarizes 12 multivariate models for GDM early prediction, mainly built on clinical and biochemical data [27].

Hereunder, we will comment on the best GDM predictive models of each category, with special emphasis on those with higher quality validation and predictive power.

(a)


Medical variables-based (65.1%)
Alternative variables-based (19.3%)
Mixed variables-based (15.6%)
Total: 109 articles

(b)


35
30
25
20
15
10
5
0
Nº of articles
< 2009
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
Year of publication
Medical variables-based
Alternative variables-based
Mixed variables-based

(c)

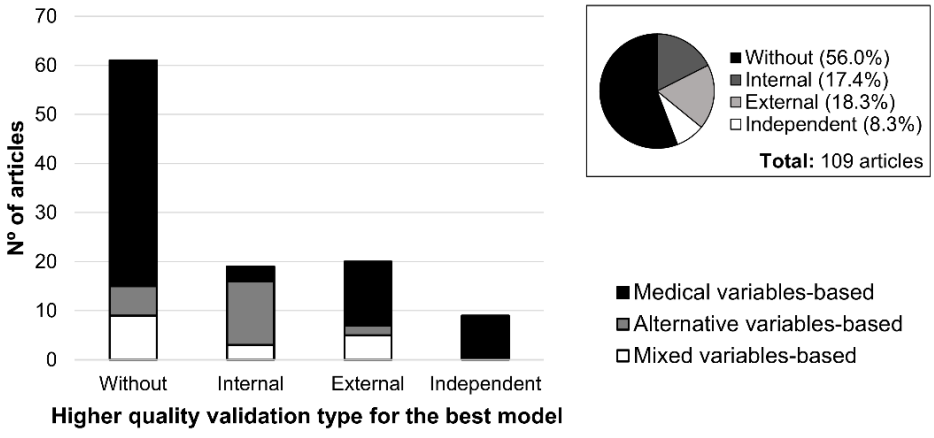

70
60
50
40
30
20
10
0
Nº of articles
Without
Internal
External
Independent
Higher quality validation type for the best model
Without (56.0%)
Internal (17.4%)
External (18.3%)
Independent (8.3%)
Total: 109 articles
Medical variables-based
Alternative variables-based
Mixed variables-based

Figure 4.2-1. Characteristics of the 109 original research articles selected for this review. A) Categorization of the original articles by the type of variables that they use in their GDM predictive ML-based models: medical (black), alternative (grey) or mixed (white). B) Number of articles published by year for the three aforementioned categories. C) Number of articles organized by the higher quality validation type that their best models have had for the same three categories. Inset: General distribution of the original articles by the higher quality validation type that their best models have had: without (black), internal (dark gray), external (light gray) or independent (white).

4.2.3.2. Medical variables-based models

Models of this category are based on clinical [28–43] or biochemical [44–53] variables, or on a combination of both [54–98]. These strategies are summarized on **Table 4.2-1**.

4.2.3.2.1. Clinical variables-based models

These strategies are based on maternal personal information, medical history, anthropometric parameters, demographic data and ultrasound measurements. The majority of them presented a moderate (AUC 0.7-0.8) predictive power in their original articles, although some showed a high (AUC 0.8-0.9) one. Many of these models have had independent validation, but with low (AUC 0.6-0.7) or moderate (AUC 0.7-0.8) performance in independent cohorts [21–25]. Two exceptions are Van Leeuwen et al. [31] and Teede et al. [34] models, which allowed to predict GDM cases requiring insulin therapy with high (AUC 0.8-0.9) predictive power in an independent study. The first approach combines maternal

body mass index (BMI), family history of diabetes mellitus (DM), prior GDM and ethnicity; and the second employs the same five variables plus maternal age. In their original articles, they yielded an unvalidated AUC of 0.77 [31] and an external validation AUC of 0.703 [34] respectively, while in the independent validation study they allowed to predict GDM needing pharmacotherapy before 20 WP with AUCs of 0.824 and 0.812 respectively [22].

Other studies of this kind presented models with high (AUC 0.8-0.9) predictive power, but without any form of validation. For instance, Sirico et al. predicted GDM at 11-14 WP with maternal age, BMI, cigarette smoking and four variables derived from ultrasound exams: nuchal translucency, crown-rump length, ductus venosus peak velocity index and fetal heart rate (FHR). Their unvalidated model achieved an AUC of 0.825, slightly higher than the obtained for FHR solo [38]. Similarly, Cremona et al. combined family history of DM, weight, previous perinatal death, overall insulin resistant condition, and four maternal ultrasonographic parameters to predict GDM at 10-16 WP. Their model yielded an unvalidated AUC of 0.860, better than the obtained for each variable on its own [41].

4.2.3.2.2. Biochemical variables-based models

This type of strategy is based on biomarkers that are measured in maternal biological fluids; typically, in blood. Some of them have a low (AUC 0.6-0.7)

performance, while others have a very high (AUC 0.9-1.0) one. None of them have had independent validation, and only one was validated in its original article, with very high predictive power: Xiong et al. combined two coagulation variables, prothrombin time and reference activated partial thromboplastin time, to predict GDM before 19 WP, and got an external validation AUC of 0.998 [46]. The other very high predictive power models of this kind have had no validation. For example, Amini et al. used serum beta-human chorionic gonadotropin, unconjugated estriol and alfa-fetoprotein to predict GDM at 14-17 WP. Single biomarkers yielded unvalidated AUCs of 0.62, 0.65 and 0.58 respectively, while their combination increased it to 0.92 [48]. Likewise, Rasanen et al. used serum high-sensitivity C-reactive protein, adiponectin, glycosylated fibronectin (GlyFn) and placental lactogen to predict GDM at 5-13 WP. They reached an unvalidated AUC of 0.919, slightly higher than the one for GlyFn solo [49].

4.2.3.2.3. Clinical and biochemical variables-based models

These strategies combine clinical and, typically, blood biochemical maternal data; and are the most common to attempt GDM early prediction. They have varying predictive power, from a low (AUC 0.6-0.7) to a very high (AUC 0.9-1.0) one. Only one model of this kind has been independently validated, and displayed a moderate (AUC 0.7-0.8) predictive power: Benhalima et al. used maternal age, family history of DM, prior GDM, ethnicity, weight, height, fasting glucose, A1c

glycated hemoglobin and triglycerides to predict GDM at 6-14 WP. This model got an internal validation AUC of 0.72 in its original article [84], and an AUC of 0.769 in a recent independent validation study [25].

Several models of this type were externally validated within their original articles, however, only some of them had a high (AUC 0.8-0.9) performance. For instance, Wang et al. combined ten clinical and three biochemical maternal parameters to predict GDM at 14-20 WP. They reported AUCs of 0.845 and 0.886 for two different external validation cohorts [92]. Another example is the study of Qiu et al., who predicted GDM before 13 WP with 50 medical variables from electronic health records (EHR). They achieved an external validation AUC of 0.893 [94]. Similarly, Wu et al. used 73 medical parameters from EHR and predicted GDM before 12 WP with an external validation AUC of 0.8 [95]. Likewise, Artzi et al. model combined 2355 variables from EHR, and accomplished GDM prediction before 20 WP with future, geographical and geo-temporal external validation AUCs of 0.854, 0.875 and 0.863 [98].

Some of these strategies have shown a very high (AUC 0.9-1.0) predictive power, but have never been validated. For example, Liu et al. used maternal BMI, fasting glucose and putrescine to predict GDM at 8-12 WP. Their unvalidated model got an AUC of 0.951, with a specificity and a sensitivity of 91.5% and 89.4%, respectively [64]. Another case is the study of Tenenbaum-Gavish et al., who used

obesity, placental and inflammatory biomarkers to predict GDM at 11-13 WP. For obese women, the combination of BMI, insulin, soluble cluster of differentiation 163 (sCD163) and tumor necrosis factor alpha (TNFα) reached an unvalidated AUC of 0.95. For non-obese women, insulin, sCD163, TNFα, pregnancy associated plasma protein A and placental protein 13 yielded an unvalidated AUC of 0.94. The addition of more parameters to these models, among them the ultrasonic variable uterine artery pulsatility index, did not improve their performance [69]. Likewise, Sweeting et al. combined eight clinical and two biochemical variables to predict GDM at 11-13 WP. For their general cohort, they got an unvalidated AUC of 0.90. Furthermore, they assessed their model in different subpopulations: by ethnicity, it performed better for South Asians, with an AUC of 0.92; by parity, for multiparous, with an AUC of 0.94; and, by time of diagnosis, for early GDM with an AUC of 0.96 [88].

Table 4.2-1. Original articles that report ML-based models for GDM early prediction based on medical variables.

Type of variables	Article	Characteristics of the most accurate ML-based model reported for GDM early prediction										
		Nº of variables	Variables	Prediction time [a]	Predictive power				Validation		
					AUC	Ac	Sp	Se	IV [b]	EV [c]	XV [d]
Clinical	[28]	2	Age, BMI	5-19	0.71	-	-	-	-	-	X
	[29]	3	Age, BMI, ethnicity	< 24	-	-	84.0	82.6	-	X	X
	[30]	3	Age, BMI, TATT	11-14	0.78	-	90	42	-	-	-
	[31]	4	BMI, family history of DM, prior GDM, ethnicity	< 20	0.77	-	-	-	-	-	X
	[32]	4	Age, gravidity, PV, VFI	11-13	0.761	-	82.9	62.3	-	-	-
	[33]	5	Age, BMI, family history of DM, prior macrosomia, prior spontaneous abortions	< 14	0.752	-	49.5	83.6	-	X	X
	[34]	5	Age, BMI, family history of DM, prior GDM, ethnicity	12-15	0.703	-	71.4	61.3	-	X	X
	[35]	5	Age, BMI, prior GDM, ethnicity, SBP	11-14	0.821	-	-	-	-	-	X
	[36]	5	Age, BMI, family history of DM, prior hypertension, expected payer for delivery	PP	0.74	-	-	-	-	X	-
	[37]	6	Age, BMI, family history of DM, prior GDM, ethnicity, parity	11-13	0.88	-	80.0	80.9	-	-	X
	[38]	7	Age, BMI, smoking, NT, CRL, DV-PVI, FHR	11-14	0.825	-	80.0	66.3	-	-	-
	[39]	7	Age, BMI, family history of DM, gravidity, PCOS, abdomen circumference, irregular menstruation	< 14	0.777	78.9	81.3	65.1	X	X	-
	[40]	8	Age, family history DM, prior GDM, ethnicity, weight, height, conception method, neonate birth weight in last pregnancy	11-13	0.823	-	80	68	X	-	X
	[41]	8	Family history of DM, weight, previous perinatal death, overall insulin resistant condition, abdominal SFT, abdominal VFT, 8-point SKT, MUAC	10-16	0.860	-	-	-	-	-	-
	[42]	12	Age, BMI, family history of DM, prior GDM, prior PE, prior abnormal pregnancies, prior hypertension, parity, smoking, drinking, education, civil status	PP	0.739	-	-	-	-	-	-

	[43]	14	Age, first degree family history of DM, second degree family history of DM, prior GDM, no prior GDM, Black ethnicity, East Asian ethnicity, South Asian ethnicity, weight, height, prior LGA, ovulation drugs, monochorionic twins, dichorionic twins	11-13	-	-	80.0	58.0	-	-	-
Biochemical	[44]	2	hs-CRP, SHBG	< 15	-	75.5	75.6	74.1	-	-	-
	[45]	2	Glucose, GCF PlGF	11-14	0.898	-	75.0	78.6	-	-	-
	[46]	2	PAT-PT, REF-APTT	< 19	0.998	-	99.2	92.5	-	X	-
	[47]	3	Fasting glucose, non-HDL-C, SHBG	< 15	0.73	-	44	80	-	-	-
	[48]	3	β-hCG, uE3, AFP	14-17	0.92	-	86	94	-	-	-
	[49]	4	hs-CRP, adiponectin, GlyFn, hPL	5-13	0.919	-	-	-	-	-	-
	[50]	4	Fasting glucose, TG, TC, LDL-C	8-12	0.643	-	82.6	38.5	-	-	-
	[51]	4	Adiponectin, chemerin, leptin, sFRP4	10-13	0.699	-	-	-	-	-	-
	[52]	5	Fasting glucose, TG, TC, TG/HDL-C, LDL-C/HDL-C	< 14	0.717	-	-	-	-	-	-
	[53]	6	TC, TG, HOMA-IR, LDL-C, t-PA, insulin	< 14	0.87	-	80.0	81.3	-	-	-
Clinical and biochemical	[54]	2	BMI, fasting glucose	8-12	0.689	-	67.8	64.7	-	-	-
	[55]	2	BMI, afamin	11-14	0.668	-	-	-	-	-	-
	[56]	2	BMI, fasting glucose	< 13	0.83	-	80.6	70.7	-	-	-
	[57]	2	BMI, PAPP-A	11-14	0.827	-	90	55	-	-	-
	[58]	2	BMI, fasting glucose	6-16	0.71	-	81	51	-	-	-
	[59]	2	BMI, fasting glucose	6-10	0.85	-	-	-	-	-	-
	[60]	3	Age, weight, PlGF	11-14	0.78	-	-	-	-	-	-
	[61]	3	BMI, family history of DM, fasting glucose	< 13	-	-	40	89	-	-	-
	[62]	3	Age, weight, IL-6	11-14	0.759	-	75.0	64.9	-	-	-
	[63]	3	BMI, adiponectin, PAPP-A	12-13	0.877	-	90.0	72.7	-	-	-
	[64]	3	BMI, fasting glucose, putrescine	8-12	0.951	-	91.5	89.4	-	-	-
	[65]	4	Age, BMI, family history of DM, PAPP-A	11-14	0.7	64.6	50.5	81.4	-	-	-
	[66]	4	Age, NT, PAPP-A, β-hCG	11-13	0.531	-	-	-	-	-	-
	[67]	4	Age, BMI, fasting glucose, TG	8-20	0.766	64	75	66	-	-	-
	[68]	4	Age, BMI, family history of DM, fasting glucose	9-13	0.7	-	-	-	X	X	-

[69]	4	BMI, insulin, sCD163, TNFα	11-13	0.95	-	90	89	-	-	-
[70]	4	Age, BMI, HbA1c, TG	7-12	0.728	-	65.2	71.6	X	-	-
[71]	4	Age, TG, afamin, platelet/lymphocyte ratio	10-12	0.748	70.2	68.3	69.2	-	-	-
[72]	5	Age, BMI, ethnicity, insurance, HbA1c	≤ 16	0.75	-	-	-	-	-	-
[73]	6	Age, parity, ethnicity, SBP, DBP, adiponectin	15-17	0.834	-	-	-	-	-	-
[74]	6	Age, BMI, family history of DM, prior GDM, ethnicity, HMW-adiponectin	12-15	0.847	-	96.2	50.0	-	-	-
[75]	6	Age, BMI, family history of DM, height, SBP, ALT	< 15	0.71	-	44.8	82.1	-	X	-
[76]	7	Age, BMI, prior GDM, prior macrosomia, ethnicity, adiponectin, SHBG	11-13	0.865	-	80.0	77.9	-	-	-
[77]	7	Age, BMI, ethnicity, parity, smoking, adiponectin, leptin	6-14	0.824	-	80.0	71.2	-	-	-
[78]	7	Age, BMI, family history of DM, season of conception, gestational week at sampling, fasting glucose, T3	< 13	0.726	-	90.0	30.8	-	-	-
[79]	8	Age, BMI, family history of DM, prior GDM, ethnicity, gestational age at sampling, HDL-C, t-PA	11-13	0.861	-	-	-	-	-	-
[80]	8	BMI, family history of DM, prior GDM, ethnicity, parity, TG, PAPP-A, lipocalin-2	11-13	0.93	-	-	-	-	-	-
[81]	8	Age, BMI, family history of DM, ethnicity, prior hypertension, PAPP-A, uE3, INH	10-20	0.722	-	-	-	-	X	-
[82]	8	BMI, folic acid intake, ALT, WBC, platelet count, albumin, bilirubin, creatinine	8-12	0.87	-	-	-	-	-	-
[83]	9	Age, BMI, family history of DM, prior macrosomia, parity, fasting glucose, GGT, HBsAg, ANGPTL8	12-16	0.772	-	-	-	-	-	-
[84]	9	Age, family history of DM, prior GDM, ethnicity, weight, height, fasting glucose, HbA1c, TG	6-14	0.72	-	-	-	X	-	X
[85]	9	Age, BMI, family history of DM, ethnicity, glucose, HMW-adiponectin, chemerin, FABP4, sOB-R	10-14	0.77	-	-	-	X	-	-
[86]	10	Age, BMI, family history of GDM, prior GDM, ethnicity, glucose, TG, TC, HDL-C, LDL-C	12-15	0.83	-	92.4	50.0	-	-	-
[87]	10	Age, family history DM, prior GDM, ethnicity, weight, height, conception method, neonate birth weight in last pregnancy, PAPP-A, PlGF	11-13	0.869	-	80	77	-	-	-

[88]	10	Age, BMI, family history of DM, prior GDM, parity, ethnicity, MAP, UtA-PI, PAPP-A, β-hCG	11-13	0.96	-	-	-	-	-	-
[89]	10	Age, BMI, parity, fasting glucose, Hb1Ac, TG, TPOAb, percentage of B lymphocytes, immunoglobulin A, progesterone	12-16	0.772	-	-	-	X	X	-
[90]	12	Age, BMI, family history of DM, prior GDM, prior macrosomia, ethnicity, pre-pregnancy exercise, pre-pregnancy soft drink intake, pregnancy (before GDM diagnosis) soft drink intake, HbA1c, hs-CRP, SHBG	14-17	0.895	-	90.0	70.5	X	-	-
[91]	13	Age, BMI, family history of DM, gravidity, weight gain, waist circumference, hip circumference, education, income, SBP, DBP, fasting glucose, ALT	< 15	0.742	75.7	76.9	61.6	X	X	-
[92]	13	Age, BMI, family history of DM, prior GDM, height, prior hypertension, PCOS, menstrual history, prior cesarean delivery, education, fasting glucose at ≤ 13 WP, fasting glucose at 14-20 WP, TG at 14–20 WP	14-20	0.845, 0.886	-	70.1, 72.5	79.2, 86.4	-	X	-
[93]	14	Age, BMI, family history of DM, prior GDM, PCOS, SBP, SFT, VFT, HbA1C, TG, TC, LDL-C, hs-CRP, visfatin	12-16	0.911	-	94.7	83.0	-	-	-
[94]	50	Among them: age, BMI, parity, education, fasting glucose, ALT, AST, WBC, RBC, TPOAb, urea, uric acid	< 13	0.893	-	-	-	X	X	-
[95]	73	Among them: family history of DM, prior GDM, multiple pregnancy, fasting glucose, HbA1c, TG	< 12	0.80	-	82	63	-	X	-
[96]	104	Among them: age, BMI, family history of DM, prior GDM, prior macrosomia, prior PTB, parity, ultrasonic screening data, fasting glucose, HbA1c, TG	16-20	0.74	-	99	15	X	X	-
[97]	Hundreds	Among them: age, prior GDM, parity, religion, glucose, HbA1c, HbA2, TG, TC, TSH, free T4, total T3, progesterone	PP	-	97.9	99.3	96.5	-	-	-
[98]	2355	Among them: age, BMI, family history of DM, prior pregnancy complications, prior hypertension, prior heart disease, PCOS, prior glucose issues, prior hypercholesterolemia, prior pregnancy GCT, glucose	< 20	0.854, 0.875, 0.863	-	-	-	X	X	-

Abbreviations. AUC: area under the ROC curve; Ac: accuracy; Sp: specificity; Se: sensitivity; IV: internal validation; EV: external validation; XV: independent validation; BMI: body mass index; TATT: total adipose tissue thickness; DM: diabetes mellitus; GDM: gestational diabetes mellitus; PV: placental volume; VFI: vascularization-flow index; SBP: systolic blood pressure; NT: nuchal translucency; CRL: crown-rump length; DV-PVI: ductus venosus peak velocity index; FHR: fetal heart rate; PCOS: polycystic ovary syndrome; SFT: subcutaneous fat thickness; VFT: visceral fat thickness; SKT: skinfold thickness; MUAC: mid-upper arm circumference; PE: preeclampsia; LGA: large for gestational age; hs-CRP: high sensitivity C-reactive protein; SHBG: sex hormone binding globulin; GCF: gingival crevicular fluid; PlGF: placental growth factor; PAT-PT: prothrombin time; REF-APTT: reference activated partial thromboplastin time; HDL-C: high density lipoprotein cholesterol; β-hCG: beta human chorionic gonadotropin; uE3: unconjugated estriol; AFP: alfa fetoprotein; GlyFn: glycosylated fibronectin; hPL: placental lactogen; TG: triglycerides; TC: total cholesterol; LDL-C: low density lipoprotein cholesterol; sFRP4: soluble frizzled-related protein 4; HOMA-IR: homeostasis model assessment of insulin resistance; t-PA: tissue plasminogen activator; PAPP-A: pregnancy associated plasma protein A; IL-6: interleukin 6; sCD163: soluble cluster of differentiation 163; TNFα: tumor necrosis factor alpha; HbA1c: A1c glycated hemoglobin; DBP: diastolic blood pressure; HMW: high molecular weight; ALT: alanine transaminase; T3: triiodothyronine; INH: dimeric inhibin A; WBC: white blood cell count; GGT: gamma glutamyl transferase; HBsAg: hepatitis B surface antigen; ANGPTL8: angiopoietin like protein 8; FABP4: fatty acid binding protein 4; sOB-R: soluble leptin receptor; MAP: mean arterial pressure; UtA-PI: uterine artery pulsatility index; TPOAb: antithyroid peroxidase autoantibody; WP: weeks of pregnancy; AST: aspartate transaminase; RBC: red blood cell count; PTB: preterm birth; HbA2: hemoglobin A2; TSH: thyroid stimulating hormone; T4: thyroxine; GCT: glucose challenge test; PP: prepregnancy.

[a] Prediction time is in gestational weeks unless “PP” is stated.

[b] Internal validation (IV) was considered as the assessment of the best model performance with the same samples that were used to train it and resampling techniques, within the same article that originally reported the model.

[c] External validation (EV) was considered as the assessment of the best model performance with samples that were not used to train it, within the same article that originally reported the model.

[d] Independent validation (XV) was considered as the assessment of the best model performance with samples that were not used to train it, in an independent study.

4.2.3.3. Alternative variables-based models

Some articles do not use medical data for their GDM predictive models, but employ alternative variables, which may derive from metabolomics [99–108], proteomics [109,110], micro-ribonucleic acid (miRNA) [111–113], cell-free deoxyribonucleic acid (cfDNA) [114], genetic variants [115–117] or microbiota [118] analysis, or other methods [119]. These strategies are summarized in **Table 4.2-2**. Notably, none of them have had independent validation.

4.2.3.3.1. Metabolomics variables-based models

In this case, variables are metabolites. Some of these models have a very low (Q^2 < 0.5) predictive power, while others have a very high (AUC 0.9-1.0) one. None of them had external validation in their original articles, although various had internal validation. Only two internally validated models had a high (AUC 0.8-0.9) performance. The first one was proposed by Liu et al., who applied urine targeted liquid chromatography-mass spectrometry (LC-MS) metabolomics to predict GDM at 6-10 WP. Using a panel of three non-polar metabolites, l-phenylalanyl-l-proline, hydroxylauroylcarnitine and levoglucosan, they yielded an AUC of 0.89 in internal validation [100]. The second one was reported by McMichael et al. They predicted GDM in overweight and obese pregnant women at 10-16 WP with plasma targeted LC-MS metabolomics data. The combination of α-hydroxybutyrate,

sphingomyelin 14:0, xanthine and hypoxanthine reached an internal validation AUC of 0.833, higher than the obtained for each metabolite on its own [103].

On the other hand, there is a model of this kind that showed a very high (AUC 0.9-1.0) predictive power, but have not been validated: Koos et al. assessed urine untargeted LC-MS and gas chromatography-mass spectrometry (GC-MS) metabolomics to predict GDM at 6-19 WP. Their model, which included seven metabolites, got an unvalidated AUC of 0.993 [104].

4.2.3.3.2. Proteomics variables-based models

Models of this type use proteins or peptides as variables. They have a very high (AUC 0.9-1.0) predictive power, however, none of them had validation in their original articles. For example, Guo et al. used urine isobaric tag for relative and absolute quantitation (iTRAQ) LC-MS to predict GDM in ≤ 35 years pregnant women. They identified 119 differential proteins between GDM and non-GDM women, two of which were verified by enzyme-linked immunosorbent assay (ELISA): cluster of differentiation 59 and interleukin 1 receptor antagonist. Single proteins allowed to predict GDM at 15-20 WP with unvalidated AUCs of 0.729 and 0.899 respectively, while their combination yielded an AUC of 0.906 [109]. Similarly, Zhao et al. used serum iTRAQ LC-MS to predict GDM at 12-16 WP. They identified 33 differentially expressed proteins, four of which were verified by

ELISA: apolipoprotein E, coagulation factor IX, fibrinogen alpha chain and insulin-like growth factor-binding protein 5. Their combination achieved an unvalidated AUC of 0.985, higher than the ones for single proteins [110].

4.2.3.3.3. miRNA variables-based models

In this case, variables are miRNA molecules. Three articles based their models on this type of data, and only one of them performed external validation, with a low (AUC 0.6-0.7) predictive power: Zhao et al. used a miRNA panel and reverse transcription-quantitative polymerase chain reaction (RT-qPCR) to analyze serum samples, and found three miRNAs which expression was significantly decreased in GDM pregnant women, miR-132, miR-29a and miR-222. Single miRNAs achieved GDM prediction at 16-19 WP with external validation AUCs of 0.642, 0.658 and 0.600 respectively, while their combination reached an AUC of 0.669 [112].

The other two articles of this kind reported internally validated models with a high (AUC 0.8-0.9) or a very high (AUC 0.9-1.0) predictive power: Yoffe at al. employed a miRNA panel and RT-qPCR to study plasma samples, and found two upregulated miRNAs in GDM women, miR-223 and miR-23a. Individual miRNAs predicted GDM at 9-11 WP with internal validation AUCs of 0.94 and 0.89 respectively, while their combination reduced it to 0.91 [111]. Similarly, Nair et al.

analyzed plasma extracellular vesicle-associated miRNAs by RT-qPCR. The combination of nine of them allowed to predict GDM before 18 WP with an AUC of 0.80 in internal validation, higher than the obtained for each miRNA solo [113].

4.2.3.3.4. cfDNA variables-based models

The variables that these approaches use derive from cfDNA analysis. The only article of this kind reported a high (AUC 0.8-0.9) predictive power and internally validated model: in order to predict GDM at 12-17 WP, Del Vecchio et al. used a novel deconvolution methodology, based on cfDNA methylation data, to assess different cfDNA sources in plasma. Total cfDNA was deconvoluted into seven sources: liver, pancreas, heart, placenta, T-cells, B-cells and neutrophils. The combination of placenta and pancreas cfDNA achieved an internal validation AUC of 0.8 [114].

4.2.3.3.5. Genetic variants variables-based models

This kind of models use genetic variants as variables. Three articles presented such models, with very low (accuracy 50%-60%) or low (AUC 0.6-0.7) predictive power. Two of them performed validation, one externally, and the other one internally. The former is from Lamri et al., who assessed different polygenic risk scores to predict GDM in South Asian women. They used data from the

genotyping of blood samples with polymerase chain reaction (PCR) and DNA-arrays. Their best model got a low performance AUC, of 0.65 in an external validation cohort [117]. That model comprised 1,290,525 single nucleotide polymorphisms (SNPs) that had previously been tested in a genome-wide association meta-analysis (GWAMA) for their connection with type 2 diabetes. Those SNPs included: *TMEM154* rs6813195, *SSR1-RREB1* rs9505118, *FAF1* rs17106184, *POU5F1-TCF19* rs3130501, *LPP* rs6808574, *ARL15* rs702634 and *MPHOSPH9* rs4275659 [120]. The latter is from Yu et al., who studied the association between the likelihood of GDM in a southern Chinese population and three SNPs of the retinoid X receptor (RXR): *RXR-α* rs4842194 G>A, *RXR-γ* rs100537 A>G and *RXR-γ* rs2134095 T>C. The authors used blood samples and PCR plus restriction fragment length polymorphism for genotyping. They found a significant association between rs2134095 and the risk of GDM, and then tested different combinations of SNPs for the prediction of this pregnancy complication. The three reported models displayed a very low predictive accuracy. The best one combined rs2134095 and rs100537, and reached an internal validation accuracy of 0.552 [115].

4.2.3.3.6. Microbiota variables-based models

In this case, variables are microbiota genera. The only article of this type reported a moderate (AUC 0.7-0.8) predictive power and internally validated model: Zheng

et al. analyzed gut microbiota to predict GDM at 8-12 WP. By 16S ribosomal ribonucleic acid (rRNA) gene amplicon sequencing, they found that the relative abundance of 10 microbiota genera differed significantly between GDM and non-GDM women. The combination of three of them, Coprococcus, Intestinimonas, and Veillonella, yielded an internal validation AUC of 0.743 [118].

4.2.3.3.7. Other variables-based models

Models of this type use variables that do not derive from any of the aforementioned methods. Only one article reported such a model, with a very high (accuracy 90%-100%) predictive power and internal validation: Boisvert et al. applied amniotic fluid capillary electrophoresis to predict GDM at 12-20 WP. Their model included two electropherogram wavelets, one with no clear identity and the other assigned to albumin, and reached internal validation specificity and sensitivity of 99% and 86%, respectively.

Table 4.2-2. Original articles that report ML-based models for GDM early prediction based on alternative variables.

Type of variables	Article	Characteristics of the most accurate ML-based model reported for GDM early prediction											
		Nº of variables	Variables	Technique	Matrix	Prediction time [a]	Predictive power					Validation	
							AUC	Ac	Sp	Se	Q^2	IV [b]	EV [c]
Metabolomics	[99]	2	17(S)-HDoHE, sebacic acid	LC-MS	Plasma	12-16	0.708	-	-	-	-	-	-
	[100]	3	l-Phe-l-Pro, hydroxylauroylcarnitine, levoglucosan	LC-MS	Urine	6-10	0.89	-	86.7	86.7	-	X	-
	[101]	3	α-HBA, β-HBA, myristic acid	GC-MS	Serum	8–14	0.791	-	68.7	84.6	-	X	-
	[102]	4	Glucose, small size HDL-C, GlycA, Leu	NMR	Serum	< 18	0.719	-	-	-	-	-	-
	[103]	4	α-HB, SM 14:0, xanthine, hypoxanthine	LC-MS	Plasma	10-16	0.833	-	-	-	-	X	-
	[104]	7	Argininate, dihydroorotate, lanthionine, 7,8-dihydroneopterin, nicotinate, ribonucleoside, phenylglucuronide	LC-MS and GC-MS	Urine	6-19	0.993	96.7	95.7	97.8	-	-	-
	[105]	9	Hippurate, trigonelline, kynurenine, 3-hydroxykynurenine, 1-methylnicotinamide, Thr, Ile, Trp, Val	NMR	Urine	≤ 12	0.796	-	79.2	78.6	-	X	-
	[106]	263	Among them: ethanolamine, 1,3-diphosphoglycerate	LC-MS	Urine	< 16	-	-	-	-	0.483	X	-

	[107]	NI	Among them: TMAO, betaine	NMR	Plasma	16-22	-	-	-	-	0.278	X	-
	[108]	NI	Among them: L-Glu, xanthine	LC-MS	Serum	10-13	-	-	-	-	0.24	X	-
Proteomics	[109]	2	CD59, IL1RA	LC-MS	Urine	15-20	0.906	-	-	-	-	-	-
	[110]	4	ApoE, F9, FGA, IGFBP5	LC-MS	Serum	12-16	0.985	-	95	80	-	-	-
Micro RNA	[111]	2	miR-223, miR-23a	miRNA array and RT-PCR	Plasma	9-11	0.91	90	80	94	-	X	-
	[112]	3	miR-132, miR-29a, miR-222	miRNA array and RT-PCR	Serum	16-19	0.669	-	63.3	66.7	-	-	X
	[113]	9	miR-let7i-5p, miR-10a-5p, miR-151b, miR-16–2-3p, miR-16-5p, miR-1910-5p, miR-423-5p, miR-92a-3p, miR-92b-3p	RT-PCR	Plasma	< 18	0.80	82	-	-	-	X	-
Cell free DNA	[114]	2	Placenta cfDNA, pancreas cfDNA	PCR-Seq	Plasma	12-17	0.8	-	-	-	-	X	-
Genetic variants	[115]	2	rs2134095, rs100537	PCR-RFLP	Blood	PP	-	0.552	-	-	-	X	-
	[116]	34	Among them: rs5015480, rs4402960, rs7936247, rs4506565	PCR-Sqn and PCR-TaqMan	Blood	PP	0.61	-	-	-	-	-	-
	[117]	1,290,525	Among them: rs6813195, rs9505118, rs17106184, rs3130501, rs6808574, rs702634, rs4275659	PCR-DNA array	Blood	PP	0.65	-	-	-	-	-	X
Microbiota	[118]	3	Coprococcus, Intestinimonas, Veillonella	PCR-Seq	Stool	8-12	0.743	-	-	-	-	X	-

Others	[119]	2	Unassigned wavelet, albumin peak	CE	Amniotic fluid	12-20	-	-	99	86	-	X	-	

Abbreviations. AUC: area under the ROC curve; Ac: accuracy; Sp: specificity; Se: sensitivity; IV: internal validation; EV: external validation; NI: not informed; HDoHE: hydroxydocosahexaenoic acid; Phe: phenylalanine; Pro: proline; HBA: hydroxybutyric acid; HDL-C: high density lipoprotein cholesterol; GlycA: glycoprotein acetylation; Leu: leucine; HB: hydroxybutyrate; SM: sphingomyelin; Thr: threonine; Ile: isoleucine; Trp: tryptophan; Val: valine; TMAO: trimethylamine N-oxide; Glu: glutamic acid; CD59: cluster of differentiation 59; IL1RA: interleukin 1 receptor antagonist; ApoE: apolipoprotein E; F9: coagulation factor IX; FGA: fibrinogen alpha chain; IGFBP5: insulin-like growth factor-binding protein 5; miR: micro RNA; cfDNA: cell free DNA; LC: liquid chromatography; MS: mass spectrometry; NMR: nuclear magnetic resonance; miRNA: micro RNA; RT: reverse transcription; PCR: polymerase chain reaction; Seq: sequencing; RFLP: restriction fragment length polymorphism; Sqn: sequenom; CE: capillary electrophoresis; PP: prepregnancy

[a] Prediction time is in gestational weeks unless “PP” is stated.

[b] Internal validation (IV) was considered as the assessment of the best model performance with the same samples that were used to train it and resampling techniques, within the same article that originally reported the model.

[c] External validation (EV) was considered as the assessment of the best model performance with samples that were not used to train it, within the same article that originally reported the model.

4.2.3.4. Mixed variables-based models

Models of this category combine medical and alternative data, which may derive from metabolomics [121–128], proteomics [129,130], or analysis of miRNA [131], cfDNA [132], genetic variants [133–135] or microbiota [136,137]. These approaches are summarized in **Table 4.2-3**. Notably, none of them have had independent validation.

4.2.3.4.1. Medical and metabolomics variables-based models

Most of these approaches have a moderate (AUC 0.7-0.8) or a high (AUC 0.8-0.9) predictive power. Only one article of this type performed external validation on its model, which showed a moderate performance: Lu et al. applied serum LC-MS lipidomics to predict GDM at 15 WP. They identified four lipid metabolites that were associated with GDM. Interestingly, when they were added to maternal age and BMI, external validation AUC increased from 0.689 to 0.741 [121].

Four articles of this kind reported high (AUC 0.8-0.9) or very high (AUC 0.9-1.0) performance GDM predictive models, but only one of them performed internal validation: Enquobahrie et al. used serum untargeted GC-MS to predict GDM at 16 WP. They identified 17 metabolites that differentiated GDM from non-GDM women. When they were added to a set of seven medical risk factors, internal

validation AUC increased from 0.712 to 0.871 [126]. The other studies with high or very high predictive power models did not perform any type of validation. For instance, Nevalainen et al. employed mass spectrometry to measure 10 amino acids and 31 acylcarnitines in first trimester of pregnancy serum samples, and found that arginine, glycine and 3-hydroxyisovalerylcarnitine were altered in GDM women. Their combination with four medical variables increased unvalidated AUC from 0.66 to 0.83 [122]. Another example is the article of Liu et al., who predicted GDM at 10 WP with serum targeted LC-MS metabolomics data. They identified 17 differential metabolites between GDM and control pregnancies. Their addition to a set of seven clinical risk factors for GDM improved unvalidated AUC from 0.69 to 0.92 [127]. Likewise, Wang et al. analyzed plasma samples with targeted LC-MS lipidomics. They identified 10 lipids that were significantly associated with GDM, and combined them with traditional clinical and biochemical risk factors to predict GDM at 6-15 WP. This combination resulted in an enhanced unvalidated AUC, which increased from 0.701 to 0.801 [128].

4.2.3.4.2. Medical and proteomics variables-based models

Two articles presented this type of strategy. Both models had external validation, one with a moderate (AUC 0.7-0.8) and the other one with a high (AUC 0.8-0.9) predictive power. In 2016, Ravnsborg et al. used serum targeted LC-MS proteomics to predict GDM at 8-13 WP. Of 47 candidate proteins, they selected

three, adiponectin, apolipoprotein M and apolipoprotein L1. Their combination with maternal age reached external validation AUCs of 0.776 for non-obese women, 0.707 for obese women and 0.725 for all [130]. In 2019, Ravnsborg et al. applied a LC-MS proteomics discovery strategy to predict GDM in obese women during their first trimester of pregnancy. They measured 548 serum proteins, of which three remained statistically significant after a multiple reaction monitoring-mass spectrometry verification experiment: afamin, serum amyloid P-component and vitronectin. Their best model combined vitronectin, maternal age and family history of DM, and yielded an external validation AUC of 0.806, markedly higher than the AUC of 0.625 for vitronectin solo, but just slightly higher than the one for clinical variables only [129].

4.2.3.4.3. Medical and miRNA variables-based models

Only one study presented this kind of model, with a high (AUC 0.8-0.9) predictive power but without validation: Sorensen et al. analyzed sera from obese pregnant women by RT-qPCR, to predict GDM before 20 WP. They found that miR-16-5p, miR-29a-3p and miR-134-5p levels were increased in GDM subjects. Together, those miRNAs reached an unvalidated AUC of 0.72. The authors also assessed the predictive performance of fasting glucose, and got an unvalidated AUC of 0.69. Interestingly, the combination of the three miRNAs with fasting glucose improved AUC to 0.81 [131].

4.2.3.4.4. Medical and cfDNA variables-based models

Only one article reported a model of this type. It had external validation and a moderate (AUC 0.7-0.8) performance: Guo et al. assessed plasma cfDNA to predict GDM before 18 WP. By cfDNA whole-genome sequencing, they found 800 genes with significantly different coverage between GDM and non-GDM women. 11 of those genes were combined to predict GDM, and reached AUCs of 0.732, 0.699 and 0.711 for different external validation cohorts. The inclusion of maternal BMI reduced AUC from 0.718 to 0.682, in a dataset that comprised the training and the three external validation cohorts [132].

4.2.3.4.5. Medical and genetic variants variables-based models

Three articles reported this kind of strategy, with a moderate (AUC 0.7-0.8) predictive power and no validation. For instance, Krishnan et al. evaluated the association between the *CREBRF* gene missense variant rs373863828 and GDM in Māori and Pacific women with obesity. They used blood samples and PCR-TaqMan for genotyping. The authors found that the *CREBRF* rs373863828 A allele was associated with a reduced likelihood of GDM in the study population. In addition, they combined rs373863828 with three clinical parameters, maternal age, BMI and family history of DM, to predict GDM at 12-17 WP, and got an unvalidated AUC of 0.76, higher than the obtained with clinical parameters only,

of 0.67 [133]. Similarly, Popova et al. tested the association between 11 common genetic variants and GDM in Russian women. They employed blood samples and PCR-TaqMan for genotyping. Their experiments showed that the *MTNR1B* gene SNP rs10830963 was significantly and independently associated with the risk of GDM. The authors assessed this genetic variant to predict GDM, and obtained an unvalidated AUC of 0.603. They also evaluated a set of eight clinical factors from prepregnancy, and got an unvalidated AUC of 0.712. Together, rs10830963 and clinical data reached an AUC of 0.729 [134]. Likewise, Dziedziejko et al. used different type 2 diabetes-related genetic variants to predict GDM in Polish women. They studied eight SNPs using blood samples and PCR-TaqMan genotyping: *PTGS2* (*COX2*) rs6681231, *FADS1* rs174550, *HNF1B* rs4430796, *ADIPOQ* rs266729, *IL18* rs187238, *CCL2* rs1024611, *HHEX* rs5015480 and *CDKN2A/2B* rs10811661. The combination of the SNPs with maternal age and prepregnancy BMI, yielded an unvalidated AUC of 0.773, greater than the obtained with maternal age and prepregnancy BMI only, of 0.683 [135].

4.2.3.4.6. Medical and microbiota variables-based models

Two studies presented this type of model, one with external validation and a low (AUC 0.6-0.7) predictive power, and the other one with internal validation and a moderate (AUC 0.7-0.8) performance. The first one was reported by Ma et al. They used 16S rRNA gene amplicon sequencing to analyze gut microbiota at 10-

15 WP, and found several differences between GDM and non-GDM women gut microbial composition. In order to predict GDM, they combined five microbiota genera, Parabacteroides, Ruminococcus 2, Ruminococcaceae UCG-014, Alloprevotella and uncultured-Ruminococcaceae, with two biochemical variables, fasting glucose and glutamine transpeptidase, and achieved an external validation AUC of 0.696, slightly higher than the one for medical variables only [136]. The second one was developed by Hu et al., who also analyzed gut microbiota by 16S rRNA gene amplicon sequencing. They used 96 microbial genera to predict GDM at 6-15 WP, and got an internal validation AUC of 0.77. Moreover, they identified six taxa that were significantly associated with a lower risk of GDM: phylum Actinobacteria, family Coriobacteriaceae, family Gemellaceae, genus 26 of Coriobacteriaceae, genus 74 of Gemellaceae, and an unknown species of Coprococcus. Together, they achieved an AUC of 0.66, slightly lower than the obtained by conventional clinical and biochemical risk factors for GDM, of 0.69. The combination of the six microbial taxa with medical risk factors reached an internal validation AUC of 0.75 [137].

Table 4.2-3. Original articles that report ML-based models for GDM early prediction based on mixed variables.

Type of variables	Article	Characteristics of the most accurate ML-based model reported for GDM early prediction										
		Nº of variables	Variables	Technique	Matrix	Prediction time [a]	Predictive power				Validation	
							AUC	Ac	Sp	Se	IV [b]	EV [c]
Medical and metabolomics	[121]	6	Age, BMI, TG(51.1), TG(48:1), PC(32:1), PCae(40:3)	LC-MS	Serum	15	0.741	-	-	-	X	X
	[122]	7	Age, BMI, smoking, PAPP-A, Arg, Gly, 3-hydroxy-isovalerylcarnitine	MS	Serum	< 13	0.83	-	80	72	-	-
	[123]	10	Age, BMI, family history of DM, parity, fasting glucose, insulin, TG, GGT, Ile, Tyr	LC-MS	Serum	12-16	0.737	-	-	-	-	-
	[124]	14	BMI, family history of DM, parity, ethnicity, gestational age at registration, weight gain, smoking, drinking, education, SBP, DBP, ALT, DCA, GUDCA	LC-MS	Serum	8-12	0.76	-	-	-	-	-
	[125]	16	Age, BMI, family history of DM, sum of skinfold thicknesses, waist/height, neck/thigh, waist/thigh, SBP, DBP, random glucose, HbA1C, TG, adiponectin, SHBG, ferritin, fructosamine	NMR	Serum	15-18	0.77	-	-	-	X	-
	[126]	24	Age, BMI, family history of DM, CRP, ferritin, hepcidin, vitamin D, linoleic acid, oleic acid, myristic acid, d-galactose, d-sorbitol, o-phosphocolamine, l-Ala, l-	GC-MS	Serum	16	0.871	-	-	-	X	-

			Val, 5-hydroxy-l-Trp, l-Ser, sarcosine, l-pyroglutamic acid, l-mimosine, l-lactic acid, glycolic acid, fumaric acid, urea									
	[127]	24	BMI, family history of DM, parity, weight gain per week, smoking, drinking, SBP, PC(40:6), LPC(16:0), LPC(17:0), LPC(18:0), LPC(18:1), LPE(16:0), LPE(18:0), Cer(36:2), FA(20:0), GCA, GDCA, GCDCA, GUDCA, TDCA, TCDCA, TMA, L-carnitine	LC-MS	Serum	10	0.92	-	-	-	-	-
	[128]	24	Age, BMI, family history of DM, prior GDM, parity, smoking, drinking, physical activity, SBP, HDL-C, TG, fasting glucose, insulin, HbA1c, PC(40:7), PC-O(36:1), PE-O(40:5), PE-P(38:6), PI(40:6), SM(34:1), DG(18:0/18:1), MHC(18:0), DHC(24:0), DHC(24:1)	LC-MS	Plasma	6-15	0.801	-	-	-	-	-
Medical and proteomics	[129]	3	Age, family history of DM, vitronectin	LC-MS	Serum	8-13	0.806	-	-	-	-	X
	[130]	4	Age, adiponectin, Apo M, Apo L1	LC-MS	Serum	8-13	0.776	-	-	-	-	X
Medical and micro RNA	[131]	4	Fasting glucose, miR-29a-3p, miR-134-5p, miR-16-5p	RT-PCR	Serum	< 20	0.81	-	82.1	68.4	-	-
Medical and cell free DNA	[132]	11	CC2D2B, NAT10, SIPA1, ZNF565, ZNF552, WDR35, MICALL1, CTNNB1, CLOCK, BCKDHB, TGIF2LY	PCR-Seq	Plasma	< 18	0.732, 0.699, 0.711	72.7, 71.0, 71.0	72.2, 72.0, 70.8	74.1, 67.7, 71.4	X	X

Medical and genetic variants	[133]	4	Age, BMI, family history of DM, rs373863828	PCR-TaqMan	Blood	12-17	0.76	-	-	-	-	-
	[134]	9	Age, BMI, family history of DM, prior GDM, prior hypertension, parity, prior IGT, PCOS, rs10830963	PCR-TaqMan	Blood	PP	0.729	-	-	-	-	-
	[135]	10	Age, BMI, rs6681231, rs174550, rs4430796, rs266729, rs187238, rs1024611, rs5015480, rs10811661	PCR-TaqMan	Blood	PP	0.773	-	70	70	-	-
Medical and microbiota	[136]	7	Fasting glucose, GGT, Parabacteroides, Ruminococcus 2, Ruminococcaceae UCG-014, Alloprevotella, uncultured-Ruminococcaceae	PCR-Seq	Stool	10-15	0.696	-	-	-	X	X
	[137]	96	Among them: genus 227 of Enterobacteriaceae, Actinomyces, genus 26 of Coriobacteriaceae, genus 27 of Coriobacteriaceae, Gemellaceae, Ruminococcus, Clostridium, genus 103 of Lachnospiraceae	PCR-Seq	Stool	6-15	0.77	-	-	-	X	-

Abbreviations. AUC: area under the ROC curve; Ac: accuracy; Sp: specificity; Se: sensitivity; IV: internal validation; EV: external validation. Apo M: apolipoprotein M; ApoL1: apolipoprotein L1; BMI: body mass index; TG: triglyceride; PC: phosphatidylcholine; PCae: choline ether phospholipid; PAPP-A: pregnancy associated plasma protein A; Arg: arginine; Gly: glycine; DM: diabetes mellitus; GGT: gamma glutamyl transferase; Ile: isoleucine; Tyr: tyrosine; SBP: systolic blood pressure; DBP: diastolic blood pressure; ALT: alanine aminotransferase; DCA: deoxycholic acid; GUDCA: glycoursodeoxycholic acid; HbA1c: glycated hemoglobin A1c; SHBG: sex hormone binding globulin; CRP: C reactive protein; Ala: alanine; Val: valine; Trp: tryptophan; Ser: serine; LPC: lysophosphatidylcholine; LPE:

lysophosphatidylethanolamine; Cer: ceramide; FA: fatty acid; GCA: glycocholic acid; GDCA: glycodeoxycholic acid; GCDCA: glycochenodeoxycholic acid; TDCA: taurodeoxycholic acid; TCDCA: taurochenodeoxycholic acid; TMA: trimethylamine; GDM: gestational diabetes mellitus; HDL-C: high density lipoprotein cholesterol; PC-O: alkylphosphatidylcholine; PE-O: alkylphosphatidylethanolamine; PE-P: phosphatidylethanolamine plasmalogen; PI: phosphatidylinositol; SM: sphingomyelin; DG: diacylglyceride; MHC: mono hexosyl ceramide; DHC: dihexosyl ceramide; Apo: apolipoprotein; miR: micro RNA; CC2D2B: coiled-coil and C2 domain containing 2B gene; NAT10: N-acetyltransferase 10 gene; SIPA1: signal induced proliferation-associated 1 gene; ZNF565: zinc finger protein 565 gene; ZNF552: zinc finger protein 552 gene; WDR35: WD repeat domain 35 gene; MICALL1: MICAL-like protein 1 gene; CTNNB1: catenin beta 1 gene; CLOCK: circadian locomotor output cycles kaput gene; BCKDHB: branched chain keto acid dehydrogenase E1 subunit beta gene; TGIF2LY: TGFB induced factor homeobox 2 like Y-linked gene; IGT: impaired glucose tolerance; PCOS: polycystic ovary syndrome; UCG: uncultured genus-level group; LC: liquid chromatography; MS: mass spectrometry; NMR: nuclear magnetic resonance; GC: gas chromatography; RT: reverse transcription; PCR: polymerase chain reaction; Seq: sequencing; PP: prepregnancy.

[a] Prediction time is in gestational weeks unless “PP” is stated.

[b] Internal validation (IV) was considered as the assessment of the best model performance with the same samples that were used to train it and resampling techniques, within the same article that originally reported the model.

[c] External validation (EV) was considered as the assessment of the best model performance with samples that were not used to train it, within the same article that originally reported the model.

4.2.4. Discussion

4.2.4.1. The importance of independent validation

Only 8.3% of the GDM predictive models presented in this review have had independent validation. This is consistent with what was reported by Lamain de Rutier et al. in 2017. Their systematic review on clinical data-based multivariate models for GDM prediction and risk estimation, found out that only four out of 14 strategies were validated in independent studies [26]. Our review demonstrates that this issue have persisted in time and is not limited to clinical variables-based models.

Independent validation is an essential step before the clinical implementation of a predictive model, since it ensures the assessment of its performance in populations that are different from the one that was involved in its calibration. Consequently, it confirms or refutes if the model will predict the pathology of interest accurately in patients from different ages, ethnicity, medical background, health centers, geographical areas and time periods.

Most of the reviewed models that have had independent validation are only based on clinical data, which can be explained by the fact that these strategies were the first to be developed. Moreover, clinical variables are the easiest to obtain, since

they do not need to be acquired invasively. Thus, this kind of approach is the easiest to implement in medical practice. However, independent validation of these models have demonstrated that they have low (AUC 0.6-0.7) or moderate (AUC 0.7-0.8) predictive power [21–25]. Therefore, other approaches must be considered and assessed.

4.2.4.2. The most promising models for GDM early prediction

The three models with higher predictive power presented in this review are not based on clinical variables. Xiong et al. strategy employs biochemical parameters and achieved an AUC of 0.998 [46], Koos et al. model uses metabolomics-derived data and yielded an AUC of 0.993 [104], and Zhao et al. approach combines proteomics-derived variables and reached an AUC of 0.985 [110]. All these strategies are based on parameters that are measured in biological fluids and, consequently, are more invasive than the ones that employ clinical variables. However, none of them are based on classic GDM risk factors, but explore the performance of new biomarkers for its prediction, which is relevant considering that 17-48% of women that develop GDM do not have evident risk factors [7].

Due to the complexity of the procedures needed to obtain the variables involved, some of these models are more likely to be implemented in routine clinical practice in the short term. For example, Xiong et al. strategy is based on two coagulation

variables. Since coagulation parameters are measured in blood, a sample that is commonly required at regular pregnancy controls, and clinical laboratories are usually equipped to measure them, the implementation of this model in medical settings is plausible. In contrast, Koos et al. approach is based on seven metabolites, which are measured in urine. The obtainment of this biological sample is less invasive than the obtainment of blood. However, metabolomics analysis implies the use of LC-MS and/or GC-MS equipment that is not usually available in clinical laboratories. Thus, the immediate and wide clinical implementation of this model is unlikely. Likewise, Zhao et al. strategy is based on four serum proteins, which were originally measured by LC-MS. Nevertheless, these proteins may be measured by antibodies-based assays, which are often available in clinical laboratories. Therefore, the implementation of this model in medical settings is feasible. In consequence, both Xiong et al. and Zhao et al. strategies have a great potential to be implemented in clinical practice. However, none of them have had independent validation. In fact, while the first model was externally validated within its original article, the second one has never been validated.

4.2.4.3. Lack of validation is an opportunity

56.0% of the GDM predictive models presented in this review were not validated in their original articles, nor have been validated in independent studies.

In this context, we encourage researchers to validate their predictive models when developing a new one. Internal and external validation can be performed easily, since they do not necessarily require more than the already available subjects. Indeed, the working dataset can be split in two, i.e. a training one for calibration and internal validation, and a testing one for external validation. Although splitting-sampling is not the best form of external validation, as the testing population is quite similar to the training one, it is better than not carrying out any. Of course, a completely different cohort should be used for external validation, if available. In addition, we motivate researchers to report model validation as recommended in the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement, a 22-item checklist that aims to improve the reporting quality of studies developing, validating or updating a prediction model [138].

Furthermore, we encourage the scientific community to share their predictive models, so they can be available for independent validation, and to carry out studies to validate the models that are already published. Independent validation demands an important effort from researchers, since it requires to recruit and follow-up a large cohort of subjects, and both time and resources are spent; however, it is essential to achieve the goal of implementing any type of predictive model in medical practice. Importantly, validation is not the only requirement that predictive models must fulfill before their clinical implementation. Once validated,

a model must go through a process called “impact analysis”, in which its effect on clinical behavior is assessed. Only the models that change physician behavior with beneficial consequences, i.e. improve patients outcomes and/or reduce costs without compromising the quality of care and patient satisfaction, are considered appropriate for clinical implementation [20]. Since only after independent validation predictive models are suitable for impact analysis and subsequent application in medical settings, the fact that most of the GDM predictive models reported in literature have never been validated is a great opportunity to continue their way towards clinical implementation.

4.2.4.4. Strengths of this review

Our search strategy was not restricted to a specific type of variable. Moreover, we considered several synonyms of the terms of interest to comprise as many models as possible. This allowed us to include 109 original research articles and their best predictive models in this review, models that were based on a wide variety of variables, i.e. medical, alternative or mixed. This is the first review to assess this amount and diversity of ML-based models for GDM prediction before 24-28 WP.

4.2.4.5. Limitations of this review

Our review was limited to PubMed English-written articles. In addition, we only included studies that reported the predictive power of their models by AUC, accuracy, specificity, sensitivity or Q^2. In other words, we excluded manuscripts that stated their models performance by other parameters, i.e. likelihood ratios, positive and negative predictive values or odds ratio. Hence, we may have missed some highly accurate and potentially promising models for GDM early prediction.

4.2.5. Conclusion

In literature, there is a wide variety of ML-based models for GDM prediction before 24-28 WP. Most of them are based only on clinical and/or biochemical variables; some of them are built exclusively on variables that derive from metabolomics, proteomics, micro-RNA or microbiota analysis, or other methods; while others combine both. This is the first review to assess this diversity of multivariate models for GDM prediction at early pregnancy stages.

Some of the reviewed models have a very high predictive power, but have never been validated. This is a great opportunity to continue their path towards clinical implementation, a process that is illustrated in **Figure 4.2-2**. The independent validation of these models will pave the way for their impact analysis and

subsequent implementation in medical settings. Once settled in routine clinical practice, they would allow to predict GDM before 24-28 WP accurately, to initiate its treatment timely and to prevent its short and long-term negative health consequences for mother and offspring.

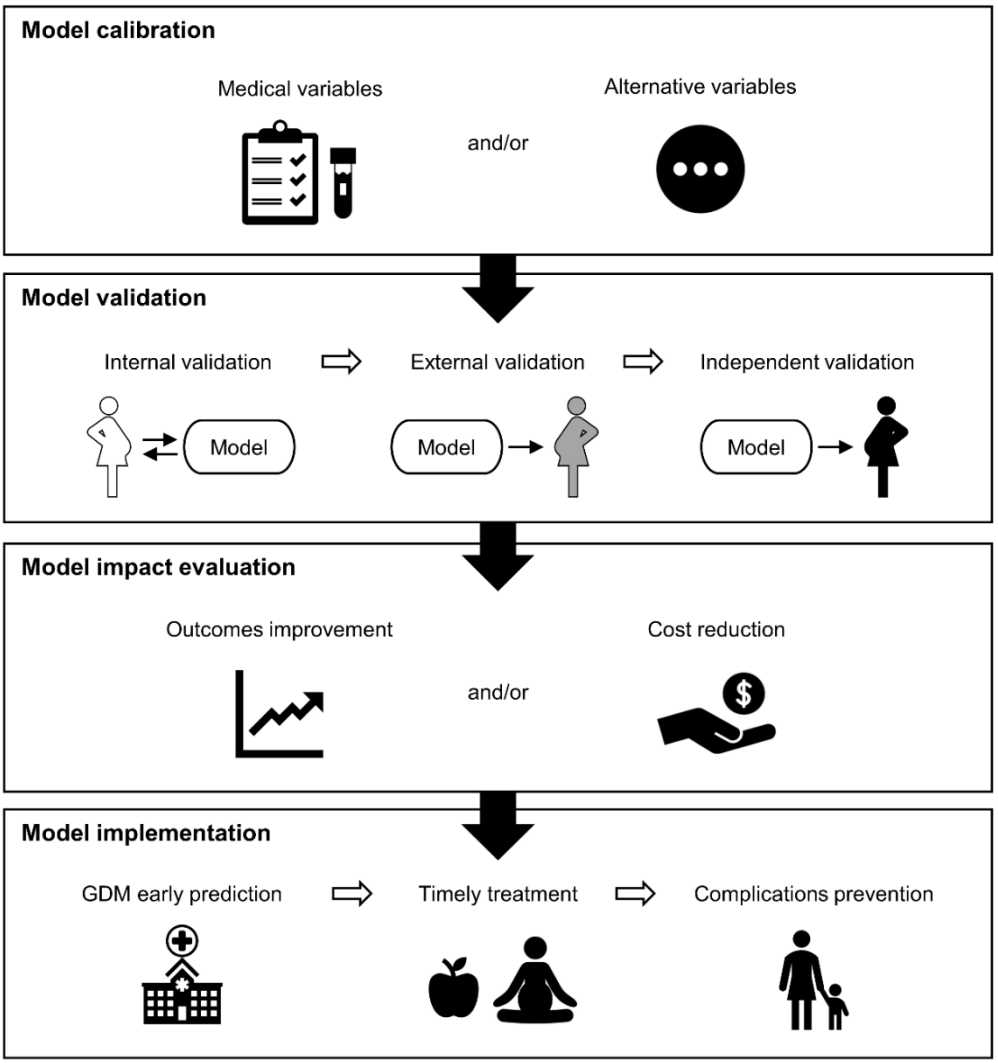


Figure 4.2-2. ML-based models for GDM prediction before 24-28 weeks of pregnancy: from calibration to implementation. The way of GDM predictive models towards clinical implementation is illustrated as a four-step flow diagram. Each step is represented as a rectangular box. The first step is model calibration, i.e. its development by combining medical and/or alternative variables. Only calibrated models can move towards (thick black arrow) the next step. The second step is model validation, i.e. the assessment of its performance by predicting different datasets. The ideal situation is that models can be internally validated, by predicting the calibration dataset (white pregnant woman icon) by resampling techniques (pair of thin black arrows); then (white arrow) externally validated, by predicting (thin black arrow) a separate dataset (gray pregnant woman icon) within the same manuscript; and finally (white arrow) independently validated, by predicting (thin black arrow) a separate dataset (black pregnant woman icon) in an independent study. Only independently validated models can move towards (thick black arrow) the next step. The third step is model impact evaluation, i.e. the assessment of its effect on clinical behavior. Only the models that improve patients outcomes and/or reduce costs can move towards (thick black arrow) the next step. The fourth step is model implementation, i.e. its application on routine clinical practice. Once implemented in medical settings, ML-based models will lead to the accurate prediction of GDM before 24-28 weeks of pregnancy, to its subsequent (white arrow) and timely treatment with lifestyle intervention including healthy diet and physical activity, and ultimately (white arrow) to the prevention of its short and long-term health complications for both mother and offspring.

Acknowledgments: This work was finantially supported by Agencia Nacional de Investigación y Desarrollo (ANID) through 21190736 PhD scholarship (DM), as well as FONDECYT 11181153 (JA) and Fomento Vinculación Internacional para Instituciones de Investigación Regionales FOVI210057 (EG-G) grants; Universidad del Bio-Bío by means of DIUBB 180709 3/R (AR) grant; and Universidad de Concepción/Ministerio de Educación del Gobierno de Chile via Proyecto UCO 1866 fund (DM). These funding sources were not involved in study design; data collection, analysis nor interpretation; report writing; nor in the decision to submit this manuscript for publication.

Conflict of interest statement: The authors have nothing to disclose.

4.2.6. References

1. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. Nat Rev Dis Primers. 2019;5: 47. doi:10.1038/s41572-019-0098-8

2. Metzger B, Lowe L, Dyer A, Trimble E, Chaovarindr U, Coustan D, et al. Hyperglycemia and Adverse Pregnancy Outcomes. New England Journal of Medicine. 2008;358: 1991–2002. doi:10.1056/NEJMoa0707943

3. Lowe WL, Scholtens DM, Lowe LP, Kuang A, Nodzenski M, Talbot O, et al. Association of gestational diabetes with maternal disorders of glucose metabolism and childhood adiposity. JAMA - Journal of the American Medical Association. 2018;320: 1005–1016. doi:10.1001/jama.2018.11628

4. Lowe WL, Scholtens DM, Kuang A, Linder B, Lawrence JM, Lebenthal Y, et al. Hyperglycemia and adverse Pregnancy Outcome follow-up study (HAPO FUS): Maternal gestational diabetes mellitus and childhood glucose metabolism. Diabetes Care. 2019;42: 372–380. doi:10.2337/dc18-1646

5. Farahvar S, Walfisch A, Sheiner E. Gestational diabetes risk factors and long-term consequences for both mother and offspring: a literature review. Expert Rev Endocrinol Metab. 2019;14: 63–74. doi:10.1080/17446651.2018.1476135

6. Popova P, Castorino K, Grineva EN, Kerr D. Gestational diabetes mellitus diagnosis and treatment goals: measurement and measures. Minerva Endocrinol. 2016;41: 421–432. Available: http://www.ncbi.nlm.nih.gov/pubmed/26824326

7. Kuo CH, Li HY. Diagnostic Strategies for Gestational Diabetes Mellitus: Review of Current Evidence. Curr Diab Rep. 2019;19: 11–16. doi:10.1007/s11892-019-1271-x

8. Tsakiridis I, Giouleka S, Mamopoulos A, Kourtis A, Athanasiadis A, Filopoulou D, et al. Diagnosis and Management of Gestational Diabetes Mellitus: An Overview of National and International Guidelines. Obstet Gynecol Surv. 2021;76: 367–381. doi:10.1097/OGX.0000000000000899

9. Sovio U, Murphy HR, Smith GCS. Accelerated fetal growth prior to diagnosis of gestational diabetes mellitus: A prospective cohort study of

nulliparous women. Diabetes Care. 2016;39: 982–987. doi:10.2337/dc16-0160

10. Venkataraman H, Ram U, Craik S, Arungunasekaran A, Seshadri S, Saravanan P. Increased fetal adiposity prior to diagnosis of gestational diabetes in South Asians: more evidence for the 'thin–fat' baby. Diabetologia. 2017;60: 399–405. doi:10.1007/s00125-016-4166-2

11. Yovera L, Zaharia M, Jachymski T, Velicu-Scraba O, Coronel C, de Paco Matallana C, et al. Impact of gestational diabetes mellitus on fetal cardiac morphology and function: cohort comparison of second- and third-trimester fetuses. Ultrasound in Obstetrics and Gynecology. 2021;57: 607–613. doi:10.1002/uog.22148

12. Moran PS, Wuytack F, Turner M, Normand C, Brown S, Begley C, et al. Economic burden of maternal morbidity – A systematic review of cost-of-illness studies. PLoS One. 2020;15: 1–18. doi:10.1371/JOURNAL.PONE.0227377

13. Immanuel J, Simmons D. Screening and Treatment for Early-Onset Gestational Diabetes Mellitus: a Systematic Review and Meta-analysis. Curr Diab Rep. 2017;17: 1–11. doi:10.1007/s11892-017-0943-7

14. Powe CE. Early Pregnancy Biochemical Predictors of Gestational Diabetes Mellitus. Curr Diab Rep. 2017;17: 12. doi:10.1007/s11892-017-0834-y

15. Lorenzo-Almorós A, Hang T, Peiró C, Soriano-Guillén L, Egido J, Tuñón J, et al. Predictive and diagnostic biomarkers for gestational diabetes and its associated metabolic and cardiovascular diseases. Cardiovasc Diabetol. 2019;18: 140. doi:10.1186/s12933-019-0935-9

16. Gan WZ, Ramachandran V, Lim CSY, Koh RY. Omics-based biomarkers in the diagnosis of diabetes. J Basic Clin Physiol Pharmacol. 2020;31: 1–21. doi:10.1515/jbcpp-2019-0120

17. van Hoorn F, Koster MPH, Naaktgeboren CA, Groenendaal F, Kwee A, Lamain-de Ruiter M, et al. Prognostic models versus single risk factor approach in first-trimester selective screening for gestational diabetes mellitus: a prospective population-based multicentre cohort study. BJOG. 2020; 1–10. doi:10.1111/1471-0528.16446

18. Goecks J, Jalili V, Heiser LM, Gray JW. How Machine Learning Will Transform Biomedicine. Cell. 2020;181: 92–101. doi:10.1016/j.cell.2020.03.022

19. Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagn Progn Res. 2019;3: 16. doi:10.1186/s41512-019-0060-y

20. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature XXII: How to use articles about clinical decision rules. J Am Med Assoc. 2000;284: 79–84. doi:10.1001/jama.284.1.79

21. van Leeuwen M, Opmeer BC, Zweers EJK, van Ballegooie E, ter Brugge HG, de Valk HW, et al. External validation of a clinical scoring system for the risk of gestational diabetes mellitus. Diabetes Res Clin Pract. 2009;85: 96–101. doi:10.1016/j.diabres.2009.04.025

22. Thériault S, Forest JC, Massé J, Giguère Y. Validation of early risk-prediction models for gestational diabetes based on clinical characteristics. Diabetes Res Clin Pract. 2014;103: 419–425. doi:10.1016/j.diabres.2013.12.009

23. Lamain - de Ruiter M, Kwee A, Naaktgeboren CA, de Groot I, Evers IM, Groenendaal F, et al. External validation of prognostic models to predict risk of gestational diabetes mellitus in one Dutch cohort: prospective multicentre cohort study. BMJ. 2016;354: i4338. doi:10.1136/bmj.i4338

24. Meertens LJE, Scheepers HCJ, van Kuijk SMJ, Roeleveld N, Aardenburg R, van Dooren IMA, et al. External validation and clinical utility of prognostic prediction models for gestational diabetes mellitus: A prospective cohort study. Acta Obstet Gynecol Scand. 2020;99: 891–900. doi:10.1111/aogs.13811

25. Kotzaeridi G, Blätter J, Eppel D, Rosicky I, Mittlböck M, Yerlikaya-Schatten G, et al. Performance of early risk assessment tools to predict the later development of gestational diabetes. Eur J Clin Invest. 2021;51: 1–11. doi:10.1111/eci.13630

26. Lamain - de Ruiter M, Kwee A, Naaktgeboren CA, Franx A, Moons KGM, Koster MPH. Prediction models for the risk of gestational diabetes: a systematic review. Diagn Progn Res. 2017;1: 1–9. doi:10.1186/s41512-016-0005-7

27. Huhn EA, Rossi SW, Hoesli I, Göbl CS. Controversies in Screening and Diagnostic Criteria for Gestational Diabetes in Early and Late Pregnancy.

Front Endocrinol (Lausanne). 2018;9: 5–12. doi:10.3389/fendo.2018.00696

28. Tran TS, Hirst JE, Do MAT, Morris JM, Jeffery HE. Early Prediction of Gestational Diabetes Mellitus in Vietnam. Diabetes Care. 2013;36: 618–624. doi:10.2337/dc12-1418

29. Naylor CD, Sermer M, Chen E, Farine D. Selective Screening for Gestational Diabetes Mellitus. New England Journal of Medicine. 1997;337: 1591–1596. doi:10.1056/NEJM199711273372204

30. Bourdages M, Demers MÉ, Dubé S, Gasse C, Girard M, Boutin A, et al. First-Trimester Abdominal Adipose Tissue Thickness to Predict Gestational Diabetes. Journal of Obstetrics and Gynaecology Canada. 2018;40: 883–887. doi:10.1016/j.jogc.2017.09.026

31. van Leeuwen M, Opmeer BC, Zweers EJK, van Ballegooie E, ter Brugge HG, de Valk HW, et al. Estimating the risk of gestational diabetes mellitus: A clinical prediction model based on patient characteristics and medical history. BJOG. 2010;117: 69–75. doi:10.1111/j.1471-0528.2009.02425.x

32. Han Z, Zhang Y, Li X, Chiu WH, Yin Y, Hou H. Investigation Into the Predictive Potential of Three-Dimensional Ultrasonographic Placental Volume and Vascular Indices in Gestational Diabetes Mellitus. Front Endocrinol (Lausanne). 2021;12: 1–8. doi:10.3389/fendo.2021.689888

33. Phaloprakarn C, Tangjitgamol S, Manusirivithaya S. A risk score for selective screening for gestational diabetes mellitus. European Journal of Obstetrics and Gynecology and Reproductive Biology. 2009;145: 71–75. doi:10.1016/j.ejogrb.2009.04.016

34. Teede HJ, Harrison CL, Teh WT, Paul E, Allan CA. Gestational diabetes: Development of an early risk prediction tool to facilitate opportunities for prevention. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2011;51: 499–504. doi:10.1111/j.1479-828X.2011.01356.x

35. Gabbay-Benziv R, Doyle LE, Blitzer M, Baschat AA. First trimester prediction of maternal glycemic status. J Perinat Med. 2015;43: 283–289. doi:10.1515/jpm-2014-0149

36. Donovan BM, Breheny PJ, Robinson JG, Baer RJ, Saftlas AF, Bao W, et al. Development and validation of a clinical model for preconception and early pregnancy risk prediction of gestational diabetes mellitus in

nulliparous women. PLoS One. 2019;14: 1–21. doi:10.1371/journal.pone.0215173

37. Sweeting AN, Appelblom H, Ross GP, Wong J, Kouru H, Williams PF, et al. First trimester prediction of gestational diabetes mellitus: A clinical model based on maternal demographic parameters. Diabetes Res Clin Pract. 2017;127: 44–50. doi:10.1016/j.diabres.2017.02.036

38. Sirico A, Lanzone A, Mappa I, Sarno L, Słodki M, Pitocco D, et al. The role of first trimester fetal heart rate in the prediction of gestational diabetes: A multicenter study. European Journal of Obstetrics and Gynecology and Reproductive Biology. 2019;243: 158–161. doi:10.1016/j.ejogrb.2019.10.019

39. Wang J, Lv B, Chen X, Pan Y, Chen K, Zhang Y, et al. An early model to predict the risk of gestational diabetes mellitus in the absence of blood examination indexes: application in primary health care centres. BMC Pregnancy Childbirth. 2021;21: 1–8. doi:10.1186/s12884-021-04295-2

40. Syngelaki A, Pastides A, Kotecha R, Wright A, Akolekar R, Nicolaides KH. First-trimester screening for gestational diabetes mellitus based on maternal characteristics and history. Fetal Diagn Ther. 2015;38: 14–21. doi:10.1159/000369970

41. Cremona A, O'Gorman CS, Ismail KI, Hayes K, Donnelly AE, Hamilton J, et al. A risk-prediction model using parameters of maternal body composition to identify gestational diabetes mellitus in early pregnancy. Clin Nutr ESPEN. 2021;45: 312–321. doi:10.1016/j.clnesp.2021.08.002

42. Garmendia ML, Mondschein S, Montiel B, Kusanovic JP. Trends and predictors of gestational diabetes mellitus in Chile. International Journal of Gynecology and Obstetrics. 2020;148: 210–218. doi:10.1002/ijgo.13023

43. Buerger O, Elger T, Varthaliti A, Syngelaki A, Wright A, Nicolaides KH. First-trimester screening for gestational diabetes mellitus in twin pregnancies. J Clin Med. 2021;10: 4–11. doi:10.3390/jcm10173814

44. Maged AM, Moety GAF, Mostafa WA, Hamed DA. Comparative study between different biomarkers for early prediction of gestational diabetes mellitus. Journal of Maternal-Fetal and Neonatal Medicine. 2014;27: 1108–1112. doi:10.3109/14767058.2013.850489

45. Chaparro A, Zúñiga E, Varas-Godoy M, Albers D, Ramírez V, Hernández M, et al. Periodontitis and placental growth factor in oral fluids are early

pregnancy predictors of gestational diabetes mellitus. J Periodontol. 2018;89: 1052–1060. doi:10.1002/JPER.17-0497

46. Xiong Y, Lin L, Chen Y, Salerno S, Li Y, Zeng X, et al. Prediction of gestational diabetes mellitus in the first 19 weeks of pregnancy using machine learning techniques. The Journal of Maternal-Fetal & Neonatal Medicine. 2022;35: 2457–2463. doi:10.1080/14767058.2020.1786517

47. Li G, Huang W, Zhang L, Tian Z, Zheng W, Wang T, et al. A prospective cohort study of early-pregnancy risk factors for gestational diabetes in polycystic ovarian syndrome. Diabetes Metab Res Rev. 2018;34: 3–9. doi:10.1002/dmrr.3003

48. Amini M, Kazemnejad A, Zayeri F, Montazeri A, Rasekhi A, Amirian A, et al. Diagnostic accuracy of maternal serum multiple marker screening for early detection of gestational diabetes mellitus in the absence of a gold standard test. BMC Pregnancy Childbirth. 2020;20: 375. doi:10.1186/s12884-020-03068-7

49. Rasanen JP, Snyder CK, Rao P v., Mihalache R, Heinonen S, Gravett MG, et al. Glycosylated fibronectin as a first-trimester biomarker for prediction of gestational diabetes. Obstetrics and Gynecology. 2013;122: 586–594. doi:10.1097/AOG.0b013e3182a0c88b

50. Zhang Y, Lan X, Cai C, Li R, Gao Y, Yang L, et al. Associations between Maternal Lipid Profiles and Pregnancy Complications: A Prospective Population-Based Study. Am J Perinatol. 2021;38: 834–840. doi:10.1055/s-0039-3402724

51. Schuitemaker JHN, Beernink RHJ, Franx A, Cremers TIFH, Koster MPH. First trimester secreted Frizzled-Related Protein 4 and other adipokine serum concentrations in women developing gestational diabetes mellitus. PLoS One. 2020;15: 1–13. doi:10.1371/journal.pone.0242423

52. Wang C, Zhu W, Wei Y, Su R, Feng H, Lin L, et al. The Predictive Effects of Early Pregnancy Lipid Profiles and Fasting Glucose on the Risk of Gestational Diabetes Mellitus Stratified by Body Mass Index. J Diabetes Res. 2016;2016: 1–8. doi:10.1155/2016/3013567

53. Correa PJ, Venegas P, Palmeiro Y, Albers D, Rice G, Roa J, et al. First trimester prediction of gestational diabetes mellitus using plasma biomarkers: a case-control study. J Perinat Med. 2019;47: 161–168. doi:10.1515/jpm-2018-0120

54. Hao M, Lin L. Fasting plasma glucose and body mass index during the first trimester of pregnancy as predictors of gestational diabetes mellitus in a Chinese population. Endocr J. 2017;64: 561–569. doi:10.1507/endocrj.EJ16-0359

55. Tramontana A, Pablik E, Stangl G, Hartmann B, Dieplinger H, Hafner E. Combination of first trimester serum afamin levels and three-dimensional placental bed vascularization as a possible screening method to detect women at-risk for adverse pregnancy complications like pre-eclampsia and gestational diabetes mellitus in low-r. Placenta. 2018;62: 9–15. doi:10.1016/j.placenta.2017.12.014

56. Kouhkan A, Khamseh ME, Moini A, Pirjani R, Arabipoor A, Zolfaghari Z, et al. Diagnostic accuracy of body mass index and fasting glucose for the prediction of gestational diabetes mellitus after assisted reproductive technology. Int J Fertil Steril. 2019;13: 32–37. doi:10.22074/ijfs.2019.5505

57. Maymon R, Meiri H, Svirski R, Weiner E, Cuckle H. Maternal serum screening marker levels in twin pregnancies affected by gestational diabetes. Arch Gynecol Obstet. 2019;299: 655–663. doi:10.1007/s00404-018-5010-3

58. Pan Y, Hu J, Zhong S. The joint prediction model of pBMI and eFBG in predicting gestational diabetes mellitus. Journal of International Medical Research. 2020;48: 030006051988919. doi:10.1177/0300060519889199

59. Rashidi H, Kalantari K, Shahbazian H, Nouhjah S. The relationship between fasting plasma glucose in the first trimester of pregnancy and the incidence of gestational diabetes in Iran. Diabetes and Metabolic Syndrome: Clinical Research and Reviews. 2021;15: 102193. doi:10.1016/j.dsx.2021.102193

60. Eleftheriades M, Papastefanou I, Lambrinoudaki I, Kappou D, Lavranos D, Akalestos A, et al. Elevated placental growth factor concentrations at 11-14 weeks of gestation to predict gestational diabetes mellitus. Metabolism. 2014;63: 1419–1425. doi:10.1016/j.metabol.2014.07.016

61. Pintaudi B, di Vieste G, Corrado F, Lucisano G, Pellegrini F, Giunta L, et al. Improvement of selective screening strategy for gestational diabetes through a more accurate definition of high-risk groups. Eur J Endocrinol. 2014;170: 87–93. doi:10.1530/EJE-13-0759

62. Hassiakos D, Eleftheriades M, Papastefanou I, Lambrinoudaki I, Kappou D, Lavranos D, et al. Increased Maternal Serum Interleukin-6

Concentrations at 11 to 14 Weeks of Gestation in Low Risk Pregnancies Complicated with Gestational Diabetes Mellitus: Development of a Prediction Model. Hormone and Metabolic Research. 2015;48: 35–41. doi:10.1055/s-0034-1395659

63. Farina A, Eklund E, Bernabini D, Paladino M, Righetti F, Monti G, et al. A First-Trimester Biomarker Panel for Predicting the Development of Gestational Diabetes. Reproductive Sciences. 2017;24: 954–959. doi:10.1177/1933719116675057

64. Liu C, Wang Y, Zheng W, Wang J, Zhang Y, Song W, et al. Putrescine as a Novel Biomarker of Maternal Serum in First Trimester for the Prediction of Gestational Diabetes Mellitus: A Nested Case-Control Study. Front Endocrinol (Lausanne). 2021;12: 1–8. doi:10.3389/fendo.2021.759893

65. Lovati E, Beneventi F, Simonetta M, Laneri M, Quarleri L, Scudeller L, et al. Gestational diabetes mellitus: Including serum pregnancy-associated plasma protein-A testing in the clinical management of primiparous women? A case-control study. Diabetes Res Clin Pract. 2013;100: 340–347. doi:10.1016/j.diabres.2013.04.002

66. Visconti F, Quaresima P, Chiefari E, Caroleo P, Arcidiacono B, Puccio L, et al. First trimester combined test (FTCT) as a predictor of gestational diabetes mellitus. Int J Environ Res Public Health. 2019;16: 1–10. doi:10.3390/ijerph16193654

67. Zheng T, Ye W, Wang X, Li X, Zhang J, Little J, et al. A simple model to predict risk of gestational diabetes mellitus from 8 to 20 weeks of gestation in Chinese women. BMC Pregnancy Childbirth. 2019;19: 252. doi:10.1186/s12884-019-2374-8

68. Guo F, Yang S, Zhang Y, Yang X, Zhang C, Fan J. Nomogram for prediction of gestational diabetes mellitus in urban, Chinese, pregnant women. BMC Pregnancy Childbirth. 2020;20: 1–8. doi:10.1186/s12884-019-2703-y

69. Tenenbaum-Gavish K, Sharabi-Nov A, Binyamin D, Møller HJ, Danon D, Rothman L, et al. First trimester biomarkers for prediction of gestational diabetes mellitus. Placenta. 2020;101: 80–89. doi:10.1016/j.placenta.2020.08.020

70. Zhang X, Zhao X, Huo L, Yuan N, Sun J, Du J, et al. Risk prediction model of gestational diabetes mellitus based on nomogram in a Chinese

population cohort study. Sci Rep. 2020;10: 1–7. doi:10.1038/s41598-020-78164-x

71. Wang X, Zheng X, Yan J, Xu R, Xu M, Zheng L, et al. The Clinical Values of Afamin, Triglyceride and PLR in Predicting Risk of Gestational Diabetes During Early Pregnancy. Front Endocrinol (Lausanne). 2021;12: 1–8. doi:10.3389/fendo.2021.723650

72. Bender W, McCarthy C, Elovitz M, Parry S, Durnwald C. Universal HbA1c screening and gestational diabetes: a comparison with clinical risk factors. Journal of Maternal-Fetal and Neonatal Medicine. 2021;0: 1–7. doi:10.1080/14767058.2021.1914578

73. Maitland RA, Seed PT, Briley AL, Homsy M, Thomas S, Pasupathy D, et al. Prediction of gestational diabetes in obese pregnant women from the UK Pregnancies Better Eating and Activity (UPBEAT) pilot trial. Diabetic Medicine. 2014;31: 963–970. doi:10.1111/dme.12482

74. Abell SK, Shorakae S, Boyle JA, de Courten B, Stepto NK, Teede HJ, et al. Role of serum biomarkers to optimise a validated clinical risk prediction tool for gestational diabetes. Australian and New Zealand Journal of Obstetrics and Gynaecology. 2019;59: 251–257. doi:10.1111/ajo.12833

75. Gao S, Leng J, Liu H, Wang S, Li W, Wang Y, et al. Development and validation of an early pregnancy risk score for the prediction of gestational diabetes mellitus in Chinese pregnant women. BMJ Open Diabetes Res Care. 2020;8: e000909. doi:10.1136/bmjdrc-2019-000909

76. Nanda S, Savvidou M, Syngelaki A, Akolekar R, Nicolaides KH. Prediction of gestational diabetes mellitus by maternal factors and biomarkers at 11 to 13 weeks. Prenat Diagn. 2011;31: 135–141. doi:10.1002/pd.2636

77. Thagaard IN, Krebs L, Holm JC, Lange T, Larsen T, Christiansen M. Adiponectin and leptin as first trimester markers for gestational diabetes mellitus: A cohort study. Clin Chem Lab Med. 2017;55: 1805–1812. doi:10.1515/cclm-2017-0427

78. Zhu B, Han Y, Deng F, Huang K, Yan S, Hao J, et al. The role of triiodothyronine (T3) and t3/free thyroxine (ft4) in glucose metabolism during pregnancy: The ma'anshan birth cohort study. Endocr Connect. 2021;10: 685–693. doi:10.1530/EC-21-0088

79. Savvidou M, Nelson SM, Makgoba M, Messow C-M, Sattar N, Nicolaides K. First-Trimester Prediction of Gestational Diabetes Mellitus: Examining

the Potential of Combining Maternal Characteristics and Laboratory Measures. Diabetes. 2010;59: 3017–3022. doi:10.2337/db10-0688

80. Sweeting AN, Wong J, Appelblom H, Ross GP, Kouru H, Williams PF, et al. A Novel Early Pregnancy Risk Prediction Model for Gestational Diabetes Mellitus. Fetal Diagn Ther. 2019;45: 76–84. doi:10.1159/000486853

81. Snyder BM, Baer RJ, Oltman SP, Robinson JG, Breheny PJ, Saftlas AF, et al. Early pregnancy prediction of gestational diabetes mellitus risk using prenatal screening biomarkers in nulliparous women. Diabetes Res Clin Pract. 2020;163: 108139. doi:10.1016/j.diabres.2020.108139

82. Zhao M, Yang S, Hung TC, Zheng W, Su X. Association of pre- and early-pregnancy factors with the risk for gestational diabetes mellitus in a large Chinese population. Sci Rep. 2021;11: 1–9. doi:10.1038/s41598-021-86818-7

83. Huang Y, Chen X, Chen X, Feng Y, Guo H, Li S, et al. Angiopoietin-like protein 8 in early pregnancy improves the prediction of gestational diabetes. Diabetologia. 2018;61: 574–580. doi:10.1007/s00125-017-4505-y

84. Benhalima K, van Crombrugge P, Moyson C, Verhaeghe J, Vandeginste S, Verlaenen H, et al. Estimating the risk of gestational diabetes mellitus based on the 2013 WHO criteria: a prediction model based on clinical and biochemical variables in early pregnancy. Acta Diabetol. 2020;57: 661–671. doi:10.1007/s00592-019-01469-5

85. Francis EC, Li M, Hinkle SN, Cao Y, Chen J, Wu J, et al. Adipokines in early and mid-pregnancy and subsequent risk of gestational diabetes: A longitudinal study in a multiracial cohort. BMJ Open Diabetes Res Care. 2020;8: 1–10. doi:10.1136/bmjdrc-2020-001333

86. Harrison CL, Lombard CB, East C, Boyle J, Teede HJ. Risk stratification in early pregnancy for women at increased risk of gestational diabetes. Diabetes Res Clin Pract. 2015;107: 61–68. doi:10.1016/j.diabres.2014.09.006

87. Syngelaki A, Kotecha R, Pastides A, Wright A, Nicolaides KH. First-trimester biochemical markers of placentation in screening for gestational diabetes mellitus. Metabolism. 2015;64: 1485–1489. doi:10.1016/j.metabol.2015.07.015

88. Sweeting AN, Wong J, Appelblom H, Ross GP, Kouru H, Williams PF, et al. A first trimester prediction model for gestational diabetes utilizing aneuploidy and pre-eclampsia screening markers. Journal of Maternal-Fetal and Neonatal Medicine. 2018;31: 2122–2130. doi:10.1080/14767058.2017.1336759

89. Kang M, Zhang H, Zhang J, Huang K, Zhao J, Hu J, et al. A Novel Nomogram for Predicting Gestational Diabetes Mellitus During Early Pregnancy. Front Endocrinol (Lausanne). 2021;12: 1–8. doi:10.3389/fendo.2021.779210

90. Thériault S, Giguère Y, Massé J, Girouard J, Forest JC. Early prediction of gestational diabetes: A practical model combining clinical and biochemical markers. Clin Chem Lab Med. 2016;54: 509–518. doi:10.1515/cclm-2015-0537

91. Liu H, Li J, Leng J, Wang H, Liu J, Li W, et al. Machine learning risk score for prediction of gestational diabetes in early pregnancy in Tianjin, China. Diabetes Metab Res Rev. 2021;37. doi:10.1002/dmrr.3397

92. Wang Y, Ge Z, Chen L, Hu J, Zhou W, Shen S, et al. Risk Prediction Model of Gestational Diabetes Mellitus in a Chinese Population Based on a Risk Scoring System. Diabetes Therapy. 2021;12: 1721–1734. doi:10.1007/s13300-021-01066-2

93. Zhang Y, Zhou L, Tian L, Li X, Zhang G, Qin J, et al. A mid-pregnancy risk prediction model for gestational diabetes mellitus based on the maternal status in combination with ultrasound and serological findings. Exp Ther Med. 2020;20: 293–300. doi:10.3892/etm.2020.8690

94. Qiu H, Yu HY, Wang LY, Yao Q, Wu SN, Yin C, et al. Electronic Health Record Driven Prediction for Gestational Diabetes Mellitus in Early Pregnancy. Sci Rep. 2017;7: 1–13. doi:10.1038/s41598-017-16665-y

95. Wu YT, Zhang CJ, Mol BW, Kawai A, Li C, Chen L, et al. Early Prediction of Gestational Diabetes Mellitus in the Chinese Population via Advanced Machine Learning. Journal of Clinical Endocrinology and Metabolism. 2021;106: E1191–E1205. doi:10.1210/clinem/dgaa899

96. Ye Y, Xiong Y, Zhou Q, Wu J, Li X, Xiao X. Comparison of Machine Learning Methods and Conventional Logistic Regressions for Predicting Gestational Diabetes Using Routine Clinical Data: A Retrospective Cohort Study. J Diabetes Res. 2020;2020: 1–10. doi:10.1155/2020/4168340

97. McLaren R, Haberman S, Moscu M, Atallah F, Friedmann H. A Novel and Precise Profiling Tool to Predict Gestational Diabetes. J Diabetes Sci Technol. 2021;15: 891–896. doi:10.1177/1932296820948883

98. Artzi NS, Shilo S, Hadar E, Rossman H, Barbash-Hazan S, Ben-Haroush A, et al. Prediction of gestational diabetes based on nationwide electronic health records. Nat Med. 2020;26: 71–76. doi:10.1038/s41591-019-0724-8

99. Zhang H, Zhao Y, Zhao D, Chen X, Khan NU, Liu X, et al. Potential biomarkers identified in plasma of patients with gestational diabetes mellitus. Metabolomics. 2021;17: 1–12. doi:10.1007/s11306-021-01851-x

100. Liu X, Wang X, Sun H, Guo Z, Liu X, Yuan T, et al. Urinary metabolic variation analysis during pregnancy and application in Gestational Diabetes Mellitus and spontaneous abortion biomarker discovery. Sci Rep. 2019;9: 2605. doi:10.1038/s41598-019-39259-2

101. Raczkowska BA, Mojsak P, Rojo D, Telejko B, Paczkowska–Abdulsalam M, Hryniewicka J, et al. Gas Chromatography–Mass Spectroscopy-Based Metabolomics Analysis Reveals Potential Biochemical Markers for Diagnosis of Gestational Diabetes Mellitus. Front Pharmacol. 2021;12: 1–9. doi:10.3389/fphar.2021.770240

102. Mokkala K, Vahlberg T, Pellonperä O, Houttu N, Koivuniemi E, Laitinen K. Distinct Metabolic Profile in Early Pregnancy of Overweight and Obese Women Developing Gestational Diabetes. Journal of Nutrition. 2020;150: 31–37. doi:10.1093/jn/nxz220

103. McMichael LE, Heath H, Johnson CM, Fanter R, Alarcon N, Quintana-Diaz A, et al. Metabolites involved in purine degradation, insulin resistance, and fatty acid oxidation are associated with prediction of Gestational diabetes in plasma. Metabolomics. 2021;17: 1–11. doi:10.1007/s11306-021-01857-5

104. Koos BJ, Gornbein JA. Early pregnancy metabolites predict gestational diabetes mellitus: implications for fetal programming. Am J Obstet Gynecol. 2021;224: 215.e1-215.e7. doi:10.1016/j.ajog.2020.07.050

105. Piras C, Neri I, Pintus R, Noto A, Petrella E, Monari F, et al. First trimester metabolomics 1H-NMR study of the urinary profile predicts gestational diabetes mellitus development in obese women. Journal of Maternal-Fetal and Neonatal Medicine. 2021;0: 1–9. doi:10.1080/14767058.2021.1970133

106. Sakurai K, Eguchi A, Watanabe M, Yamamoto M, Ishikawa K, Mori C. Exploration of predictive metabolic factors for gestational diabetes mellitus in Japanese women using metabolomic analysis. J Diabetes Investig. 2019;10: 513–520. doi:10.1111/jdi.12887

107. Diaz SO, Pinto J, Graça G, Duarte IF, Barros AS, Galhano E, et al. Metabolic biomarkers of prenatal disorders: An exploratory NMR metabonomics study of second trimester maternal urine and blood plasma. J Proteome Res. 2011;10: 3732–3742. doi:10.1021/pr200352m

108. Tian M, Ma S, You Y, Long S, Zhang J, Guo C, et al. Serum Metabolites as an Indicator of Developing Gestational Diabetes Mellitus Later in the Pregnancy: A Prospective Cohort of a Chinese Population. J Diabetes Res. 2021;2021. doi:10.1155/2021/8885954

109. Guo Y, Han Z, Guo L, Liu Y, Li G, Li H, et al. Identification of urinary biomarkers for the prediction of gestational diabetes mellitus in early second trimester of young gravidae based on iTRAQ quantitative proteomics. Endocr J. 2018;65: 727–735. doi:10.1507/endocrj.EJ17-0471

110. Zhao D, Shen L, Wei Y, Xie J, Chen S, Liang Y, et al. Identification of candidate biomarkers for the prediction of gestational diabetes mellitus in the early stages of pregnancy using iTRAQ quantitative proteomics. Proteomics Clin Appl. 2017;11: 7–8. doi:10.1002/prca.201600152

111. Yoffe L, Polsky A, Gilam A, Raff C, Mecacci F, Ognibene A, et al. Early diagnosis of gestational diabetes mellitus using circulating microRNAs. Eur J Endocrinol. 2019;181: 565–577. doi:10.1530/EJE-19-0206

112. Zhao C, Dong J, Jiang T, Shi Z, Yu B, Zhu Y, et al. Early Second-Trimester Serum MiRNA Profiling Predicts Gestational Diabetes Mellitus. PLoS One. 2011;6: e23925. doi:10.1371/journal.pone.0023925

113. Nair S, Guanzon D, Jayabalan N, Lai A, Scholz-Romero K, Kalita de Croft P, et al. Extracellular vesicle-associated miRNAs are an adaptive response to gestational diabetes mellitus. J Transl Med. 2021;19: 1–17. doi:10.1186/s12967-021-02999-9

114. del Vecchio G, Li Q, Li W, Thamotharan S, Tosevska A, Morselli M, et al. Cell-free DNA Methylation and Transcriptomic Signature Prediction of Pregnancies with Adverse Outcomes. Epigenetics. 2021;16: 642–661. doi:10.1080/15592294.2020.1816774

115. Yu XY, Song LP, Zheng HT, Wei SD, Wen XL, Huang B, et al. Association between functional genetic variants in retinoid X receptor-α/γ and the risk of gestational diabetes mellitus in a southern Chinese population. Biosci Rep. 2021;41: 1–10. doi:10.1042/BSR20211338

116. Kawai VK, Levinson RT, Adefurin A, Kurnik D, Collier SP, Conway D, et al. A genetic risk score that includes common type 2 diabetes risk variants is associated with gestational diabetes. Clin Endocrinol (Oxf). 2017;87: 149–155. doi:10.1111/cen.13356

117. Lamri A, Mao S, Desai D, Gupta M, Paré G, Anand SS. Fine-tuning of Genome-Wide Polygenic Risk Scores and Prediction of Gestational Diabetes in South Asian Women. Sci Rep. 2020;10: 8941. doi:10.1038/s41598-020-65360-y

118. Zheng W, Xu Q, Huang W, Yan Q, Chen Y, Zhang L, et al. Gestational Diabetes Mellitus Is Associated with Reduced Dynamics of Gut Microbiota during the First Half of Pregnancy. mSystems. 2020;5: 1–13. doi:10.1128/msystems.00109-20

119. Boisvert MR, Koski KG, Burns DH, Skinner CD. Prediction of gestational diabetes mellitus based on an analysis of amniotic fluid by capillary electrophoresis. Biomark Med. 2012;6: 645–653. doi:10.2217/bmm.12.53

120. Mahajan A, Go MJ, Zhang W, Below JE, Gaulton KJ, Ferreira T, et al. Genome-wide trans-ancestry meta-analysis provides insight into the genetic architecture of type 2 diabetes susceptibility. Nat Genet. 2014;46: 234–244. doi:10.1038/ng.2897

121. Lu L, Koulman A, Petry CJ, Jenkins B, Matthews L, Hughes IA, et al. An unbiased lipidomics approach identifies early second trimester lipids predictive of maternal glycemic traits and gestational diabetes mellitus. Diabetes Care. 2016;39: 2232–2239. doi:10.2337/dc16-0863

122. Nevalainen J, Sairanen M, Appelblom H, Gissler M, Timonen S, Ryynänen M. First-trimester maternal serum amino acids and acylcarnitines are significant predictors of gestational diabetes. Review of Diabetic Studies. 2016;13: 236–245. doi:10.1900/RDS.2016.13.236

123. Jiang R, Wu S, Fang C, Wang C, Yang Y, Liu C, et al. Amino acids levels in early pregnancy predict subsequent gestational diabetes. J Diabetes. 2020;12: 503–511. doi:10.1111/1753-0407.13018

124. Li J, Huo X, Cao YF, Li SN, Du Z, Shao P, et al. Bile acid metabolites in early pregnancy and risk of gestational diabetes in Chinese women: A nested case-control study. EBioMedicine. 2018;35: 317–324. doi:10.1016/j.ebiom.2018.08.015

125. White SL, Lawlor DA, Briley AL, Godfrey KM, Nelson SM, Oteng-Ntim E, et al. Early antenatal prediction of gestational diabetes in obese women: Development of prediction tools for targeted intervention. PLoS One. 2016;11: 1–17. doi:10.1371/journal.pone.0167846

126. Enquobahrie DA, Denis M, Tadesse MG, Gelaye B, Ressom HW, Williams MA. Maternal early pregnancy serum metabolites and risk of gestational diabetes mellitus. Journal of Clinical Endocrinology and Metabolism. 2015;100: 4348–4356. doi:10.1210/jc.2015-2862

127. Liu J, Li J, Li W, Li N, Huo X, Wang H, et al. Predictive values of serum metabolites in early pregnancy and their possible pathways for gestational diabetes: A nested case-control study in Tianjin, China. J Diabetes Complications. 2021;35: 108048. doi:10.1016/j.jdiacomp.2021.108048

128. Wang Y, Huang Y, Wu P, Ye Y, Sun F, Yang X, et al. Plasma lipidomics in early pregnancy and risk of gestational diabetes mellitus: a prospective nested case-control study in Chinese women. Am J Clin Nutr. 2021;114: 1763–1773. doi:10.1093/ajcn/nqab242

129. Ravnsborg T, Svaneklink S, Andersen LLT, Larsen MR, Jensen DM, Overgaard M. First-trimester proteomic profiling identifies novel predictors of gestational diabetes mellitus. PLoS One. 2019;14: 1–13. doi:10.1371/journal.pone.0214457

130. Ravnsborg T, Andersen LLT, Trabjerg ND, Rasmussen LM, Jensen DM, Overgaard M. First-trimester multimarker prediction of gestational diabetes mellitus using targeted mass spectrometry. Diabetologia. 2016;59: 970–979. doi:10.1007/s00125-016-3869-8

131. Sørensen A, van Poppel M, Desoye G, Damm P, Simmons D, Jensen D, et al. The Predictive Value of miR-16, -29a and -134 for Early Identification of Gestational Diabetes: A Nested Analysis of the DALI Cohort. Cells. 2021;10: 170. doi:10.3390/cells10010170

132. Guo Z, Yang F, Zhang J, Zhang Z, Li K, Tian Q, et al. Whole-Genome Promoter Profiling of Plasma DNA Exhibits Diagnostic Value for Placenta-Origin Pregnancy Complications. Advanced Science. 2020;7: 1901819. doi:10.1002/advs.201901819

133. Krishnan M, Murphy R, Okesene-Gafa KAM, Ji M, Thompson JMD, Taylor RS, et al. The Pacific-specific CREBRF rs373863828 allele protects against gestational diabetes mellitus in Māori and Pacific women with obesity. Diabetologia. 2020;63: 2169–2176. doi:10.1007/s00125-020-05202-8

134. Popova P v., Klyushina AA, Vasilyeva LB, Tkachuk AS, Vasukova EA, Anopova AD, et al. Association of Common Genetic Risk Variants With Gestational Diabetes Mellitus and Their Role in GDM Prediction. Front Endocrinol (Lausanne). 2021;12: 1–7. doi:10.3389/fendo.2021.628582

135. Dziedziejko V, Safranow K, Tarnowski M, Pawlik A. Common Type 2 Diabetes Genetic Risk Variants Improve the Prediction of Gestational Diabetes. Hormone and Metabolic Research. 2019;51: 655–660. doi:10.1055/a-0945-0328

136. Ma S, You Y, Huang L, Long S, Zhang J, Guo C, et al. Alterations in Gut Microbiota of Gestational Diabetes Patients During the First Trimester of Pregnancy. Front Cell Infect Microbiol. 2020;10. doi:10.3389/fcimb.2020.00058

137. Hu P, Chen X, Chu X, Fan M, Ye Y, Wang Y, et al. Association of gut microbiota during early pregnancy with risk of incident gestational diabetes mellitus. Journal of Clinical Endocrinology and Metabolism. 2021;106: E4128–E4141. doi:10.1210/clinem/dgab346

138. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement. Ann Intern Med. 2015;162: 55–63. doi:10.7326/M14-0697

4.3. Evaluation of first and second trimester maternal thyroid profile on the prediction of gestational diabetes mellitus and post load glycemia

En esta sección se presentan y discuten los resultados que dieron origen al artículo “Evaluation of first and second trimester maternal thyroid profile on the prediction of gestational diabetes mellitus and post load glycemia”, publicado en la revista *PLOS ONE* (https://doi.org/10.1371/journal.pone.0280513).

Evaluation of first and second trimester maternal thyroid profile on the prediction of gestational diabetes mellitus and post load glycemia

Daniela Mennickent [1,2,3], Bernel Ortega-Contreras [1], Sebastián Gutiérrez-Vega [1], Erica Castro [4], Andrés Rodríguez [3,5,6], Juan Araya [2,3,*], Enrique Guzmán-Gutiérrez [1,3,6,7,*]

[1] Departamento de Bioquímica Clínica e Inmunología, Facultad de Farmacia, Universidad de Concepción, Concepción, Chile. [2] Departamento de Análisis Instrumental, Facultad de Farmacia, Universidad de Concepción, Concepción, Chile. [3] Machine Learning Applied in Biomedicine (MLAB), Chile. [4] Departamento de Obstetricia y Puericultura, Facultad de Ciencias de la Salud, Universidad de Atacama, Copiapó, Chile. [5] Departamento de Ciencias Básicas, Facultad de Ciencias, Universidad del Bío-Bío, Chillán, Chile. [6] Group of Research and Innovation in Vascular Health (GRIVAS-Health), Chile. [7] Millennium Institute on Immunology and Immunotherapy (IMII), Chile.

* Corresponding authors: eguzman@udec.cl (EG-G); jarayaq@udec.cl (JA)

Abstract

Maternal thyroid alterations have been widely associated with the risk of gestational diabetes mellitus (GDM). This study aims to 1) test the first and the second trimester full maternal thyroid profile on the prediction of GDM, both alone and combined with non-thyroid data; and 2) make that prediction independent of the diagnostic criteria, by evaluating the effectiveness of the different maternal variables on the prediction of oral glucose tolerance test (OGTT) post load glycemia. Pregnant women were recruited in Concepción, Chile. GDM diagnosis was performed at 24-28 weeks of pregnancy by an OGTT (n=54 for normal glucose tolerance, n=12 for GDM). 75 maternal thyroid and non-thyroid parameters were recorded in the first and the second trimester of pregnancy. Various combinations of variables were assessed for GDM and post load glycemia prediction through different classification and regression machine learning techniques. The best predictive models were simplified by variable selection. Every model was subjected to leave-one-out cross-validation. Our results indicate that thyroid markers are useful for the prediction of GDM and post load glycemia, especially at the second trimester of pregnancy. Thus, they could be used as an alternative screening tool for GDM, independently of the diagnostic criteria used. The final classification models predict GDM with cross-validation areas under the receiver operating characteristic curve of 0.867 ($p<0.001$) and 0.920 ($p<0.001$) in the first and the second trimester of pregnancy, respectively.

The final regression models predict post load glycemia with cross-validation Spearman r correlation coefficients of 0.259 ($p=0.036$) and 0.457 ($p<0.001$) in the first and the second trimester of pregnancy, respectively. This investigation constitutes the first attempt to test the performance of the whole maternal thyroid profile on GDM and OGTT post load glycemia prediction. Future external validation studies are needed to confirm these findings in larger cohorts and different populations.

4.3.1. Introduction

Gestational diabetes mellitus (GDM) is a hyperglycemia state that is diagnosed at the second or third trimester of pregnancy, with glycemia values that vary depending on the diagnostic criteria [1]. This disorder has been linked to several negative effects on maternal and fetal health, short and long-term [2–4]. The improvement of its diagnosis, whether with early detection tools or with alternative screening tools, could help to prevent those adverse pregnancy outcomes. Predictive models, with “prediction" understood as the forecast of future behaviors or unobserved outcomes [5], are an auspicious means to achieve that goal.

GDM prediction has been typically addressed with strategies based on classic clinical risk factors for this pregnancy complication, i.e. maternal age, body mass index (BMI), family history of diabetes, previous GDM, among others.

Nevertheless, approaches that consider only in this type of parameters usually have a poor predictive performance for GDM [6–8]. Thus, other biomarkers should be investigated and tested in order to complement and boost them.

Maternal thyroid alterations in the first and the second trimester of pregnancy have been widely related to the risk of GDM [9–12]. Moreover, certain authors have assessed first and second trimester thyroid parameters on the prediction of this pregnancy disorder [13–19]. However, the majority of them considered only a few thyroid markers, instead of the whole thyroid profile [13–18]. Furthermore, various of them didn't evaluate these markers in association with other variables [13–15], which is usually related to an improved predictive accuracy [20]. In addition, even though some authors used thyroid data in multivariate analysis, they didn't assess the particular effect of the thyroid profile on GDM prediction, whether by testing it solo or by adding/removing it from the analysis [16–19].

Therefore, this investigation aims to: 1) evaluate the first and the second trimester maternal thyroid profile on GDM prediction, both alone and combined with non-thyroid parameters; and 2) make that prediction independent of the diagnostic criteria, by testing the performance of the different maternal variables on post load glycemia prediction.

4.3.2. Materials and methods

4.3.2.1. Ethical aspects

This biomedical study was approved by the Ethics Committee of the Servicio de Salud Concepción (17-12-88) and was performed in accordance to the World Medical Association Declaration of Helsinki.

4.3.2.2. Subjects recruitment

Pregnant women with 10-14 gestational weeks were recruited between 2017 and 2019 at three primary health centers (CESFAM Victor Manuel Fernández, CESFAM Santa Sabina and CESFAM Tucapel) in Concepción, Chile. Subjects with pregestational diabetes mellitus, pregestational thyroid disorders, or any pregnancy pathologies different than GDM were excluded. Women that gave their written informed consent were incorporated into the study, and followed through pregnancy until 24-28 gestational weeks.

4.3.2.3. Blood samples collection

Blood samples were collected on the first (at 10-14 gestational weeks) and the second (at 24-28 gestational weeks) trimester of pregnancy, either after 12 hours

of fasting or after 2 hours of a 75 g glucose load. Samples were transported to laboratory at 4˚C. Sera (obtained from tubes with no anticoagulant) and plasma (obtained from tubes with sodium fluoride/citrate as anticoagulant) were aliquoted and stored at -80˚C. White blood cells, obtained from EDTA tubes, were stored at -20ºC.

4.3.2.4. GDM diagnosis and study groups

Pregnant women were subjected to an oral glucose tolerance test (OGTT) at 24-28 gestational weeks. Subjects with post load glycemia > 140 mg/dL (75 g, 2 h) were diagnosed with GDM. In this study, out of 66 subjects, 12 had GDM. The remaining 54 exhibited no alteration in fasting or post load glycemia, i.e. had normal glucose tolerance (NGT).

4.3.2.5. Clinical data collection

First trimester maternal clinical variables were obtained from both health centers records and pregnant women statements. Institution-derived parameters involved: age, height, weight, BMI, systolic and diastolic pressure at the first trimester of pregnancy. Self-reported data comprised: supplement consumption, hyperemesis, vaginal bleeding and drug use at the first trimester of pregnancy; and preconception information, i.e. drug use before pregnancy, prior pregnancy

complications, prior non-viable pregnancy, fertility problems, polycystic ovary syndrome (PCOS) history, first period age, last period date, personal and family morbid history.

4.3.2.6. Biochemical determinations

Fasting and post load plasma glucose were measured using the hexokinase method [21]. Thyroid stimulating hormone (TSH), total triiodothyronine (TT3), total thyroxine (TT4), free T4 (FT4), thyroglobulin (TG), TG antibody (aTG), thyroid peroxidase antibody (aTPO), TSH receptor antibody (TRAb) and fat mass and obesity-associated (FTO) rs9939609 genotype were determined as described by Araya et al [22]. Briefly, serum thyroid hormones and antibodies were measured by different immunoassays. For FTO genotyping, genomic DNA was isolated from white blood cells, and then analyzed by a previously optimized PCR-HRM method. Melt curve results were normalized to identify the subject's genotype by allele comparison with control samples of known genotype.

4.3.2.7. Univariate data analysis

Between the first and second trimester of pregnancy, a total of 75 maternal clinical and biochemical variables were recorded. Qualitative data were compared by two-sided Fisher exact test. Quantitative data normality was assessed through

Shapiro-Wilk test. Normally distributed variables were compared by unpaired Student t test. Not-normally distributed variables were compared by Mann-Whitney or Wilcoxon matched-pairs signed rank test in unpaired or paired analyses, respectively. Values of $p<0.05$ were considered statistically significant. Univariate analyses were performed with the software GraphPad Prism 9.2.0 (GraphPad Software Inc, USA).

4.3.2.8. Multivariate data analysis

For multivariate analyses, almost all the registered maternal variables were used as predictors, except for height, weight and OGTT glycemia. Height and weight weren't considered as individual predictors. Instead, the BMI, which comprises both, was utilized. The OGTT glycemia was employed as a label (i.e. NGT or GDM, for exploration) or as the property to predict (i.e. NGT or GDM, for classification; or the OGTT post load glycemia value, for regression).

For classification and regression analyses, the predictors were divided into four categories: thyroid from the first (Thy1T) and the second (Thy2T) trimester of pregnancy, which comprised TSH, TT3, TT4, FT4, TG, aTG, aTPO and TRAb; and non-thyroid from the first (NoThy1T) and the second (NoThy2T) trimester of pregnancy, which included every other recorded predictor. All the possible

combinations of these four categories were assessed for GDM and post load glycemia prediction.

Prior to multivariate analyses, qualitative variables were transformed into categorical and all data were normalized by auto-scaling.

4.3.2.8.1. Exploratory analysis

Full auto-scaled data were explored by principal components analysis (PCA). To detect possible outliers, four parameters were evaluated at this stage: sample residual, F ratio, probability and Mahalanobis distance. Subjects were considered outliers if two or more parameters suggested such behavior. Data exploration was performed with the software Pirouette 4.5 (Infometrix Inc, USA).

4.3.2.8.2. Classification analysis

Different subgroups of auto-scaled data were assessed for GDM prediction through five classification ML techniques: logistic regression (LR), linear support vector machine (L-SVM), partial least squares discriminant analysis (PLS-DA), classification and regression tree (CART) and extreme gradient boosting (XGB). Every model was subjected to leave-one-out cross-validation. LR, PLS-DA and XGBoost were executed in PLS_Toolbox R8.9.2 (Eigenvector Research Inc,

USA). L-SVM and CART were implemented by coding in MATLAB R2021a (The MathWorks Inc, USA).

The classification predictive performance was assessed by the determination of the models sensitivity (Se), specificity (Sp) and non-error rate (NER) in both calibration and cross-validation. These parameters were calculated as follows:

$$Se\ (\%) = \frac{TP}{TP + FN} \cdot 100 \qquad (1)$$

$$Sp\ (\%) = \frac{TN}{TN + FP} \cdot 100 \qquad (2)$$

$$NER\ (\%) = \frac{Se + Sp}{2} \qquad (3)$$

Where TP, FN, TN and FP are the number of true positives, false negatives, true negatives and false positives, respectively.

The area under the receiver operating characteristic curve (AUC) of the final classification models was determined through the software GraphPad Prism 9.2.0 (GraphPad Software Inc, USA).

4.3.2.8.3. Regression analysis

Different subgroups of auto-scaled data were assessed for post load glycemia prediction through the regression ML technique partial least squares (PLS). Every model was subjected to leave-one-out cross-validation. PLS was executed in PLS_Toolbox R8.9.2 (Eigenvector Research Inc, USA).

The regression predictive performance was assessed by the determination of the models root mean square error (RMSE) and relative error (RE) in both calibration (RMSEC and REC) and cross validation (RMSECV and RECV). These parameters were calculated as follows:

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n-1}} \qquad (4)$$

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \qquad (5)$$

$$REC\ (\%) = \frac{RMSEC}{\bar{y}} \cdot 100 \qquad (6)$$

$$RECV\ (\%) = \frac{RMSECV}{\bar{y}} \cdot 100 \qquad (7)$$

Where $\hat{y}_i$ is the predicted post load glycemia for subject i, y_i is the actual post load glycemia for subject i, n is the number of subjects in calibration and cross-validation, and $\bar{y}$ is the mean of the actual post load glycemia values in calibration. The Spearman r correlation coefficient of the final regression models was determined by the software GraphPad Prism 9.2.0 (GraphPad Software Inc, USA).

4.3.2.8.4. Variable selection

Variable importance was retrieved from PLS-DA and PLS regression vectors. For every model, the contribution of each variable was calculated as follows:

$$Variable\ contribution\ (\%) = \frac{|b_i|}{\sum_{i=1}^{n} |b_i|} \cdot 100 \qquad (8)$$

Where b_i is the regression vector value for variable i, and n is the number of variables in the classification or regression model.

For each model, variables were sorted according to their percentage contribution. Toward model simplification, the j variables with the highest percentage contribution were selected, with j ranging from 10 to 1.

4.3.3. Results and discussion

4.3.3.1. Description of the study groups

To characterize and compare the study groups, classic univariate techniques were used. **Table 4.3-1** shows only the first and second trimester maternal variables that differed significantly between NGT and GDM pregnancies in the univariate approach. The full list, which contains the 75 maternal variables that were recorded and analyzed in this study, is presented in **Table 4.3-S1**.

Table 4.3-1. Maternal variables that differ significantly between normal glucose tolerance and gestational diabetes mellitus pregnancies in univariate analysis.

Variable		Unit	NGT (n=54)	GDM (n=12)	p value		All (n=66)
First trimester							
BMI		Kg/m^2	27.4 (23.5-31.1)	30.3 (27.1-31.6)	0.046	*	27.8 (24.5-31.2)
Systolic pressure		mmHg	109 ± 11	116 ± 7	0.042	*	110 ± 11
Preconception data							
Prior GDM		%	1.9 (1/54)	33.3 (4/12)	0.003	**	7.6 (5/66)
Personal history of anemia		%	0.0 (0/54)	16.7 (2/12)	0.031	*	3.0 (2/66)
Family history of DM2		%	20.4 (11/54)	66.7 (8/12)	0.003	**	28.8 (19/66)
Family history of hypertension		%	33.3 (18/54)	66.7 (8/12)	0.049	*	39.4 (26/66)
FTO genotype (rs9939609)	TT	%	53.7 (29/54)	25.0 (3/12)	0.047	*	48.5 (32/66)
	TA	%	44.4 (24/54)	66.7 (8/12)			48.5 (32/66)
	AA	%	1.9 (1/54)	8.3 (1/12)			3.0 (2/66)
Second trimester							
TSH		µIU/mL	1.42 (1.10-2.09)	3.25 (2.28-4.38)	0.001	**	1.54 (1.10-2.53)
TT3		ng/mL	1.92 (1.90-2.04)	2.02 (1.92-2.37)	0.027	*	1.92 (1.90-2.11)
FT4		ng/dL	0.79 (0.76-0.81)	0.68 (0.63-0.78)	0.006	**	0.77 (0.76-0.81)
aTG		IU/mL	4.50 (4.30-7.05)	16.93 (8.42-18.45)	0.020	*	4.89 (4.31-16.85)
Fasting glycemia		mg/dL	79 ± 8	86 ± 10	0.013	*	80 ± 9
OGTT glycemia (75 g, 2 h)		mg/dL	102 (94-109)	150 (141-172)	<0.001	****	105 (96-134)

*NGT: normal glucose tolerance. GDM: gestational diabetes mellitus. BMI: body mass index. DM2: type 2 diabetes. FTO: fat mass and obesity-associated gene. TSH: thyroid stimulating hormone. TT3: total triiodothyronine. FT4: free thyroxine. aTG: thyroglobulin antibody. OGTT: oral glucose tolerance test. Qualitative variables are presented as percentage (proportion); quantitative variables with normal distribution, as mean ± standard deviation; and quantitative variables with non-normal distribution, as median (interquartile range). * p<0.05. ** p<0.01. **** p<0.0001.*

In the first trimester of pregnancy, the BMI and the systolic pressure were higher in the GDM group, than in the NGT group. Moreover, prior history of GDM, family history of type 2 diabetes (DM2) and family history of hypertension were more frequent in GDM pregnancies. These observations are consistent with a recent Chilean epidemiologic study, in which subjects with a higher pre-pregnancy BMI, previous record of hypertension or GDM, or family history of DM2 or hypertension, presented a greater risk of GDM [23]. Likewise, personal history of anemia was more frequent in GDM women. The relationship between anemia and GDM appears to be dependent on the underlying cause of the former. Indeed, whereas iron deficiency decreases the risk of GDM [24], B12 deficiency increases it [25]. Interestingly, although iron deficiency was the most common cause of anemia among Chilean women between 1981 and 2010 [26], the 2010–2011 National Food Consumption Survey revealed that, while approximately 30% of women had a deficient consumption of iron, nearly 60% of them had insufficient intake of vitamin B12 [27]. The FTO genotype was significantly different between the two groups as well: while the TT genotype was more common in NGT women, the TA and the AA genotypes were more frequent in GDM women. Contradictory results have been reported regarding to the association between FTO rs9939609 polymorphism and GDM [28–30]. However, in Chilean adults the risk allele A was linked to an increased BMI [31], which is a risk factor for GDM in this population [23].

In the second trimester of pregnancy, TSH, TT3 and aTG were higher in the GDM group, when compared to the NGT group. In addition, FT4 was lower in GDM pregnancies. Preceding studies describe similar results. For instance, Gutierrez-Vega et al reported that GDM subjects had greater levels of TSH and TT3 than NGT subjects at 24-28 weeks of pregnancy [32]. Likewise, Tang et al found increased levels of aTG and decreased levels of FT4 in GDM pregnant women at 17.0 ± 4.0 weeks of gestation [14]. Furthermore, second trimester fasting glycemia and OGTT post load glycemia were higher in the GDM group, which is in agreement with its diagnostic criteria. The prevalence of GDM in this study was 18.2%, moderately superior than the 13.0% described for Chile in 2015 [23] and to the 11.2% reported for South and Central America in 2015-2018 [1].

To better understand the differences observed between NGT and GDM groups in terms of thyroid markers, a complementary analysis was performed. For each group, TSH, TT3, FT4 and aTG levels were compared between the first and the second trimester of pregnancy. The results are displayed in **Table 4.3-S2**. TSH didn't vary significantly in any of the groups, although it tended to increase from the first to the second trimester, especially in GDM subjects. The latter accounts for the statistical difference between NGT and GDM pregnancies regarding TSH2T (**Table 4.3-1**). TT3 raised and FT4 decreased from the first to the second trimester in both groups. The magnitude of those changes was greater in GDM women, which explains the significant difference amidst the two groups in terms

of TT32T and FT42T (**Table 4.3-1**). aTG lowered from the first to the second trimester in the NGT subjects, whereas it didn't vary in GDM subjects. This contrasting behavior gives a rational for the statistical difference between NGT and GDM pregnancies in relation to aTG2T (**Table 4.3-1**). The changes observed in thyroid hormones within the NGT group are consistent with the physiology of normal pregnancy, which is characterized by is a slight rise in TSH, a considerable increase in TT3 and a substantial decrease in FT4 between early- and mid-gestation [33]. The fact that these changes are exacerbated in GDM is not yet fully understood, however it has been proposed that it may be linked to maternal weight. High maternal weight, as high maternal BMI, is an important risk factor for GDM. Haddow et al proposed that higher maternal weight would induce higher peripheral deiodinase activity, which in turn would lead to reduced FT4 and increased TT3. Moreover, higher TT3 would induce higher plasma glucose levels, thereby contributing to GDM pathophysiology [34]. This hypothesis, although just beginning to be tested by other authors [35–37], provides a feasible explanation for the differences observed between GDM and NGT pregnancies in this study, especially considering that the GDM group presented a higher BMI than the NGT group. In a similar manner, the changes observed in aTG within the NGT group are in line with previous reports, in which the levels of this thyroid antibody decrease from early- to mid-gestation in normal pregnancies [38–40]. This decline is due to the physiological immunosuppression that occurs in normal pregnancy to tolerate the fetus and placenta [41]. That process is disrupted in GDM [42],

which could explain why the levels of aTG do not decrease amidst trimesters in the GDM group as they do in the NGT group.

These results indicate that, when considered individually, the main differential variables between the two study groups are non-thyroid in the first trimester of pregnancy, and thyroid in the second trimester of pregnancy.

4.3.3.2. Exploration of maternal data by PCA

To assess the potential of a multivariate approach to distinguish GDM from NGT pregnancies, a PCA was performed. The unsupervised analysis comprised 72 maternal variables and all the 66 subjects. No outliers were found. The PCA results are displayed in **Figure 4.3-1**.

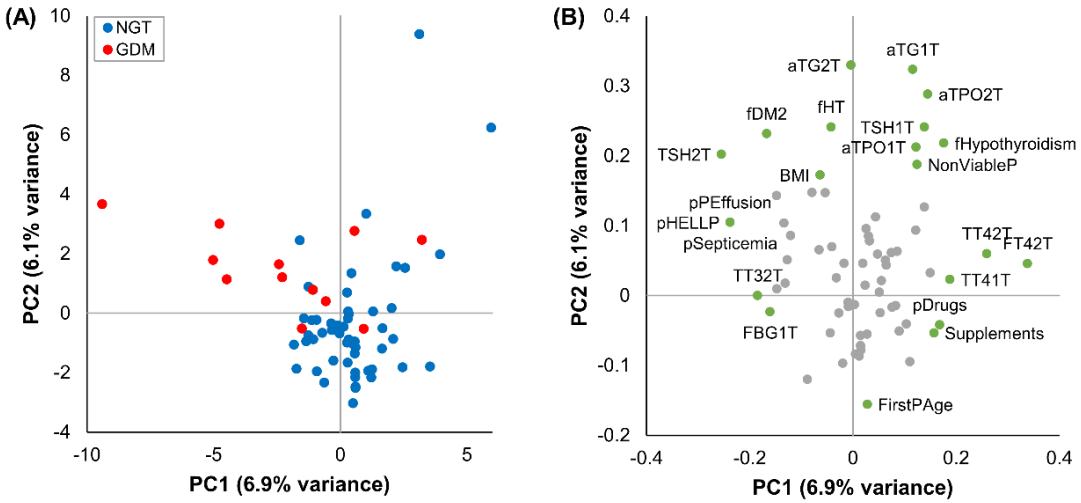


Figure 4.3-1. Exploration of maternal data by principal component analysis. (A) The scores plot shows the NGT (blue) and GDM (red) subjects projected in the space constituted by PC1 and PC2. (B) The loadings plot shows the variables contribution to PC1 and PC2. The ones that contribute the most to each PC are highlighted (green) and labeled. *NGT: normal glucose tolerance. GDM: gestational diabetes mellitus. PC: principal component. 1T: first trimester. 2T: second trimester. pHELLP: prior hemolysis elevated liver enzymes and low platelets syndrome. pPEffusion: prior pleural effusion. pSepticemia: prior septicemia. fDM2: family history of type 2 diabetes. fHT: family history of hypertension. BMI: body mass index. fHypothyroidism: family history of hypothyroidism. NonViableP: prior non-viable pregnancy. pDrugs: drug use before pregnancy. Supplements: supplement consumption. FirstPAge: first period age. TSH: thyroid stimulating hormone. aTG: thyroglobulin antibody. aTPO: thyroid peroxidase antibody. TT4: total thyroxine. FT4: free thyroxine. TT3: total triiodothyronine. FBG: fasting blood glucose.*

Figure 4.3-1A shows that NGT and GDM women are partially separated by the conjunction of principal components 1 (PC1) and 2 (PC2). **Figure 4.3-1B** highlights the maternal variables that, when combined, contribute the most to the separation between the two study groups. They are mainly non-thyroid variables from the first trimester, and thyroid variables from the second trimester of pregnancy, which is consistent with the univariate analysis results. However, although more discreetly, some thyroid variables from the first trimester also

appear to contribute to the separation between NGT and GDM subjects. A previous study reported similar results. Araya et al used PCA to explore clinical and biochemical maternal data from NGT and GDM pregnancies. After removing the OGTT post load glycemia from the analysis, the separation between two groups was mostly attributable to first trimester clinical variables and second trimester thyroid parameters, with a minor but noticeable contribution of first trimester thyroid data [22]. These observations are also in agreement with other authors, who described that different second trimester thyroid markers are more strongly associated with the risk of GDM, than their first trimester counterparts [14,16,43–45].

These results suggest that a supervised multivariate model, i.e. a ML-based model, built with first and second trimester thyroid parameters could be successful in the prediction of GDM.

4.3.3.3. Prediction of GDM with ML techniques

To evaluate the ability of maternal thyroid markers to predict GDM, different classification models were developed, covering various subgroups of predictors and ML techniques. **Table 4.3-2** displays the predictive performance of each model in calibration and cross-validation. It focuses on the NER, which is a global measure of classification quality that doesn't depend on the relative sizes of the

groups of interest [46]. In other words, the NER isn't biased by the presence of unbalanced groups, which is the case of this study: the NGT and GDM groups comprise 54 and 12 subjects, respectively. The sensitivity and specificity values of each model are presented in **Table 4.3-S3**.

Table 4.3-2. Non-error rate of machine learning models that predict gestational diabetes mellitus based on different maternal variables.

Maternal predictors	Calibration NER					Cross-validation NER				
	LR	L-SVM	PLS-DA	CART	XGB	LR	L-SVM	PLS-DA	CART	XGB
Thy1T	57.4	54.2	62.0	50.0	75.0	51.4	50.0	52.8	50.0	52.3
Thy2T	86.6	83.3	83.8	79.2	100.0	77.3	74.1	81.9	79.2	82.4
Thy1T + Thy2T	95.8	79.2	93.1	79.2	100.0	59.3	74.1	81.0	79.2	83.3
NoThy1T	100.0	83.3	93.1	78.2	100.0	75.0	73.1	80.1	73.1	59.7
NoThy1T + Thy1T	100.0	91.7	100.0 [a]	65.7	100.0	75.9	73.1	81.0 [a]	61.6	61.6
NoThy1T + Thy2T	100.0	91.7	88.9	79.2	100.0	74.1	83.3	86.1	79.2	79.2
NoThy1T + Thy1T + Thy2T	100.0	87.5	93.1 [b]	79.2	100.0	75.0	83.3	87.0 [b]	79.2	79.2
NoThy2T	54.2	50.0	69.0	50.0	50.0	54.2	50.0	64.8	50.0	50.0
NoThy2T + Thy1T	62.5	54.2	58.3	50.0	58.3	54.6	54.2	48.1	50.0	49.1
NoThy2T + Thy2T	91.7	87.5	86.1	79.2	100.0	77.3	83.3	81.0	79.2	83.3
NoThy2T + Thy1T + Thy2T	95.8	87.5	92.1	79.2	100.0	65.7	78.2	84.3	79.2	83.3
NoThy1T + NoThy2T	100.0	95.8	89.8	78.2	100.0	75.9	77.3	85.2	65.7	63.0
NoThy1T + NoThy2T + Thy1T	100.0	91.7	94.0	65.7	100.0	73.1	80.6	79.2	61.6	60.6
NoThy1T + NoThy2T + Thy2T	100.0	91.7	88.9	79.2	100.0	69.0	83.3	86.1	79.2	79.2
NoThy1T + NoThy2T + Thy1T + Thy2T	100.0	95.8	93.1	79.2	100.0	67.1	83.3	86.1	79.2	79.2

[a] *Model with the highest NER for GDM prediction, using 1T data only.*
[b] *Model with the highest NER for GDM prediction, including 2T data.*

NER: non-error rate. LR: logistic regression. L-SVM: linear support vector machine. PLS-DA: partial least squares discriminant analysis. CART: classification and regression tree. XGB: extreme gradient boosting. Thy: thyroid predictors. NoThy: non-thyroid predictors. 1T: first trimester. 2T: second trimester.

First trimester thyroid variables (Thy1T) solo displayed a poor performance for the prediction of GDM. Their models got an average cross-validation NER (CV-NER) of 51.3%. Furthermore, in most cases the addition of Thy1T to other variables didn't improve the classification quality. However, the model built only on first trimester predictors that had the highest CV-NER comprises Thy1T (PLS-DA, Thy1T + NoThy1T, CV-NER of 81.0%). The model with second trimester predictors that showed the greatest CV-NER also encompasses Thy1T (PLS-DA, NoThy1T + Thy1T + Thy2T, CV-NER of 87.0%). The top 15 most important variables of those models are presented in **Figure 4.3-2**. **Figure 4.3-2A** concerns the one covering first trimester data only, and **Figure 4.3-2B** the one containing second trimester data. Notably, the 10 major predictors of both models include first trimester thyroid markers, which didn't come up as relevant in univariate analysis. This supports the idea that emerged in exploratory analysis: although moderately, first trimester thyroid variables contribute to the differentiation between NGT and GDM subjects when they are in association with other data. This view is consistent with the study of Zhu et al, who used LR and clinical and biochemical maternal parameters to predict GDM. They observed that the addition of first trimester TT3 and TT3/FT4 improved the AUC of their classification model, from 0.710 to 0.726 and 0.724, respectively [16].

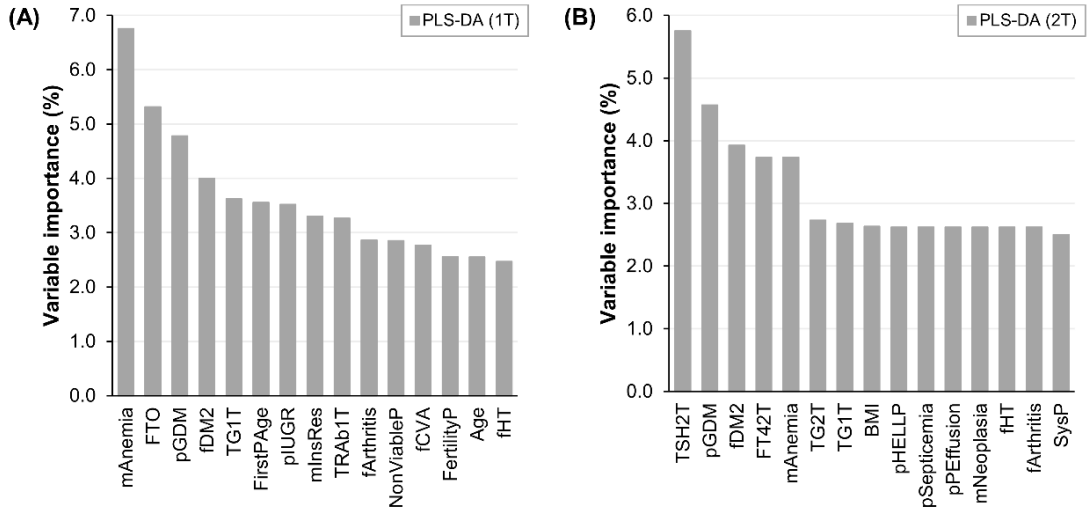


Figure 4.3-2. Variable importance of the machine learning models that predict gestational diabetes mellitus with best performance. (A) Built on first trimester data only. (B) Including second trimester data. *1T: first trimester. 2T: second trimester. PLS-DA: partial least squares discriminant analysis. mAnemia: personal anemia. FTO: fat mass and obesity-associated genotype (rs9939609). pGDM: prior gestational diabetes mellitus. fDM2: family history of type 2 diabetes. TG: thyroglobulin. FirstPAge: first period age. pIUGR: prior intrauterine growth restriction. mInsRes: personal insulin resistance. TRAb: TSH receptor antibody. fArthritis: family history of rheumatoid arthritis. NonViableP: prior non-viable pregnancy. fCVA: family history of cerebrovascular accident. FertilityP: fertility problems. fHT: family history of hypertension. TSH: thyroid stimulating hormone. FT4: free thyroxine. BMI: body mass index. pHELLP: prior hemolysis elevated liver enzymes and low platelets syndrome. pSepticemia: prior septicemia. pPEffusion: prior pleural effusion. mNeoplasia: personal neoplasia. SysP: systolic pressure.*

The particular first trimester thyroid variables that are among the top 10 most relevant predictors of the presented models are TG1T, in the first and the second trimester approaches, and TRAb1T, in the first trimester approach. The levels of TG1T and TRAb1T tend to be higher in the GDM group than in the NGT group (**Table 4.3-S1**), yet they didn't reach statistical significance in univariate analysis. Remarkably, in both models they are more important than age and BMI, two known risk factors for GDM [23]. Few articles have studied the association

between these two thyroid markers and the risk of GDM. Bell et al assessed the role of TG1T on the risk of GDM in Finnish pregnant women, and found no association between them, neither solo nor combined with blood iodine levels [47]. Likewise, Wang et al evaluated the association between TRAb1T and the risk of GDM in Shanghainese pregnant women, and found that positive TRAb1T was related to a decreased risk of this pregnancy disorder [48]. Pregnant women from both Finland [47] and Shanghai [48] are known to be iodine deficient. Differently from those populations, Chilean pregnant women have an adequate iodine intake [49], which could partially explain these contrasting results. It is also important to mention that, in the present study, TG1T and TRAb1T only came out as relevant for the prediction of GDM in association with other maternal variables, which wasn't comprehensively assessed in the mentioned studies. As far as we know, neither TG1T nor TRAb1T had been tested for the prediction of GDM before.

In contrast to Thy1T, second trimester thyroid variables (Thy2T) solo showed an acceptable performance for GDM prediction. Their models reached an average CV-NER of 79.0%. Moreover, the addition of Thy2T to other variables improved the classification performance in most cases. Interestingly, Thy2T presented a higher CV-NER than second trimester non-thyroid variables (NoThy2T, average CV-NER of 53.8%), and the only NoThy2T assessed in this study was fasting blood glucose (FBG2T). This means that Thy2T had a better predictive

performance than FBG2T, a parameter that is often used as a screening tool for GDM [50]. Moreover, several second trimester thyroid parameters figure among the top 15 most important predictors of the second trimester model, whereas FBG2T doesn't. Therefore, Thy2T could be employed as an alternative screening tool for GDM. To our knowledge, this is the first time that thyroid markers are proposed to be used with that purpose.

The particular second trimester thyroid variables that are part of the top 10 most relevant predictors of the second trimester model are TSH2T, FT42T and TG2T. TSH2T and FT42T arose as significant in univariate analysis, but TG2T didn't. They are more important to the model than age and BMI. Furthermore, TSH2T is more relevant than prior history of GDM and family history of DM2, other known risk factors for this pregnancy complication [23]. Higher TSH2T [43,44] and lower FT42T [14,16,45] levels have been associated to an increased risk of GDM, whereas TG2T hadn't been linked with the risk of this disorder until now. The levels of this hormone precursor tend to be higher in the GDM group, than in the control group (**Table 4.3-S1**). FT42T has been used for its prediction in univariate analysis, with poor classification performance [14]. However, as far as we can tell, none of these thyroid markers had been used for the prediction of GDM in multivariate analysis before.

These results prove that maternal thyroid variables are useful for the prediction of GDM, especially at the second trimester of pregnancy.

4.3.3.4. Prediction of post load glycemia with ML techniques

In order to make the prediction of GDM independent from the diagnostic criteria, the prediction of post load glycemia was conducted. Whereas the prediction of GDM is a binary classification task, the prediction of the continuous parameter post load glycemia is a regression task. Given that PLS-DA stood out in classification, its related ML technique PLS was used for regression. The prediction performance was evaluated by RMSE and RE (also known as normalized RMSE or coefficient of variation of the RMSE), which are usually employed to assess the error of regression models in ML-based studies [51]. **Table 4.3-3** presents the RE of each PLS model in calibration and cross-validation. Their RMSEs are shown in **Table 4.3-S4**.

Table 4.3-3. Relative error of machine learning models that predict post load glycemia based on different maternal variables.

Maternal predictors	PLS	
	Calibration RE (%)	Cross-validation RE (%)
Thy1T	22.3	24.7
Thy2T	16.3	20.1
Thy1T + Thy2T	17.7	22.9
NoThy1T	16.3 [a]	21.4 [a]
NoThy1T + Thy1T	15.7	21.6
NoThy1T + Thy2T	15.2	19.1
NoThy1T + Thy1T + Thy2T	14.8	20.0

NoThy2T	21.6	22.4
NoThy2T + Thy1T	21.0	23.7
NoThy2T + Thy2T	17.4	20.1
NoThy2T + Thy1T + Thy2T	17.1	22.2
NoThy1T + NoThy2T	15.5	20.9
NoThy1T + NoThy2T + Thy1T	15.0	21.2
NoThy1T + NoThy2T + Thy2T	14.7 [b]	18.7 [b]
NoThy1T + NoThy2T + Thy1T + Thy2T	14.3	19.6

[a] Model with the lowest RE for post load glycemia prediction, using 1T data only.
[b] Model with the lowest RE for post load glycemia prediction, including 2T data.

PLS: partial least squares. RE: relative error. 1T: first trimester. 2T: second trimester. Thy: thyroid predictors. NoThy: non-thyroid predictors.

Thy1T and Thy2T exhibited a similar behavior than the observed in the classification analysis. Thy1T alone presented the highest error for post load glycemia prediction (RECV of 24.7%) and their addition to other variables didn't improve the regression performance. Thy2T alone displayed a better performance (RECV of 20.1%) and their incorporation to other predictors enhanced the regression quality. Moreover, Thy2T had a lower RECV than FBG2T (NoThy2T, RECV of 22.4%). This is in agreement with the study of Ouzilleau et al, who found a very weak correlation between FBG and OGTT post load glycemia (75 g, 2 h) at 24-28 weeks of pregnancy [52].

In opposition with the observed in the classification analysis, the regression models with the lowest RECV don't contain Thy1T, neither when built on first trimester predictors only (PLS, NoThy1T, RECV of 21.4%) nor when include second trimester predictors (PLS, NoThy1T + NoThy2T + Thy2T, RECV of

18.7%). However, the latter does include Thy2T. The top 15 most relevant variables of these models are presented in **Figure 4.3-3**. **Figure 4.3-3A** concerns the one comprising first trimester data only, and **Figure 4.3-3B** the one involving second trimester data.

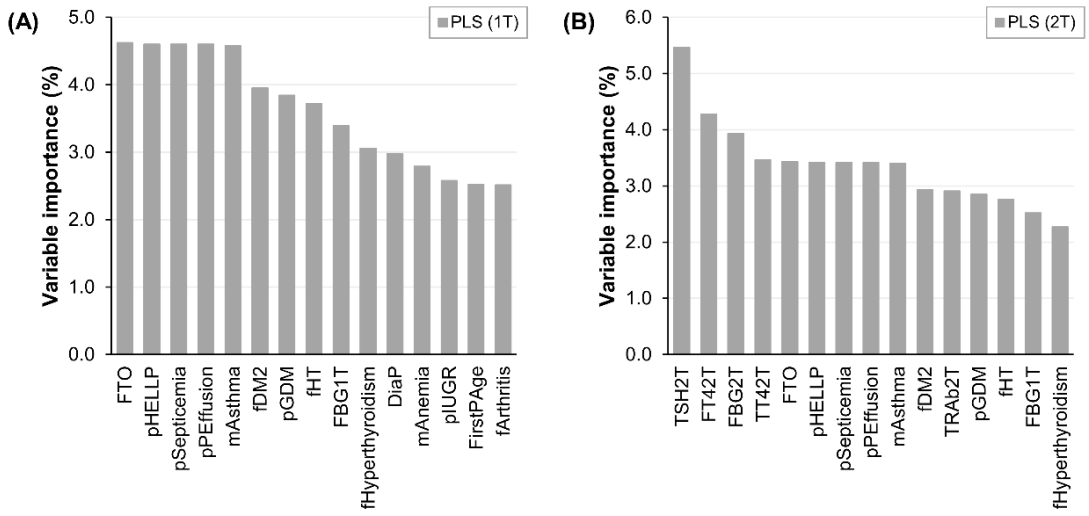


Figure 4.3-3. Variable importance of the machine learning models that predict post load glycemia with best performance. (A) Built on first trimester data only. (B) Including second trimester data. *1T: first trimester. 2T: second trimester. PLS: partial least squares. FTO: fat mass and obesity-associated genotype (rs9939609). pHELLP: prior hemolysis elevated liver enzymes and low platelets syndrome. pSepticemia: prior septicemia. pPEffusion: prior pleural effusion. mAsthma: personal asthma. fDM2: family history of type 2 diabetes. pGDM: prior gestational diabetes mellitus. fHT: family history of hypertension. FBG: fasting blood glucose. fHyperthyroidism: family history of hyperthyroidism. DiaP: diastolic pressure. mAnemia: personal anemia. pIUGR: prior intrauterine growth restriction. FirstPAge: first period age. fArthritis: family history of rheumatoid arthritis. TSH: thyroid stimulating hormone. FT4: free thyroxine. TT4: total thyroxine. TRAb: TSH receptor antibody.*

The particular thyroid variables that are among the top 10 most important predictors of the second trimester model are TSH2T, FT42T and TT42T. In line

with the results described above, TSH2T and FT42T are more relevant to the model than FBG2T. Moreover, the three thyroid markers are more relevant than family history of DM2 and prior GDM, two recognized risk factors for GDM in Chile [23] and worldwide [1]. In contrast to TSH2T and FT42T, TT42T didn't emerge as significant in univariate analysis and hadn't been linked to the risk of GDM before. TT42T levels tend to be lower in the GDM group, than in the NGT group (**Table 4.3-S1**). To our knowledge, this is the first time that these thyroid parameters are used to predict post load glycemia in pregnancy. In agreement with was already suggested, they could be used as an alternative screening tool for GDM, independently of its diagnostic criteria.

These results demonstrate that second trimester maternal thyroid variables are useful for the prediction of OGTT post load glycemia.

4.3.3.5. Models simplification by variable selection

In order to ease the application of the best classification and regression models, variable selection was performed. Prior ML-based studies carried out similar approaches [6,19]. For each full model, different numbers of variables were selected, which resulted in various simplified models. Their performance is presented in **Table 4.3-S5**. The first and second trimester simplified models with the highest CV-NER (in the case of the prediction of GDM) and the lowest RECV

(in the case of the prediction of post load glycemia) were chosen as the final ones. **Table 4.3-4** displays the NER or RE of the full and the final simplified models, in calibration and cross-validation.

Table 4.3-4. Non-error rate and relative error of machine learning models that predict gestational diabetes mellitus and post load glycemia before and after variable selection.

Maternal predictors	Calibration		Cross-validation	
	PLS-DA	PLS	PLS-DA	PLS
	NER (%)	RE (%)	NER (%)	RE (%)
First trimester				
Full [a]	100.0	16.3	81.0	21.4
Simplified [b]	87.0	17.6	87.0	20.7
Second trimester				
Full [c]	93.1	14.7	87.0	18.7
Simplified [d]	89.8	16.1	89.8	18.4

[a] For PLS-DA: NoThy1T + Thy1T. For PLS: NoThy1T.
[b] For PLS-DA: 4 predictors (mAnemia, FTO, pGDM, fDM2). For PLS: 7 predictors (FTO, pHELLP, pSepticemia, pPEffusion, mAsthma, fDM2, pGDM).
[c] For PLS-DA: NoThy1T + Thy1T + Thy2T. For PLS: NoThy1T + NoThy2T + Thy2T.
[d] For PLS-DA: 9 predictors (TSH2T, pGDM, fDM2, FT42T, mAnemia, TG2T, TG1T, BMI, pHELLP). For PLS: 9 predictors (TSH2T, FT42T, FBG2T, TT42T, FTO, pHELLP, pSepticemia, pPEffusion, mAsthma).

PLS-DA: partial least squares discriminant analysis. PLS: partial least squares. NER: non-error rate. RE: relative error. NoThy: non-thyroid predictors. Thy: thyroid predictors. 1T: first trimester. 2T: second trimester. mAnemia: personal anemia. FTO: fat mass and obesity-associated genotype (rs9939609). pGDM: prior gestational diabetes mellitus. fDM2: family history of type 2 diabetes. pHELLP: prior hemolysis elevated liver enzymes and low platelets syndrome. pSepticemia: prior septicemia. pPEffusion: prior pleural effusion. mAsthma: personal asthma. TSH: thyroid stimulating hormone. FT4: free thyroxine. TG: thyroglobulin. BMI: body mass index. FBG: fasting blood glucose. TT4: total thyroxine.

In classification with first trimester data only, the full model achieved a CV-NER of 81.0%. The selection of the top 4 most important variables increased the CV-NER to 87.0%. In the case of classification including second trimester data, the

full model reached a CV-NER of 87.0%. The selection of the top 9 most relevant variables improved the predictive performance, but in a more moderate extent than in the first trimester case (CV-NER of 89.8%). The former simplified model doesn't contain thyroid predictors, but the latter includes both first and second trimester thyroid markers. It is important to notice that the second trimester model has a better predictive performance than the first trimester model, both before and after variable selection. This could be related to the fact that GDM's hyperglycemia state is already established at the second trimester of pregnancy. In contrast, at the first trimester of pregnancy the pathophysiological events that lead to the clinical manifestation of GDM are likely to be incipient. Thus, the phenotypic alterations induced by this pregnancy complication will probably be more evident in the second trimester than in the first trimester of pregnancy, which is reflected in the variables' behavior and the models' performance. This idea is consistent with was described by other authors. For instance, Catalano et al observed an impaired insulin response and a reduced hepatic insulin sensitivity in GDM women at late gestation, but not at early gestation [53]. Likewise, Wong et al found an increased placental volume in GDM subjects at the second trimester, but not at the first trimester of pregnancy [54]. Similarly, Sovio et al noticed an altered fetal biometry in GDM pregnancies at 28 weeks of gestation, but not at 20 weeks of gestation [55]. On the other hand, the main differential predictors between the first and second trimester models are thyroid. Therefore, the comportment of the models' performance is in agreement with what was

already discussed: thyroid markers are useful for the prediction of GDM, especially at the second trimester of pregnancy.

In regression, the full model with first trimester data alone got a RECV of 21.4%. The selection of the 7 most relevant variables decreased the RECV to 20.7%. In the case of regression comprising second trimester data, the full model yielded a RECV of 18.7%. The selection of the 9 most important variables improved the predictive performance, but in a more discreet degree than in the first trimester case (RECV of 18.4%). Even though none of the simplified models include first trimester thyroid predictors, the latter does contain second trimester thyroid markers. Similar to what was observed in classification, the second trimester model has a better predictive performance than the first trimester model, both before and after variable selection. Together, these observations reaffirm the idea that second trimester thyroid variables are useful for the prediction of OGTT post load glycemia.

Figure 4.3-4 presents the cross-validation predictive performance of the final simplified models in a graphical manner. **Figures 4.3-4A** and **4.3-4B** display the receiver operating characteristic (ROC) curve for the prediction of GDM in the first and the second trimester of pregnancy, respectively. **Figures 4.3-4C** and **4.3-4D** show the correlation between the actual and the predicted value of post load

glycemia in the first and second trimester of pregnancy, respectively. Calibration plots are exhibited in **Figure 4.3-S1**.

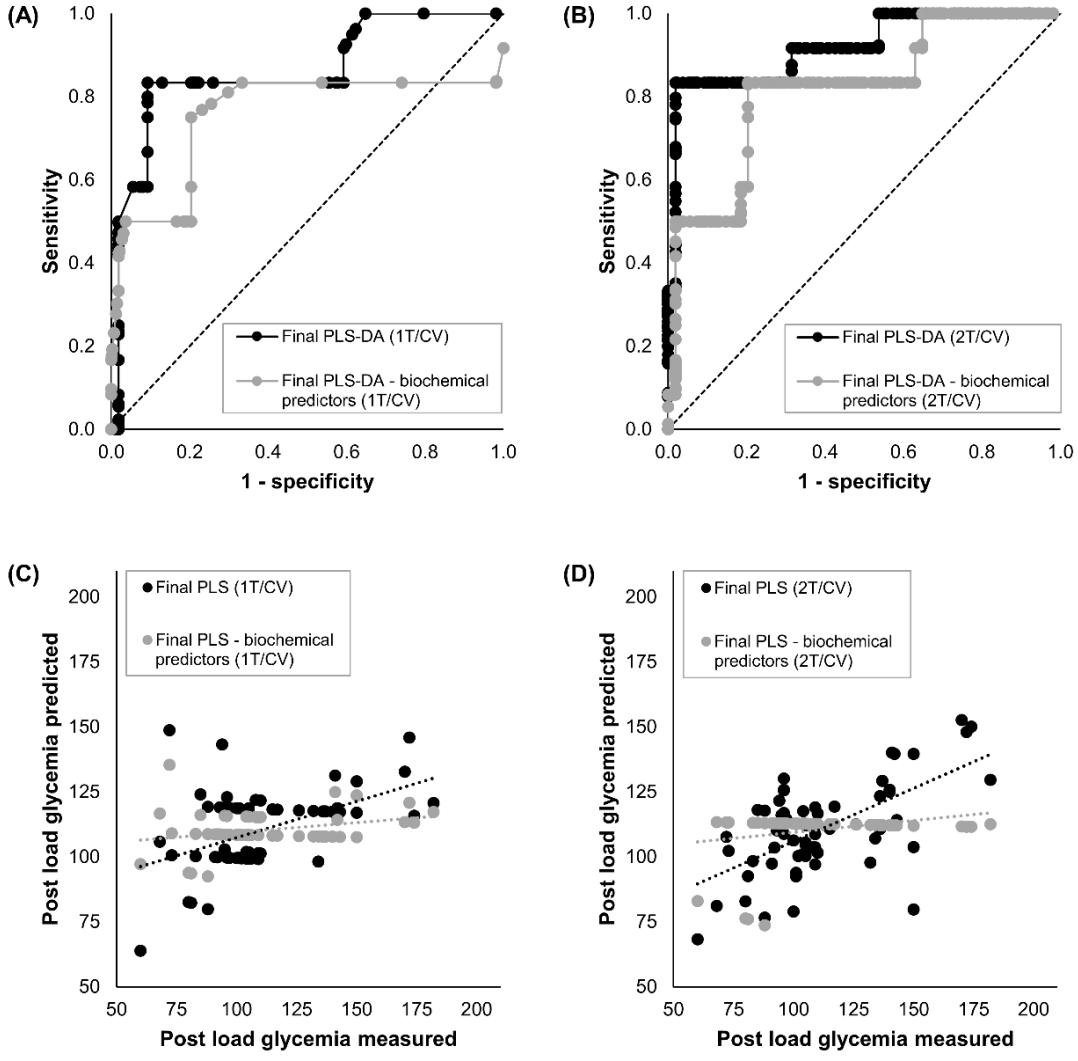


Figure 4.3-4. Cross-validation performance of the final machine learning models that predict gestational diabetes mellitus and post load glycemia. A-B) For the prediction of GDM in the first (A) or second (B) trimester of pregnancy by PLS-DA. C-D) For the prediction of post load glycemia in the first (C) or second (D) trimester of pregnancy by PLS. The four plots show

the cross-validation performance of the final models with their full set of predictors (black) and after the remotion of biochemical markers (grey). *1T: first trimester. 2T: second trimester. PLS-DA: partial least squares discriminant analysis. PLS: partial least squares. CV: cross-validation.*

The prediction of GDM with the first trimester model reaches calibration and cross-validation AUCs of 0.914 (95% CI: 0.821-1.000, p<0.001****) and 0.867 (95% CI: 0.741-0.994, p<0.001****), respectively. This classification model covers the top 4 most important predictors presented in **Figure 4.3-2A**, i.e. personal history of anemia, FTO genotype, prior GDM and family history of DM2. The contribution of each predictor to the final model is shown on Panel A in **Figure 4.3-S2**. The relationship between them and GDM was already discussed. Remarkably, the deletion of FTO genotype from the model decreases calibration and cross-validation AUCs to 0.860 (95% CI: 0.722-0.998, p<0.001****) and 0.759 (95% CI: 0.556-0.962, p<0.01**), respectively.

GDM prediction with the second trimester model achieves AUCs of 0.940 (95% CI: 0.860-1.000, p<0.001****) and 0.920 (95% CI: 0.823-1.000, p<0.001****) in calibration and cross-validation, respectively. This classification model comprises the top 9 most relevant variables displayed in **Figure 4.3-2B**, i.e. TSH2T, prior GDM, family history of DM2, FT42T, personal history of anemia, TG2T, TG1T, BMI and past HELLP syndrome. The contribution of each predictor to the final model is presented on Panel B in **Figure 4.3-S2**. The association between the majority of them and GDM was formerly discussed, except for past hemolysis,

elevated liver enzymes and low platelet (HELLP) syndrome. Past HELLP syndrome tends to be more frequent in GDM than in NGT pregnancies (**Table 4.3-S1**). The relationship between this syndrome and the development of GDM in a future pregnancy has not been studied. However, HELLP has been linked to long-term glucose intolerance [56]. Furthermore, severe hypertensive pregnancy disorders, including the HELLP syndrome, have been associated with an increased risk of DM2 [57]. Interestingly, the remotion of all biochemical variables from the model, i.e. TSH2T, FT42T, TG2T and TG1T, reduces the predictive accuracy in both calibration and cross-validation, with AUCs of 0.914 (95% CI: 0.836-0.992, p<0.001****) and 0.819 (95% CI: 0.682-0.957, p<0.001***), respectively. This result reinforces the idea that thyroid markers are useful to predict GDM at the second trimester of pregnancy.

It is important to note that the classification models presented in this study were calibrated (i.e. developed or trained), optimized and then validated using the Chilean diagnostic criteria. Therefore, they should only be used to predict GDM under that criteria. Very few articles have proposed strategies to predict GDM in Chilean pregnant women. They used clinical parameters [23], plasma biomarkers [58], plasma extracellular vesicles [59], oral extracellular vesicles [60] and the combination of glycemia with BMI or gingival crevicular fluid placental growth factor (GCF-PlGF) [61]; and got AUCs of 0.739, 0.870, 0.798, 0.814, 0.828 and 0.898, respectively. Those studies employed preconception [23] or first trimester

data [58–61], and the only ML technique used was LR. None of them validated their models, hence, the reported AUCs are calibration AUCs. The first and second trimester models presented here have similar or better calibration AUCs. Moreover, they are cross-validated. As far as we know, this is the first study that develops and validates ML models for the prediction of GDM in Chilean subjects. Post load glycemia prediction with the first trimester model yields a Spearman r correlation coefficient of 0.496 (95% CI: 0.282-0.663, p<0.001****) and 0.259 (95% CI: 0.011-0.477, p=0.036*) in calibration and cross-validation, respectively. This regression model encompasses the top 7 most important predictors exhibited in **Figure 4.3-3A**, i.e. FTO genotype, past HELLP syndrome, prior septicemia, prior pleural effusion, past asthma, family history of DM2 and prior GDM. The contribution of each predictor to the final model is presented on Panel A in **Figure 4.3-S3**. The link between most of them and GDM was discussed above, except for prior septicemia, prior pleural effusion and past asthma. Prior septicemia tends to be more frequent in GDM than in NGT pregnancies (**Table 4.3-S1**). Even though its association with GDM has not been investigated, septicemia has been related to a higher risk of DM2 [62,63]. Likewise, prior pleural effusion leans to be more frequent in GDM than in NGT subjects (**Table 4.3-S1**). There are several long-term follow-up studies of patients affected with pleural effusion [64–69], but none of them assessed diabetes or related conditions as outcomes. Hence, more studies are needed to explain the link between pleural effusion and GDM. At last, past asthma tends to be less common in GDM than in NGT pregnant women

(**Table 4.3-S1**). In contrast, a meta-analysis reported that maternal asthma was associated with a higher risk of GDM [70]; however, it didn't evaluate the risk of GDM by ethnicity. A different study revealed that Hispanic subjects with and without asthma had an increased risk of GDM, when compared to White subjects. Interestingly, in the Hispanic cohort, the non-asthmatics had a higher odds ratio (OR) for GDM than the asthmatics [71]. It is interesting to note that, as in the classification first trimester final model, the deletion of FTO genotype from the regression model is associated with an impaired predictive performance, with calibration and cross-validation Spearman r correlation coefficients of 0.375 (95% CI: 0.139-0.571, $p<0.01$**) and -0.081 (95% CI: -0.324-0.171, $p>0.05$ ns), respectively.

The prediction of post load glycemia with the second trimester model gets calibration and cross-validation Spearman r correlation coefficients of 0.520 (95% CI: 0.311-0.681, $p<0.001$****) and 0.457 (95% CI: 0.234-0.634, $p<0.001$***), respectively. This regression model involves the top 9 most relevant variables shown in **Figure 4.3-3B**, i.e. TSH2T, FT42T, FBG2T, TT42T, FTO genotype, past HELLP syndrome, prior septicemia, prior pleural effusion and past asthma. The contribution of each predictor to the final model is exhibited on Panel B in **Figure 4.3-S3**. Their relation with GDM was already discussed. Remarkably, as in the classification second trimester final model, the remotion of all biochemical variables from the regression model, i.e. TSH2T, FT42T, FBG2T, TT42T and FTO

genotype, leads to a reduced predictive performance, with Spearman r correlation coefficients of 0.425 (95% CI: 0.197-0.609, p<0.001***) and -0.679 (95% CI: -0.794--0.518, p<0.001****) in calibration and cross validation, respectively. This result supports the idea that thyroid-related markers are useful to predict post load glycemia at the second trimester of pregnancy.

It is important to note that, contrary to what happens in the case of classification models, the regression models presented in this work are not restricted to a specific diagnostic criteria. This is because they allow to predict the post load glycemia very value, and then that value can be interpreted as NGT or GDM depending on the criteria applied. The error associated with the prediction of post load glycemia (given in terms of RECV) using the best regression models of this study is 20.7% in the first trimester of pregnancy and 18.4% in the second trimester of pregnancy, regardless of the diagnostic criteria employed. Worldwide, few studies have proposed strategies to predict glycemia in the context of pregnancy. Most of them are focused on blood glucose monitoring for the management of GDM [72–74], and not on the prediction of post load glycemia as an early detection or alternative screening tool for this pregnancy disorder. Furthermore, although some of them evaluate the correlation between different variables and fasting glycemia [75–77], postprandial glycemia [78], glucose challenge test (GCT) post load glycemia [79,80] and OGTT post load glycemia [81–83], they don’t test their predictive performance through error metrics like

RMSE and RE. To our knowledge, this is the first study that introduces and properly assesses ML models to predict OGTT post load glycemia so as to make the early detection and screening of GDM independent of the diagnostic criteria.

4.3.3.6. Strengths and limitations of this study

This study has some limitations. The sample size is small, thus, the generalization of the findings described here should be made with moderation. Models built on a restricted number of subjects are prone to overfitting. Even though internal validation was carried out to reduce the risk of it, the possibility that the presented models may be associated with some degree of overfitting cannot be ruled out. Future external validation studies are needed to evaluate that possibility. In addition, some of the clinical variables used were self-reported. Hence, the association between those particular parameters and GDM should be interpreted with caution.

Despite its limitations, this investigation has several strengths. Differently from previous studies, it focuses on the capability of the whole thyroid profile to predict GDM, both separately and in association with other maternal variables. Moreover, this work considered thyroid markers from the first and the second trimester of pregnancy, i.e. it tested their potential for both the early detection and the alternative screening of GDM. In addition, it assessed the power of thyroid and

other maternal parameters on the prediction of OGTT post load glycemia, which hadn't been done before. All the presented models were cross-validated, which reduces the chance of overestimating their predictive performance. To our knowledge, this is the first study that performs such a systematic analysis to evaluate the effectiveness of thyroid variables for the prediction of GDM and OGTT post load glycemia.

4.3.4. Conclusion

In this work, the maternal thyroid profile from the first and the second trimester of pregnancy was methodically assessed for the prediction of GDM and OGTT post load glycemia, both alone and combined with other variables. Our results indicate that thyroid parameters are useful for the prediction of GDM and OGTT post load glycemia, especially at the second trimester of pregnancy. Therefore, these markers could be employed as an alternative screening tool for GDM, independently of the diagnostic criteria used. Further studies are required to confirm these findings in larger cohorts and different populations.

Despite its limitations, this investigation represents the first attempt to evaluate the effectiveness of the whole thyroid profile on the prediction of GDM and OGTT post load glycemia. Our best models achieved a good predictive performance for

both outcomes. Future external validation studies are needed so as to corroborate their potential and bring them closer to clinical practice.

Acknowledgments

The authors would like to thank all the pregnant women who voluntarily participated in this study; and to the health staff of the primary health centers (CESFAM) Victor Manuel Fernández, Santa Sabina and Tucapel, who collaborated in the collection and the initial storage of blood samples, and the retrieval of institution-derived maternal data.

4.3.5. References

1. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. Nat Rev Dis Prim. 2019;5: 47. doi:10.1038/s41572-019-0098-8

2. Metzger B, Lowe L, Dyer A, Trimble E, Chaovarindr U, Coustan D, et al. Hyperglycemia and Adverse Pregnancy Outcomes. N Engl J Med. 2008;358: 1991–2002. doi:10.1056/NEJMoa0707943

3. Lowe WL, Scholtens DM, Lowe LP, Kuang A, Nodzenski M, Talbot O, et al. Association of gestational diabetes with maternal disorders of glucose metabolism and childhood adiposity. JAMA - J Am Med Assoc. 2018;320: 1005–1016. doi:10.1001/jama.2018.11628

4. Lowe WL, Scholtens DM, Kuang A, Linder B, Lawrence JM, Lebenthal Y, et al. Hyperglycemia and adverse Pregnancy Outcome follow-up study (HAPO FUS): Maternal gestational diabetes mellitus and childhood glucose metabolism. Diabetes Care. 2019;42: 372–380. doi:10.2337/dc18-1646

5. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2018;15: 233–234. doi:10.1038/nmeth.4642

6. Artzi NS, Shilo S, Hadar E, Rossman H, Barbash-Hazan S, Ben-Haroush A, et al. Prediction of gestational diabetes based on nationwide electronic health records. Nat Med. 2020;26: 71–76. doi:10.1038/s41591-019-0724-8

7. Farrar Di, Simmonds M, Bryant M, Lawlor DA, Dunne F, Tuffnell D, et al. Risk factor screening to identify women requiring oral glucose tolerance testing to diagnose gestational diabetes: A systematic review and meta-analysis and analysis of two pregnancy cohorts. PLoS One. 2017;12: 1–17. doi:10.1371/journal.pone.0175288

8. Mennickent D, Rodríguez A, Farías-Jofré M, Araya J, Guzmán-Gutiérrez E. Machine learning-based models for gestational diabetes mellitus prediction before 24–28 weeks of pregnancy: A review. Artif Intell Med. 2022;132: 102378. doi:10.1016/j.artmed.2022.102378

9. Yang Y, Li Q, Wang Q, Ma X. Thyroid antibodies and gestational diabetes mellitus: A meta-analysis. Fertil Steril. 2015;104: 665-671.e3. doi:10.1016/j.fertnstert.2015.06.003

10. Jia M, Wu Y, Lin B, Shi Y, Zhang Q, Lin Y, et al. Meta-analysis of the association between maternal subclinical hypothyroidism and gestational diabetes mellitus. Int J Gynecol Obstet. 2019;144: 239–247. doi:10.1002/ijgo.12751

11. Kent NL, Young SL, Akison LK, Cuffe JSM. Is the link between elevated TSH and gestational diabetes mellitus dependant on diagnostic criteria and thyroid antibody status: a systematic review and meta-analysis. Endocrine. 2021;74: 38–49. doi:10.1007/s12020-021-02733-x

12. Luo J, Wang X, Yuan L, Guo L. Association of thyroid disorders with gestational diabetes mellitus: a meta-analysis. Endocrine. 2021;73: 550–560. doi:10.1007/s12020-021-02712-2

13. Sert UY, Buyuk GN, Engin Ustun Y, Ozgu Erdinc AS. Is there any relationship between thyroid function abnormalities, thyroid antibodies and development of gestational diabetes mellitus (GDM) in pregnant women? Medeni Med J. 2020;35: 195–201. doi:10.5222/MMJ.2020.29964

14. Tang L, Li P, Zhou H, Li L. A longitudinal study of thyroid markers during pregnancy and the risk of gestational diabetes mellitus and post-partum

glucose metabolism. Diabetes Metab Res Rev. 2021;37. doi:10.1002/dmrr.3441

15. Yanachkova V, Kamenov Z. The relationship between thyroid dysfunction during pregnancy and gestational diabetes mellitus. Endokrynol Pol. 2021;72: 226–231. doi:10.5603/EP.A2021.0016

16. Zhu B, Han Y, Deng F, Huang K, Yan S, Hao J, et al. The role of triiodothyronine (T3) and t3/free thyroxine (ft4) in glucose metabolism during pregnancy: The ma'anshan birth cohort study. Endocr Connect. 2021;10: 685–693. doi:10.1530/EC-21-0088

17. Qiu H, Yu HY, Wang LY, Yao Q, Wu SN, Yin C, et al. Electronic Health Record Driven Prediction for Gestational Diabetes Mellitus in Early Pregnancy. Sci Rep. 2017;7: 1–13. doi:10.1038/s41598-017-16665-y

18. McLaren R, Haberman S, Moscu M, Atallah F, Friedmann H. A Novel and Precise Profiling Tool to Predict Gestational Diabetes. J Diabetes Sci Technol. 2021;15: 891–896. doi:10.1177/1932296820948883

19. Wu YT, Zhang CJ, Mol BW, Kawai A, Li C, Chen L, et al. Early Prediction of Gestational Diabetes Mellitus in the Chinese Population via Advanced Machine Learning. J Clin Endocrinol Metab. 2021;106: E1191–E1205. doi:10.1210/clinem/dgaa899

20. Grobman WA, Stamilio DM. Methods of clinical prediction. Am J Obstet Gynecol. 2006;194: 888–894. doi:10.1016/j.ajog.2005.09.002

21. Villena Gonzales W, Mobashsher A, Abbosh A. The Progress of Glucose Monitoring—A Review of Invasive to Minimally and Non-Invasive Techniques, Devices and Sensors. Sensors. 2019;19: 800. doi:10.3390/s19040800

22. Araya J, Rodriguez A, Lagos-SanMartin K, Mennickent D, Gutiérrez-Vega S, Ortega-Contreras B, et al. Maternal thyroid profile in first and second trimester of pregnancy is correlated with gestational diabetes mellitus through machine learning. Placenta. 2021;103. doi:10.1016/j.placenta.2020.10.015

23. Garmendia ML, Mondschein S, Montiel B, Kusanovic JP. Trends and predictors of gestational diabetes mellitus in Chile. Int J Gynecol Obstet. 2020;148: 210–218. doi:10.1002/ijgo.13023

24. Tiongco RE, Arceo E, Clemente B, Pineda-Cortel MR. Association of maternal iron deficiency anemia with the risk of gestational diabetes mellitus: a meta-analysis. Arch Gynecol Obstet. 2019;299: 89–95. doi:10.1007/s00404-018-4932-0

25. Behere R V., Deshmukh AS, Otiv S, Gupte MD, Yajnik CS. Maternal Vitamin B12 Status During Pregnancy and Its Association With Outcomes of Pregnancy and Health of the Offspring: A Systematic Review and Implications for Policy in India. Front Endocrinol (Lausanne). 2021;12: 1–18. doi:10.3389/fendo.2021.619176

26. Ríos-Castillo I, Brito A, Olivares M, López-de Romaña D, Pizarro F. Low prevalence of iron deficiency anemia between 1981 and 2010 in chilean women of childbearing age. Salud Publica Mex. 2013;55: 478–483. doi:10.21149/spm.v55i5.7247

27. Atalah E, Amigo H, Bustos P. Does Chile's nutritional situation constitute a double burden? Am J Clin Nutr. 2014;100: 1623S-1627S. doi:10.3945/ajcn.114.083790

28. Lin Z, Wang Y, Zhang B, Jin Z. Association of type 2 diabetes susceptible genes GCKR, SLC30A8, and FTO polymorphisms with gestational diabetes mellitus risk: a meta-analysis. Endocrine. 2018;62: 34–45. doi:10.1007/s12020-018-1651-z

29. He H, Cao W ting, Zeng Y hui, Huang Z qing, Du W rong, Guan N di, et al. Lack of associations between the FTO polymorphisms and gestational diabetes: A meta-analysis and trial sequential analysis. Gene. 2018;677: 169–175. doi:10.1016/j.gene.2018.07.064

30. Guo F, Long W, Zhou W, Zhang B, Liu J, Yu B. FTO, GCKR, CDKAL1 and CDKN2A/B gene polymorphisms and the risk of gestational diabetes mellitus: a meta-analysis. Arch Gynecol Obstet. 2018;298: 705–715. doi:10.1007/s00404-018-4857-7

31. Petermann F, Villagrán M, Troncoso C, Mardones L, Leiva AM, Martínez MA, et al. Asociación entre el polimorfismo rs9939609 del gen FTO y marcadores de adiposidad en población adulta chilena. Rev Med Chil. 2018;146: 717–726. doi:10.4067/s0034-98872018000600717

32. Gutierrez-Vega S, Armella A, Mennickent D, Loyola M, Covarrubias A, Ortega-Contreras B, et al. High levels of maternal total tri-iodothyronine, and low levels of fetal free L-thyroxine and total tri-iodothyronine, are associated with altered deiodinase expression and activity in placenta with

gestational diabetes mellitus. PLoS One. 2020;15: 1–15. doi:10.1371/journal.pone.0242743

33. Moleti M, Trimarchi F, Vermiglio F. Thyroid physiology in pregnancy. Endocr Pract. 2014;20: 589–596. doi:10.4158/EP13341.RA

34. Haddow JE, Lambert-Messerlian G, Eklund E, Neveux LM, Palomaki GE. Peripheral deiodinase activity: A potential explanation for the association between maternal weight and gestational hyperglycemia. Obstet Med. 2018;11: 73–78. doi:10.1177/1753495X17733223

35. Haddow JE, Metzger BE, Lambert-Messerlian G, Eklund E, Coustan D, Catalano P, et al. Maternal BMI, Peripheral Deiodinase Activity, and Plasma Glucose: Relationships between White Women in the HAPO Study. J Clin Endocrinol Metab. 2019;104: 2593–2600. doi:10.1210/jc.2018-02328

36. Wang Y, Lu S, Xu X, Zhang L, Yang J, Hu W. The interactive effects of pre-pregnancy body mass index, thyroid function, and blood lipid levels on the risk of gestational diabetes mellitus: a crossover analysis. BMC Pregnancy Childbirth. 2022;22: 1–10. doi:10.1186/s12884-022-04908-4

37. Zhang S, Wu Y, Pan W, Li G, Zhang D, Li S, et al. Free-Triiodothyronine to Free-Thyroxine Ratio Mediated the Effect of Prepregnancy Body Mass Index or Maternal Weight Gain During Early Pregnancy on Gestational Diabetes Mellitus. Endocr Pract. 2022;28: 398–404. doi:10.1016/j.eprac.2021.12.016

38. Han SM, Han JH, Park JA, Quinn FA, Park J, Oh E. Longitudinal evaluation of thyroid autoimmunity and function in pregnant Korean women. Clin Chem Lab Med. 2013;51: 2295–2301. doi:10.1515/cclm-2013-0598

39. Ekinci EI, Chiu WL, Lu ZX, Sikaris K, Churilov L, Bittar I, et al. A longitudinal study of thyroid autoantibodies in pregnancy: The importance of test timing. Clin Endocrinol (Oxf). 2015;82: 604–610. doi:10.1111/cen.12571

40. Ollero MD, Pineda J, Martínez de Esteban JP, Toni M, Espada M, Anda E. Optimization of the follow-up of pregnant women with autoimmune thyroid disease. Endocrinol Diabetes y Nutr (English ed). 2019;66: 305–311. doi:10.1016/j.endien.2018.08.010

41. Balucan FS, Morshed SA, Davies TF. Thyroid autoantibodies in pregnancy: Their role, regulation and clinical relevance. J Thyroid Res. 2013;2013. doi:10.1155/2013/182472

42. Sharma S, Banerjee S, Krueger PM, Blois SM. Immunobiology of Gestational Diabetes Mellitus in Post-Medawar Era. Front Immunol. 2022;12: 1–10. doi:10.3389/fimmu.2021.758267

43. Ying H, Tang YP, Bao YR, Su XJ, Cai XY, Li YH, et al. Maternal TSH level and TPOAb status in early pregnancy and their relationship to the risk of gestational diabetes mellitus. Endocrine. 2016;54: 742–750. doi:10.1007/s12020-016-1022-6

44. Deng S-Q, Chen H-T, Wang D-Y, Liu B, Chen H-Q, Wang Z-L. Maternal Thyroid-Stimulating Hormone Level and Thyroid Peroxidase Antibody Status in the First and Second Trimester of Pregnancy and Their Relationship with the Risk of Gestational Diabetes Mellitus. Matern Med. 2019;1: 81–85. doi:10.1097/fm9.0000000000000016

45. Haddow JE, Craig WY, Neveux LM, Palomaki GE, Lambert-Messerlian G, Malone FD, et al. Free thyroxine during early pregnancy and risk for gestational diabetes. PLoS One. 2016;11: 1–11. doi:10.1371/journal.pone.0149065

46. Ballabio D, Grisoni F, Todeschini R. Multivariate comparison of classification performance measures. Chemom Intell Lab Syst. 2018;174: 33–44. doi:10.1016/j.chemolab.2017.12.004

47. Bell GA, Männistö T, Liu A, Kannan K, Yeung EH, Kim UJ, et al. The joint role of thyroid function and iodine concentration on gestational diabetes risk in a population-based study. Acta Obstet Gynecol Scand. 2019;98: 500–506. doi:10.1111/aogs.13523

48. Wang J, Gong XH, Peng T, Wu JN. Association of Thyroid Function During Pregnancy With the Risk of Pre-eclampsia and Gestational Diabetes Mellitus. Endocr Pract. 2021;27: 819–825. doi:10.1016/j.eprac.2021.03.014

49. Castillo C, Lustig N, Margozzini P, Gomez A, Rojas MP, Muzzo S, et al. Thyroid-stimulating hormone reference ranges in the first trimester of pregnancy in an iodine-sufficient country. Endocrinol Metab. 2018;33: 466–472. doi:10.3803/EnM.2018.33.4.466

50. Li M, Lan JR, Liang JL, Xiong XL. Diagnostic accuracy of fasting plasma glucose as a screening test for gestational diabetes mellitus: A systematic review and meta-analysis. Eur Rev Med Pharmacol Sci. 2020;24: 11172–11186. doi:10.26355/eurrev_202011_23605

51. Botchkarev A. A new typology design of performance metrics to measure errors in machine learning regression algorithms. Interdiscip J Information, Knowledge, Manag. 2019;14: 45–76. doi:10.28945/4184

52. Ouzilleau C, Roy MA, Leblanc L, Carpentier A, Maheux P. An observational study comparing 2-hour 75-g oral glucose tolerance with fasting plasma glucose in pregnant women: Both poorly predictive of birth weight. Can Med Assoc J. 2003;168: 403–409.

53. Catalano PM, Tyzbir ED, Wolfe RR, Calles J, Roman NM, Amini SB, et al. Carbohydrate metabolism during pregnancy in control subjects and women with gestational diabetes. Am J Physiol - Endocrinol Metab. 1993;264. doi:10.1152/ajpendo.1993.264.1.e60

54. Wong CH, Chen CP, Sun FJ, Chen CY. Comparison of placental three-dimensional power Doppler indices and volume in the first and the second trimesters of pregnancy complicated by gestational diabetes mellitus. J Matern Neonatal Med. 2019;32: 3784–3791. doi:10.1080/14767058.2018.1472226

55. Sovio U, Murphy HR, Smith GCS. Accelerated fetal growth prior to diagnosis of gestational diabetes mellitus: A prospective cohort study of nulliparous women. Diabetes Care. 2016;39: 982–987. doi:10.2337/dc16-0160

56. Weitgasser R, Spitzer D, Kartnig I, Zajc M, Staudach A, Sandhofer F. Association of HELLP syndrome with autoimmune antibodies and glucose intolerance. Diabetes Care. 2000;23: 786–790. doi:10.2337/diacare.23.6.786

57. Lykke JA, Langhoff-Roos J, Sibai BM, Funai EF, Triche EW, Paidas MJ. Hypertensive pregnancy disorders and subsequent cardiovascular morbidity and type 2 diabetes mellitus in the mother. Hypertension. 2009;53: 944–951. doi:10.1161/HYPERTENSIONAHA.109.130765

58. Correa PJ, Venegas P, Palmeiro Y, Albers D, Rice G, Roa J, et al. First trimester prediction of gestational diabetes mellitus using plasma biomarkers: a case-control study. J Perinat Med. 2019;47: 161–168. doi:10.1515/jpm-2018-0120

59. Arias M, Monteiro LJ, Acuña-Gallardo S, Varas-Godoy M, Rice GE, Monckeberg M, et al. Vesículas extracelulares como predictores tempranos de diabetes gestacional. Rev Med Chil. 2019;147: 1503–1509. doi:10.4067/S0034-98872019001201503

60. Monteiro LJ, Varas-Godoy M, Monckeberg M, Realini O, Hernández M, Rice G, et al. Oral extracellular vesicles in early pregnancy can identify patients at risk of developing gestational diabetes mellitus. PLoS One. 2018;14: 1–13. doi:10.1371/journal.pone.0218616

61. Chaparro A, Zúñiga E, Varas-Godoy M, Albers D, Ramírez V, Hernández M, et al. Periodontitis and placental growth factor in oral fluids are early pregnancy predictors of gestational diabetes mellitus. J Periodontol. 2018;89: 1052–1060. doi:10.1002/JPER.17-0497

62. Dombrovskiy VY, Martin AA, Sunderram J, Paz HL. Occurrence and outcomes of sepsis: Influence of race. Crit Care Med. 2007;35: 763–768. doi:10.1097/01.CCM.0000256726.80998.BF

63. Hsu CW, Lin CS, Chen SJ, Lin SH, Lin CL, Kao CH. Risk of type 2 diabetes mellitus in patients with acute critical illness: a population-based cohort study. Intensive Care Med. 2016;42: 38–45. doi:10.1007/s00134-015-4044-2

64. Ferrer JS, Muñoz XG, Orriols RM, Light RW, Morell FB. Evolution of Idiopathic Pleural Effusion: A Prospective, Long-term Follow-up Study. Chest. 1996;109: 1508–1513. doi:10.1378/chest.109.6.1508

65. Janssen JP, Ramlal S, Mravunac M. The long-term follow up of exudative pleural effusion after nondiagnostic thoracoscopy. J Bronchol. 2004;11: 169–174. doi:10.1097/01.lab.0000131024.20254.81

66. Gunluoglu G, Olcmen A, Gunluoglu MZ, Dincer I, Sayar A, Camsari G, et al. Long-term Outcome of Patients With Undiagnosed Pleural Effusion. Arch Bronconeumol. 2015;51: 632–636. doi:10.1016/j.arbr.2015.09.021

67. Chan CHS, Arnold M, Chan CY, Mak TWL, Hoheisel GB. Clinical and pathological features of tuberculous pleural effusion and its long-term consequences. Respiration. 1991;58: 171–175. doi:10.1159/000195921

68. Kataoka H. Ultrasound Pleural Effusion Sign as a Useful Marker for Identifying Heart Failure Worsening in Established Heart Failure Patients During Follow-Up. Congest Hear Fail. 2012;18: 272–277. doi:10.1111/j.1751-7133.2012.00285.x

69. Hung TH, Tseng CW, Tsai CC, Tsai CC, Tseng KC, Hsieh YH. The long-term outcomes of cirrhotic patients with pleural effusion. Saudi J Gastroenterol. 2018;24: 46–51. doi:10.4103/sjg.SJG_336_17

70. Wang G, Murphy VE, Namazy J, Powell H, Schatz M, Chambers C, et al. The risk of maternal and placental complications in pregnant women with asthma: A systematic review and meta-analysis. J Matern Neonatal Med. 2014;27: 934–942. doi:10.3109/14767058.2013.847080

71. Flores KF, Robledo CA, Hwang BS, Leishear K, Laughon Grantz K, Mendola P. Does maternal asthma contribute to racial/ethnic disparities in obstetrical and neonatal complications? Ann Epidemiol. 2015;25: 392-397.e1. doi:10.1016/j.annepidem.2015.01.011

72. Pustozerov E, Popova P, Tkachuk A, Bolotko Y, Yuldashev Z, Grineva E. Development and evaluation of a mobile personalized blood glucose prediction system for patients with gestational diabetes mellitus. JMIR mHealth uHealth. 2018;6. doi:10.2196/mhealth.9236

73. Pustozerov E, Tkachuk A, Vasukova E, Dronova A, Shilova E, Anopova A, et al. The role of glycemic index and glycemic load in the development of real-time postprandial glycemic response prediction models for patients with gestational diabetes. Nutrients. 2020;12. doi:10.3390/nu12020302

74. Seo W, Park SW, Kim N, Jin SM, Park SM. A personalized blood glucose level prediction model with a fine-tuning strategy: A proof-of-concept study. Comput Methods Programs Biomed. 2021;211: 106424. doi:10.1016/j.cmpb.2021.106424

75. Raúl Delgado M, Victoria Novik A, Felipe Cardemil M, Diego Santander A. Utilidad de la medición de fructosamina como indicador de control en pacientes con diabetes gestacional y pregestacional. Rev Med Chil. 2011;139: 1444–1450. doi:10.4067/S0034-98872011001100008

76. Kizilgul M, Kan S, Beysel S, Apaydin M, Ozcelik O, Caliskan M, et al. Is fibroblast growth factor 23 a new cardiovascular risk marker in gestational diabetes? Arch Endocrinol Metab. 2017;61: 562–566. doi:10.1590/2359-3997000000287

77. Ren Z, Zhe D, Li Z, Sun X-P, Yang K, Lin L. Study on the correlation and predictive value of serum pregnancy-associated plasma protein A, triglyceride and serum 25-hydroxyvitamin D levels with gestational diabetes mellitus. World J Clin Cases. 2020;8: 864–873. doi:10.12998/wjcc.v8.i5.864

78. Scioscia M, Kunjara S, Gumaa K, McLean P, Rodeck CH, Rademacher TW. Urinary excretion of inositol phosphoglycan P-type in gestational

diabetes mellitus. Diabet Med. 2007;24: 1300–1304. doi:10.1111/j.1464-5491.2007.02267.x

79. Shushan A, Samueloff A. Correlation between fasting glucose in the first trimester and glucose challenge test in the second. Obstet Gynecol. 1998;91: 596–599. doi:10.1016/S0029-7844(98)00042-8

80. Kianpour M, Saadatmand F, Nematbakhsh M, Fahami F. Relationship between c-reactive protein and screening test results of gestational diabetes in pregnant women referred to health centers in Isfahan in 2013-2014. Iran J Nurs Midwifery Res. 2019;24: 360–364. doi:10.4103/ijnmr.IJNMR_352_14

81. Eid J, Kechichian T, Benavides E, Thibodeaux L, Salazar AE, Saade GR, et al. The Quantose Insulin Resistance Test for Maternal Insulin Resistance: A Pilot Study. Am J Perinatol. 2020;1. doi:10.1055/s-0040-1716730

82. Erbağcı MO, Tuna G, Köse S, Dal-Bekar NE, Akış M, Kant M, et al. Association between early oxidative DNA damage and iron status in women with gestational diabetes mellitus. Reprod Toxicol. 2021;103: 171–180. doi:10.1016/j.reprotox.2021.05.009

83. Wu Y, Bible PW, Long S, Ming W kit, Ding W, Long Y, et al. Metagenomic analysis reveals gestational diabetes mellitus-related microbial regulators of glucose tolerance. Acta Diabetol. 2020;57: 569–581. doi:10.1007/s00592-019-01434-2

4.3.6. Supporting information

Table 4.3-S1. Maternal variables that were recorded in both normal glucose tolerance and gestational diabetes mellitus pregnancies.

Variable	Unit	NGT (n=54)	GDM (n=12)	p value		All (n=66)
First trimester						
Age	years	30 ± 5	30 ± 7	0.724	NS	30 ± 6
Height	m	1.60 ± 0.06	1.60 ± 0.05	0.804	NS	1.60 ± 0.06
Weight	Kg	70.0 (60.4-77.1)	77.0 (66.0-81.8)	0.084	NS	70.5 (61.9-78.2)
BMI	Kg/m^2	27.4 (23.5-31.1)	30.3 (27.1-31.6)	0.046	*	27.8 (24.5-31.2)
Systolic pressure	mmHg	109 ± 11	116 ± 7	0.042	*	110 ± 11
Diastolic pressure	mmHg	67 ± 7	66 ± 6	0.443	NS	67 ± 7
Supplement consumption	%	57.4 (31/54)	41.7 (5/12)	0.355	NS	54.5 (36/66)
Hyperemesis	%	24.1 (13/54)	16.7 (2/12)	0.719	NS	22.7 (15/66)
Vaginal bleeding	%	7.4 (4/54)	16.7 (2/12)	0.298	NS	9.1 (6/66)
Drug use at pregnancy	%					
Cigarette		3.7 (2/54)	0.0 (0/12)	>0.999	NS	3.0 (2/66)
Alcohol		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Other drugs		3.7 (2/54)	0.0 (0/12)	>0.999	NS	3.0 (2/66)
Fasting glycemia	mg/dL	81 ± 8	83 ± 4	0.272	NS	81 ± 7
TSH	µIU/mL	1.41 (1.03-2.36)	1.58 (1.01-2.40)	0.752	NS	1.41 (1.02-2.36)
TT3	ng/mL	1.62 (1.55-1.80)	1.61 (1.54-1.85)	0.789	NS	1.62 (1.55-1.82)
TT4	µg/dL	11.30 (10.55-12.50)	10.90 (9.57-12.80)	0.436	NS	11.30 (10.40-12.50)
FT4	ng/dL	0.99 (0.85-1.11)	0.99 (0.89-1.08)	0.905	NS	0.99 (0.86-1.11)
TG	ng/mL	16.69 (9.70-18.13)	17.63 (9.90-22.17)	0.394	NS	16.69 (9.74-19.14)
aTG	IU/mL	9.71 (5.37-11.99)	12.83 (6.12-16.89)	0.239	NS	10.09 (5.40-13.29)
aTPO	IU/mL	1.32 (0.43-3.51)	1.22 (0.38-1.71)	0.407	NS	1.32 (0.43-2.51)
TRAb	IU/L	0.25 (0.25-0.29)	0.25 (0.25-0.43)	0.338	NS	0.25 (0.25-0.32)
Preconception data						
Drug use before pregnancy	%					
Cigarette		37.0 (20/54)	33.3 (4/12)	>0.999	NS	36.4 (24/66)
Alcohol		57.4 (31/54)	50.0 (6/12)	0.752	NS	56.1 (37/66)

Other drugs		20.4 (11/54)	0.0 (0/12)	0.193	NS	16.7 (11/66)
Prior pregnancy issues	%					
GDM		1.9 (1/54)	33.3 (4/12)	0.003	**	7.6 (5/66)
Hypertension		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Preeclampsia		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
HELLP		0.0 (0/54)	8.3 (1/12)	0.182	NS	1.5 (1/66)
IUGR		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Preterm birth		3.7 (2/54)	8.3 (1/12)	0.458	NS	4.5 (3/66)
Septicemia		0.0 (0/54)	8.3 (1/12)	0.182	NS	1.5 (1/66)
Pleural effusion		0.0 (0/54)	8.3 (1/12)	0.182	NS	1.5 (1/66)
Fetal arrhythmia		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Hernia		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Prior non-viable pregnancy	%	18.5 (10/54)	8.3 (1/12)	0.673	NS	16.7 (11/66)
Fertility problems	%	14.8 (8/54)	0.0 (0/12)	0.333	NS	12.1 (8/66)
PCOS	%	24.1 (13/54)	16.7 (2/12)	0.719	NS	22.7 (15/66)
First period age	years	13 (12-14)	12 (11-13)	0.095	NS	13 (12-14)
Last period month	%			0.699	NS	
January		3.7 (2/54)	0.0 (0/12)			3.0 (2/66)
February		5.6 (3/54)	25.0 (3/12)			9.1 (6/66)
March		9.3 (5/54)	0.0 (0/12)			7.6 (5/66)
April		9.3 (5/54)	8.3 (1/12)			9.1 (6/66)
May		11.1 (6/54)	8.3 (1/12)			10.6 (7/66)
June		13.0 (7/54)	8.3 (1/12)			12.1 (8/66)
July		11.1 (6/54)	16.7 (2/12)			12.1 (8/66)
August		9.3 (5/54)	0.0 (0/12)			7.6 (5/66)
September		3.7 (2/54)	16.7 (2/12)			6.1 (4/66)
October		11.1 (6/54)	8.3 (1/12)			10.6 (7/66)
November		11.1 (6/54)	8.3 (1/12)			10.6 (7/66)
December		1.9 (1/54)	0.0 (0/12)			1.5 (1/66)
Personal morbid history	%					
Insulin resistance		3.7 (2/54)	0.0 (0/12)	>0.999	NS	3.0 (2/66)
Hypertension		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Thyroid disfunction		7.4 (4/54)	0.0 (0/12)	>0.999	NS	6.1 (4/66)
Anemia		0.0 (0/54)	16.7 (2/12)	0.031	*	3.0 (2/66)
Asthma		7.4 (4/54)	0.0 (0/12)	>0.999	NS	6.1 (4/66)
Rhinitis		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)

Osteopenia		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Von Willebrand		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Neoplasia		0.0 (0/54)	8.3 (1/12)	0.182	NS	1.5 (1/66)
Family morbid history	%					
Insulin resistance		3.7 (2/54)	8.3 (1/12)	0.458	NS	4.5 (3/66)
Prediabetes		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
DM1		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
DM2		20.4 (11/54)	66.7 (8/12)	0.003	**	28.8 (19/66)
Hypertension		33.3 (18/54)	66.7 (8/12)	0.049	*	39.4 (26/66)
Thrombosis		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Heart surgery		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Myocardial infarction		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Cerebrovascular accident		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Hypothyroidism		16.7 (9/54)	25.0 (3/12)	0.679	NS	18.2 (12/66)
Hyperthyroidism		5.6 (3/54)	16.7 (2/12)	0.222	NS	7.6 (5/66)
Gout		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Chronic kidney disease		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
Rheumatoid arthritis		0.0 (0/54)	8.3 (1/12)	0.182	NS	1.5 (1/66)
Asthma		9.3 (5/54)	0.0 (0/12)	0.575	NS	7.6 (5/66)
Neoplasia		1.9 (1/54)	0.0 (0/12)	>0.999	NS	1.5 (1/66)
FTO genotype (rs9939609)	%			0.047	*	
TT		53.7 (29/54)	25.0 (3/12)			48.5 (32/66)
TA		44.4 (24/54)	66.7 (8/12)			48.5 (32/66)
AA		1.9 (1/54)	8.3 (1/12)			3.0 (2/66)
Second trimester						
TSH	µIU/mL	1.42 (1.10-2.09)	3.25 (2.28-4.38)	0.001	**	1.54 (1.10-2.53)
TT3	ng/mL	1.92 (1.90-2.04)	2.02 (1.92-2.36)	0.027	*	1.92 (1.90-2.11)
TT4	µg/dL	12.90 (12.59-13.33)	11.15 (10.63-13.30)	0.088	NS	12.90 (11.35-13.33)
FT4	ng/dL	0.79 (0.76-0.81)	0.68 (0.63-0.78)	0.006	**	0.77 (0.76-0.81)
TG	ng/mL	17.37 (11.92-20.91)	21.44 (13.63-23.88)	0.067	NS	18.81 (12.28-20.96)
aTG	IU/mL	4.50 (4.30-7.05)	16.93 (8.42-18.45)	0.020	*	4.89 (4.31-16.85)
aTPO	IU/mL	1.18 (1.09-1.40)	1.22 (1.11-2.05)	0.280	NS	1.18 (1.09-1.41)
TRAb	IU/L	0.25 (0.25-0.52)	0.25 (0.25-0.25)	0.228	NS	0.25 (0.25-0.38)
Fasting glycemia	mg/dL	79 ± 8	86 ± 10	0.013	*	80 ± 9
OGTT glycemia (75 g, 2 h)	mg/dL	102 (94-109)	150 (141-172)	<0.001	****	105 (96-134)

*NGT: normal glucose tolerance. GDM: gestational diabetes mellitus. BMI: body mass index. HELLP: hemolysis elevated liver enzymes and low platelets syndrome. IUGR: intrauterine growth restriction. PCOS: polycystic ovary syndrome. DM1: type 1 diabetes. DM2: type 2 diabetes. FTO: fat mass and obesity-associated gene. TSH: thyroid stimulating hormone. TT3: total triiodothyronine. TT4: total thyroxine. FT4: free T4. TG: thyroglobulin. aTG: TG antibody. aTPO thyroid peroxidase antibody. TRAb: TSH receptor antibody. OGTT: oral glucose tolerance test. Qualitative variables are presented as percentage (proportion); quantitative variables with normal distribution, as mean ± standard deviation; and quantitative variables with non-normal distribution, as median (interquartile range). NS: not significant. * p<0.05. ** p<0.01. **** p<0.0001.*

Table 4.3-S2. Maternal thyroid variables that differed significantly between normal glucose tolerance and gestational diabetes mellitus pregnancies.

(A) Full cohort (n=66)

Variable	Unit	First trimester	Second trimester	p value	
TSH	µIU/mL	1.41 (1.02-2.36)	1.54 (1.10-2.53)	0.798	NS
TT3	ng/mL	1.62 (1.55-1.82)	1.92 (1.90-2.11)	<0.001	****
FT4	ng/dL	0.99 (0.86-1.11)	0.77 (0.76-0.81)	<0.001	****
aTG	IU/mL	10.09 (5.40-13.29)	4.89 (4.31-16.85)	0.030	*

(B) Normal glucose tolerance (n=54)

Variable	Unit	First trimester	Second trimester	p value	
TSH	µIU/mL	1.41 (1.03-2.36)	1.42 (1.10-2.09)	0.458	NS
TT3	ng/mL	1.62 (1.55-1.80)	1.92 (1.90-2.04)	<0.001	****
FT4	ng/dL	0.99 (0.85-1.11)	0.79 (0.76-0.81)	<0.001	****
aTG	IU/mL	9.71 (5.37-11.99)	4.50 (4.30-7.05)	0.010	**

(C) Gestational diabetes mellitus (n=12)

Variable	Unit	First trimester	Second trimester	p value	
TSH	µIU/mL	1.58 (1.01-2.40)	3.25 (2.28-4.38)	0.067	NS
TT3	ng/mL	1.61 (1.54-1.85)	2.02 (1.92-2.36)	0.005	**
FT4	ng/dL	0.99 (0.89-1.08)	0.68 (0.63-0.78)	0.001	***
aTG	IU/mL	12.83 (6.12-16.89)	16.93 (8.42-18.45)	0.966	NS

*TSH: thyroid stimulating hormone. TT3: total triiodothyronine. FT4: free thyroxine. aTG: thyroglobulin antibody. Quantitative variables with non-normal distribution are presented as median (interquartile range). NS: not significant. * p<0.05. ** p<0.01. *** p<0.001. **** p<0.0001.*

Table 4.3-S3. Sensitivity and specificity of machine learning models that predict gestational diabetes mellitus based on different maternal variables.

Parameter	Maternal predictors	Calibration					Cross-validation				
		LR	L-SVM	PLS-DA	CART	XGB	LR	L-SVM	PLS-DA	CART	XGB
Sensitivity (%)	Thy1T	16.7	8.3	33.3	0.0	50.0	8.3	0.0	16.7	0.0	8.3
	Thy2T	75.0	66.7	75.0	58.3	100.0	58.3	50.0	75.0	58.3	66.7
	Thy1T + Thy2T	91.7	58.3	91.7	58.3	100.0	33.3	50.0	75.0	58.3	66.7
	NoThy1T	100.0	66.7	91.7	58.3	100.0	66.7	50.0	75.0	50.0	25.0
	NoThy1T + Thy1T	100.0	83.3	100.0 [a]	33.3	100.0	66.7	50.0	75.0 [a]	25.0	25.0
	NoThy1T + Thy2T	100.0	83.3	83.3	58.3	100.0	66.7	66.7	83.3	58.3	58.3
	NoThy1T + Thy1T + Thy2T	100.0	75.0	91.7 [b]	58.3	100.0	66.7	66.7	83.3 [b]	58.3	58.3
	NoThy2T	8.3	0.0	58.3	0.0	0.0	8.3	0.0	50.0	0.0	0.0
	NoThy2T + Thy1T	25.0	8.3	33.3	0.0	16.7	16.7	8.3	16.7	0.0	0.0
	NoThy2T + Thy2T	83.3	75.0	83.3	58.3	100.0	58.3	66.7	75.0	58.3	66.7
	NoThy2T + Thy1T + Thy2T	91.7	75.0	91.7	58.3	100.0	50.0	58.3	83.3	58.3	66.7
	NoThy1T + NoThy2T	100.0	91.7	83.3	58.3	100.0	66.7	58.3	83.3	33.3	33.3
	NoThy1T + NoThy2T + Thy1T	100.0	83.3	91.7	33.3	100.0	66.7	66.7	75.0	25.0	25.0
	NoThy1T + NoThy2T + Thy2T	100.0	83.3	83.3	58.3	100.0	58.3	66.7	83.3	58.3	58.3
	NoThy1T + NoThy2T + Thy1T + Thy2T	100.0	91.7	91.7	58.3	100.0	58.3	66.7	83.3	58.3	58.3
Specificity (%)	Thy1T	98.1	100.0	90.7	100.0	100.0	94.4	100.0	88.9	100.0	96.3
	Thy2T	98.1	100.0	92.6	100.0	100.0	96.3	98.1	88.9	100.0	98.1
	Thy1T + Thy2T	100.0	100.0	94.4	100.0	100.0	85.2	98.1	87.0	100.0	100.0
	NoThy1T	100.0	100.0	94.4	98.1	100.0	83.3	96.3	85.2	96.3	94.4
	NoThy1T + Thy1T	100.0	100.0	100.0 [a]	98.1	100.0	85.2	96.3	87.0 [a]	98.1	98.1
	NoThy1T + Thy2T	100.0	100.0	94.4	100.0	100.0	81.5	100.0	88.9	100.0	100.0
	NoThy1T + Thy1T + Thy2T	100.0	100.0	94.4 [b]	100.0	100.0	83.3	100.0	90.7 [b]	100.0	100.0
	NoThy2T	100.0	100.0	79.6	100.0	100.0	100.0	100.0	79.6	100.0	100.0
	NoThy2T + Thy1T	100.0	100.0	83.3	100.0	100.0	92.6	100.0	79.6	100.0	98.1
	NoThy2T + Thy2T	100.0	100.0	88.9	100.0	100.0	96.3	100.0	87.0	100.0	100.0
	NoThy2T + Thy1T + Thy2T	100.0	100.0	92.6	100.0	100.0	81.5	98.1	85.2	100.0	100.0
	NoThy1T + NoThy2T	100.0	100.0	96.3	98.1	100.0	85.2	96.3	87.0	98.1	92.6
	NoThy1T + NoThy2T + Thy1T	100.0	100.0	96.3	98.1	100.0	79.6	94.4	83.3	98.1	96.3
	NoThy1T + NoThy2T + Thy2T	100.0	100.0	94.4	100.0	100.0	79.6	100.0	88.9	100.0	100.0
	NoThy1T + NoThy2T + Thy1T + Thy2T	100.0	100.0	94.4	100.0	100.0	75.9	100.0	88.9	100.0	100.0

[a] Model with the highest cross-validation non-error rate (CV-NER) for GDM prediction, using 1T data only. [b] Model with the highest CV-NER for GDM prediction, including 2T data. LR: logistic regression. L-SVM: linear support vector machine. PLS-DA: partial least squares discriminant analysis. CART: classification and regression tree. XGB: extreme gradient boosting. 1T: first trimester. 2T: second trimester. Thy: thyroid predictors. NoThy: non-thyroid predictors.

Table 4.3-S4. Root mean square error of machine learning models that predict post load glycemia based on different maternal variables.

Maternal predictors	PLS	
	Calibration RMSE	Cross-validation RMSE
Thy1T	24.9	27.5
Thy2T	18.1	22.4
Thy1T + Thy2T	19.8	25.6
NoThy1T	18.1 [a]	23.8 [a]
NoThy1T + Thy1T	17.5	24.1
NoThy1T + Thy2T	17.0	21.3
NoThy1T + Thy1T + Thy2T	16.5	22.3
NoThy2T	24.1	25.0
NoThy2T + Thy1T	23.5	26.4
NoThy2T + Thy2T	19.4	22.4
NoThy2T + Thy1T + Thy2T	19.0	24.7
NoThy1T + NoThy2T	17.3	23.3
NoThy1T + NoThy2T + Thy1T	16.7	23.6
NoThy1T + NoThy2T + Thy2T	16.4 [b]	20.9 [b]
NoThy1T + NoThy2T + Thy1T + Thy2T	16.0	21.9

[a] Model with the lowers RE for GDM prediction, using 1T data only. [b] Model with the lowest RE for GDM prediction, including 2T data. PLS: partial least squares. RMSEC: root mean square error of calibration. RMSECV: root mean square error of cross-validation. 1T: first trimester. 2T: second trimester. Thy: thyroid predictors. NoThy: non-thyroid predictors.

Table 4.3-S5. Sensitivity, specificity, non-error rate, root mean square error and relative error of machine learning models that predict gestational diabetes mellitus and post load glycemia before and after variable selection.

Maternal predictors	Calibration					Cross-validation				
	PLS-DA			PLS		PLS-DA			PLS	
	Se (%)	Sp (%)	NER (%)	RMSE	RE (%)	Se (%)	Sp (%)	NER (%)	RMSE	RE (%)
First trimester										
Full [a]	100.0	100.0	100.0	18.1	16.3	75.0	87.0	81.0	23.8	21.4
Top 10	75.0	94.4	84.7	21.1	19.0	75.0	88.9	81.9	23.2	20.8
Top 9	83.3	94.4	88.9	21.0	18.8	75.0	88.9	81.9	23.3	20.9
Top 8	83.3	92.6	88.0	19.4	17.4	75.0	87.0	81.0	23.2	20.8
Top 7	83.3	90.7	87.0	19.6 [b]	17.6 [b]	75.0	87.0	81.0	23.1 [b]	20.7 [b]
Top 6	83.3	94.4	88.9	20.9	18.7	75.0	90.7	82.9	23.2	20.8
Top 5	83.3	90.7	87.0	21.4	19.2	66.7	90.7	78.7	23.6	21.2
Top 4	83.3 [c]	90.7 [c]	87.0 [c]	23.3	20.9	83.3 [c]	90.7 [c]	87.0 [c]	25.5	22.9
Top 3	58.3	96.3	77.3	23.3	20.9	50.0	96.3	73.1	25.5	22.9
Top 2	25.0	98.1	61.6	23.3	20.9	25.0	98.1	61.6	25.4	22.8
Top 1	16.7	100.0	58.3	24.6	22.1	16.7	100.0	58.3	25.5	22.9
Second trimester										
Full [d]	91.7	94.4	93.1	16.4	14.7	83.3	90.7	87.0	20.9	18.7
Top 10	83.3	98.1	90.7	17.9	16.1	83.3	96.3	89.8	20.7	18.6
Top 9	83.3 [e]	96.3 [e]	89.8 [e]	17.9 [f]	16.1 [f]	83.3 [e]	96.3 [e]	89.8 [e]	20.5 [f]	18.4 [f]
Top 8	83.3	94.4	88.9	18.9	17.0	83.3	94.4	88.9	21.6	19.4
Top 7	83.3	96.3	89.8	19.6	17.6	83.3	94.4	88.9	21.5	19.3
Top 6	83.3	96.3	89.8	19.5	17.5	83.3	94.4	88.9	21.5	19.3
Top 5	83.3	96.3	89.8	20.2	18.1	75.0	96.3	85.6	21.6	19.3
Top 4	58.3	94.4	76.4	20.7	18.5	58.3	94.4	76.4	22.0	19.7
Top 3	66.7	94.4	80.6	20.5	18.4	66.7	94.4	80.6	21.8	19.6
Top 2	66.7	98.1	82.4	21.6	19.4	66.7	98.1	82.4	22.8	20.4
Top 1	75.0	79.6	77.3	22.1	19.8	66.7	79.6	73.1	22.8	20.4

[a] For PLS-DA: NoThy1T + Thy1T. For PLS: NoThy1T. [b] Model with the lowest RE for post load glycemia prediction, using 1T data only. Predictors: FTO, pHELLP, pSepticemia, pPEffusion, mAsthma, fDM2, pGDM. [c] Model with the highest NER for GDM prediction, using 1T data only. Predictors: mAnemia, FTO, pGDM, fDM2. [d] For PLS-DA: NoThy1T + Thy1T + Thy2T. For PLS: NoThy1T + NoThy2T + Thy2T.

[e] Model with the highest NER for GDM prediction, including 2T data. Predictors: TSH2T, pGDM, fDM2, FT42T, mAnemia, TG2T, TG1T, BMI, pHELLP. [f] Model with the lowest RE for post load glycemia prediction, including 2T data. Predictors: TSH2T, FT42T, FBG2T, TT42T, FTO, pHELLP, pSepticemia, pPEffusion, mAsthma. PLS-DA: partial least squares discriminant analysis. PLS: partial least squares. Se: sensitivity. Sp: specificity. NER: non-error rate. RMSE: root mean square error. RE: relative error. NoThy: non-thyroid predictors. Thy: thyroid predictors. 1T: first trimester. 2T: second trimester. FTO: fat mass and obesity-associated genotype (rs9939609). pHELLP: prior hemolysis elevated liver enzymes and low platelets syndrome. pSepticemia: prior septicemia. pPEffusion: prior pleural effusion. mAsthma: personal asthma. fDM2: family history of type 2 diabetes. pGDM: prior gestational diabetes mellitus. mAnemia: personal anemia. TSH: thyroid stimulating hormone. FT4: free thyroxine. TG: thyroglobulin. BMI: body mass index. FBG: fasting blood glucose. TT4: total thyroxine.

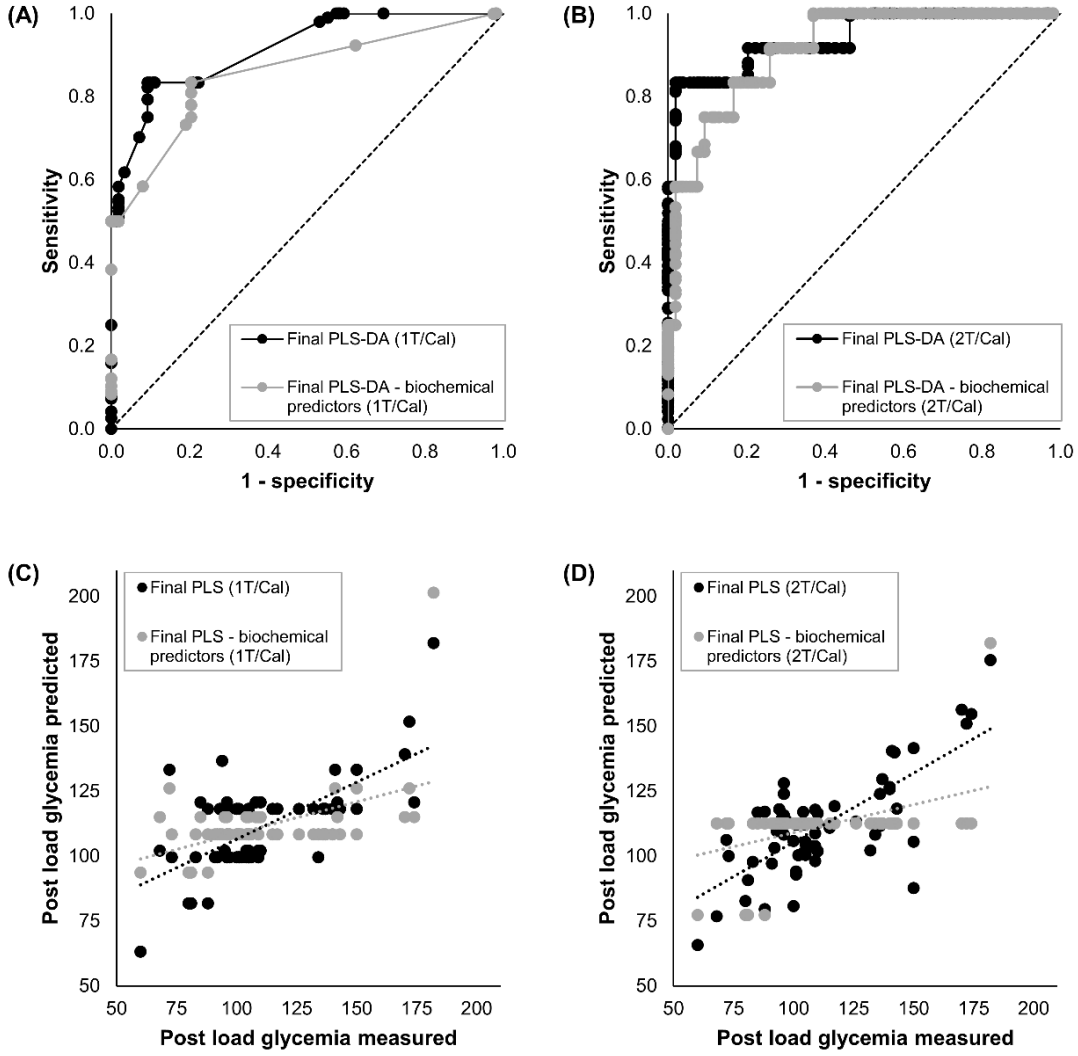


Figure 4.3-S1. Calibration performance of the final machine learning models that predict gestational diabetes mellitus and post load glycemia. A-B) For the prediction of GDM in the first (A) or second (B) trimester of pregnancy by PLS-DA. C-D) For the prediction of post load glycemia in the first (C) or second (D) trimester of pregnancy by PLS. The four plots show the calibration performance of the final models with their full set of predictors (black) and after the remotion of biochemical markers (grey). *1T: first trimester. 2T: second trimester. PLS-DA: partial least squares discriminant analysis. PLS: partial least squares. CV: cross-validation.*

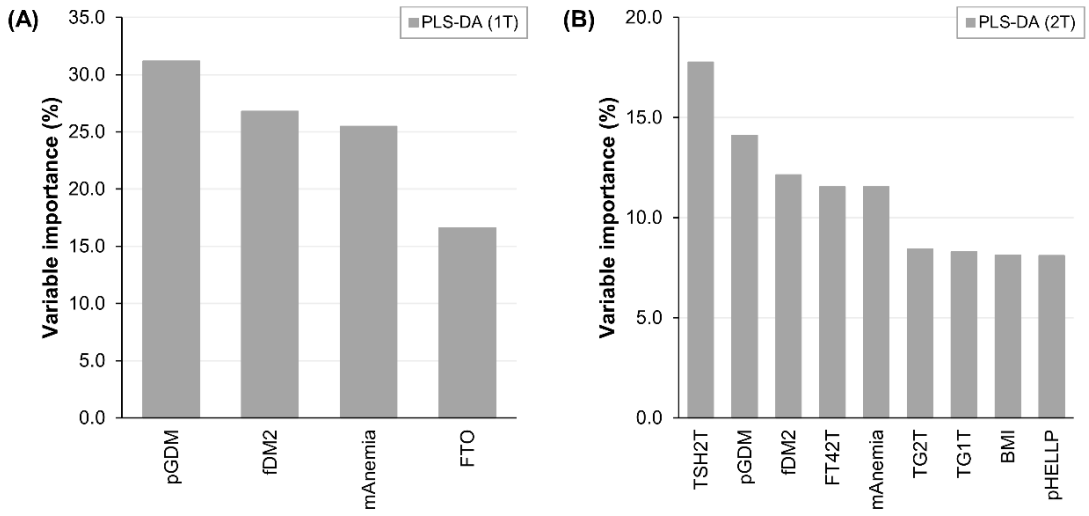


Figure 4.3-S2. Variable importance of the machine learning models that predict gestational diabetes mellitus with best performance after variable selection. (A) Built on first trimester data only. (B) Including second trimester data. *1T: first trimester. 2T: second trimester. PLS-DA: partial least squares discriminant analysis. pGDM: prior gestational diabetes mellitus. fDM2: family history of type 2 diabetes. mAnemia: personal anemia. FTO: fat mass and obesity-associated genotype (rs9939609). TSH: thyroid stimulating hormone. FT4: free thyroxine. TG: thyroglobulin. BMI: body mass index. pHELLP: prior hemolysis elevated liver enzymes and low platelets syndrome.*

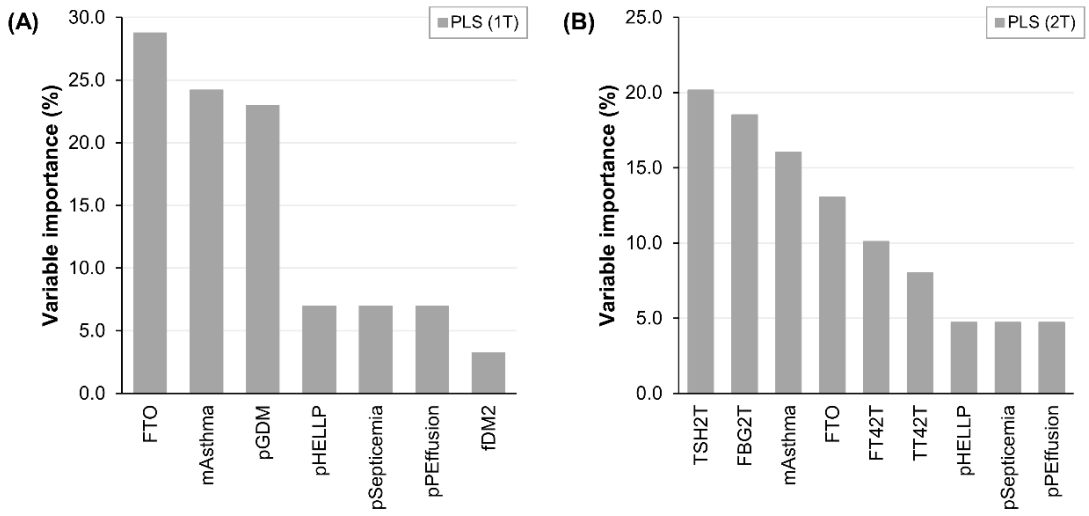


Figure 4.3-S3. Variable importance of the machine learning models that predict post load glycemia with best performance after variable selection. (A) Built on first trimester data only. (B) Including second trimester data. *1T: first trimester. 2T: second trimester. PLS: partial least squares. FTO: fat mass and obesity-associated genotype (rs9939609). mAsthma: personal asthma. pGDM: prior gestational diabetes mellitus. pHELLP: prior hemolysis elevated liver enzymes and low platelets syndrome. pSepticemia: prior septicemia. pPEffusion: prior pleural effusion. fDM2: family history of type 2 diabetes. TSH: thyroid stimulating hormone. FBG: fasting blood glucose. FT4: free thyroxine. TT4: total thyroxine.*

4.4. Near-infrared spectroscopy-based machine learning models as a simple and fast tool for the prediction of gestational diabetes mellitus at different stages of pregnancy

En esta sección se presentan y discuten los resultados que dieron origen al artículo “Near-infrared spectroscopy-based machine learning models as a simple and fast tool for the prediction of gestational diabetes mellitus at different stages of pregnancy”, en preparación para ser enviado a la revista *Analytical and Bioanalytical Chemistry*.

Near-infrared spectroscopy-based machine learning models as a simple and fast tool for the prediction of gestational diabetes mellitus at different stages of pregnancy

Daniela Mennickent [1,2,3], Federico Marini [4], Lucas Romero-Albornoz [2], Enrique Guzmán-Gutiérrez [1,3,*], Juan Araya [2,3,*]

[1] Departamento de Bioquímica Clínica e Inmunología, Facultad de Farmacia, Universidad de Concepción, Concepción, Chile. [2] Departamento de Análisis Instrumental, Facultad de Farmacia, Universidad de Concepción, Concepción, Chile. [3] Machine Learning Applied in Biomedicine (MLAB), Chile. [4] Dipartimento di Chimica, Facoltà di Scienze Matematiche, Fisiche e Naturali, Sapienza Università di Roma, Roma, Italia

* Corresponding authors: jarayaq@udec.cl (JA); eguzman@udec.cl (EG-G)

Abstract

Introduction: Gestational diabetes mellitus (GDM) is a hyperglycemia state that is typically diagnosed by an oral glucose tolerance test (OGTT). The OGTT is unpleasant, time-consuming, poorly reproducible, and tardy. Therefore, the diagnosis of GDM can be improved with early detection tools, or alternative screening tools to the OGTT. Many of the machine learning (ML) predictive models that have been previously proposed for this purpose are based on very time-consuming instrumental methods. Near-infrared (NIR) spectroscopy, a simple, fast and low-cost analytical technique that has been widely applied in clinical diagnostics, has never been assessed for the prediction of GDM. **Aim:** To develop ML predictive models for GDM based on NIR spectroscopy, and to evaluate their potential as early detection or alternative screening tools according to their predictive power and time of analysis. **Methods:** Serum samples from the first (n=67 for control, n=15 for GDM; before GDM diagnosis) and the second (n=39 for control, n=8 for GDM; at the time of GDM diagnosis) trimester of pregnancy were analyzed by NIR spectroscopy. Four spectral ranges were considered (R1: 10500-7600 cm^{-1}, R2: 7600-5100 cm^{-1}, R3: 5100-4000 cm^{-1}, and Full: 10500-4000 cm^{-1}) and 80 combinations of pretreatments were tested for each. NIR data-based ML models were built with single- and multi-block classification ML techniques. Every model was subjected to double cross-validation (DCV). **Results:** In the case of first trimester sera, the best model was

obtained with R1, normalization, mean centering and single-block analysis; and got an area under the receiver operating characteristic curve (AUC-ROC) of 0.5768 ± 0.0635 in DCV. In the case of second trimester sera, the best model was obtained with R3, first derivative (width=15), mean centering and single-block analysis; and achieved an AUC-ROC of 0.8836 ± 0.0259 in DCV. **Discussion:** The developed methods take 32 minutes to predict GDM, including sample preparation and spectral acquisition. This is much faster than the instrumental strategies proposed in literature for GDM prediction, which usually take hours to get a result. Although simple and fast, the best first trimester model showed a moderate predictive performance. Modifying sample preparation, or modelling with high-level data fusion multi-block ML techniques could help improve the results, and bring this method closer to clinical practice for GDM early detection. In contrast, the best second trimester model displayed a very high predictive power. This, coupled with its simplicity and rapidity, makes it an ideal alternative screening method for GDM. **Conclusion:** This is the first study reporting NIR spectroscopy-based methods for the prediction of GDM. The developed methods are simple and fast, and have a great potential to be applied in clinical practice, especially as alternative screening tools to the OGTT at the time of GDM diagnosis. Future external validation studies are needed to confirm the predictive performance of these methods in different populations.

4.4.1. Introduction

Gestational diabetes mellitus (GDM) is a hyperglycemia state that is first diagnosed during pregnancy [1], with negative short- and long-term consequences on both maternal and fetal health [2]. The diagnosis of this disease is typically made by an oral glucose tolerance test (OGTT) in the second or third trimester of pregnancy [3]. The OGTT is unpleasant [4,5], time-consuming [6,7], and has low reproducibility [8,9]. Moreover, by the time of its use, the fetal phenotype is already altered in GDM pregnancies [10–12]. Therefore, the diagnosis of GDM can be improved. Machine learning (ML) predictive modeling is a powerful means of meeting that goal [13], either by early detection tools, or alternative screening tools to the OGTT.

Numerous models have been proposed to predict GDM at different stages of pregnancy. Many of them derive from data acquired by instrumental techniques, such as liquid or gas chromatography coupled to mass spectrometry (LC-MS or GC-MS, respectively), nuclear magnetic resonance (NMR) spectroscopy, polymerase chain reaction (PCR), among others [14]. Methods based on these techniques are very time-consuming, as they require tedious sample preparation procedures or prolonged instrumental runs. Consequently, simpler and faster strategies should be developed.

Near-infrared (NIR) spectroscopy is an analytical technique that is based on the absorption, emission, scattering, reflection, or diffuse-reflection of light in the NIR range of the electromagnetic spectrum, i.e. between 12500 and 4000 cm^{-1} [15]. Biomolecules are capable of interacting with NIR radiation, therefore, a NIR spectrum constitutes the biochemical fingerprint of a biological sample [16]. NIR spectroscopy has multiple advantages, some of which are typical of vibrational spectroscopy, e.g. is noninvasive, nondestructive, reagent-free, waste-free, simple, fast, low-cost, and requires minimal sample preparation [17]. Moreover, NIR spectroscopy is more versatile and less expensive than other vibrational spectroscopy techniques [15]. Due to its advantageous analytical features, this technique has been widely applied in different fields of science, including clinical diagnostics [18].

NIR spectroscopy has never been used as a diagnostic support tool for GDM. Therefore, its capability for GDM prediction at particular stages of pregnancy, such as before or at the time of GDM diagnosis, remains unexplored. This study aims to develop ML predictive models for GDM based on NIR spectroscopy, and to evaluate their potential as early detection or alternative screening tools according to their predictive power and time of analysis.

4.4.2. Materials and methods

4.4.2.1. Ethical aspects

This work was approved by the Ethics Committee of Servicio de Salud Concepción (17-12-88) and was carried out in accordance with the Declaration of Helsinki.

4.4.2.2. Subjects recruitment

First trimester pregnant women were recruited at three primary health centers in Concepción, Chile: CESFAM Victor Manuel Fernández, CESFAM Santa Sabina, and CESFAM Tucapel. Recruitment was conducted between 2017 and 2019. Individuals with pregestational diabetes or any pregnancy alterations different than GDM were excluded. Subjects who gave their written informed consent were included in the study and followed until the second trimester of pregnancy.

4.4.2.3. Medical data collection

28 medical variables were retrieved from CESFAM records and subject statements. The former comprised age and body mass index (BMI) in the first trimester of gestation. The latter encompassed first trimester information, i.e.

supplement consumption, hyperemesis and vaginal bleeding; and preconception information, i.e. drug use, prior pregnancy diseases or complications, prior pregnancy non-viability, fertility issues, history of polycystic ovary syndrome (PCOS), age at menarche, the month of last period, personal morbid history, and family morbid history.

4.4.2.4. Blood samples collection

Blood samples were collected in the first and the second trimester of pregnancy, after fasting (12 hours) or after a 75 g glucose load (2 hours). First and second trimester samples were taken before and at the time of GDM diagnosis, respectively. They were transported to laboratory at 4°C. Sera and NaF/citrate plasma were obtained by centrifugation (10000 g, 5 minutes, 4°C). They were aliquoted and stored at -80°C.

4.4.2.5. NIR spectra acquisition

Sera were randomized before analyses. Each sample was thawed at room temperature, homogenized, and 10 µl were deposited and dried (37°C, 30 minutes) on a MirrIR slide (Kevley Technologies, USA). NIR spectra (range 10500-4000 cm^{-1}, resolution 4 cm^{-1}) were acquired in transflectance mode using a FT-IR Spectrum Frontier/Spotlight 400 Microscopy System (Perkin Elmer,

USA). The acquisition time was 2 minutes per spectrum. Five NIR spectra, i.e. instrumental replicates, were recorded and averaged per sample.

4.4.2.6. GDM diagnosis, cohorts and study groups

In the second trimester of pregnancy, pregnant women were subjected to an OGTT. Fasting and post load plasma glucose were quantified by the hexokinase method [19]. The Chilean diagnostic criteria was used, i.e. subjects with fasting glycemia between 100 and 125 mg/dL, or post load glycemia higher than 140 mg/dL (75 g, 2 hours) were diagnosed with GDM [20]. In this study, two cohorts were considered. The first cohort had first trimester serum samples and, then, first trimester NIR spectra. This cohort, from now called the first trimester cohort, consisted of 82 pregnant women: 15 with GDM and 67 with normal glucose tolerance (NGT). The second cohort had second trimester serum samples and, then, second trimester NIR spectra. This cohort, from now called the second trimester cohort, consisted of 47 subjects: 8 with GDM and 39 with NGT.

4.4.2.7. Classical statistics analyses

Qualitative medical data were compared by two-sided Fisher exact test. The normality of quantitative medical data was evaluated by Shapiro-Wilk test. Normally distributed parameters were compared by unpaired Student t test. Not-

normally distributed parameters were compared by Mann-Whitney test. P values less than 0.05 were considered statistically significant. These analyses were carried out using GraphPad Prism 9.5.1 (GraphPad Software Inc, USA).

4.4.2.8. ML analyses

4.4.2.8.1. Data pretreatment

Prior to ML analyses, qualitative medical parameters were transformed into categorical, and NIR spectra into absorbance. Besides the full NIR spectral range, three shorter wavenumber regions were analyzed: 10500-7600 cm^{-1}, 7600-5100 cm^{-1} and 5100-4000 cm^{-1}. For each spectral range, 80 different combinations of transformations were tested, including Savitzky-Golay smoothing or first/second derivative with varying filter width, standard normal variate scattering correction, weighted least squares baseline correction, and 2-norm normalization. The order in which these transformations were applied was based on recent specialized literature [21,22]. Medical and NIR data were preprocessed by autoscaling and mean centering, respectively.

4.4.2.8.2. Single- and multi-block analyses

For single-block analyses, pretreated data were analyzed by partial least squares linear discriminant analysis (PLS-LDA). For multi-block analyses, pretreated data were analyzed by sequentially orthogonalized PLS-LDA (SO-PLS-LDA). The mathematical description of these classification ML techniques can be found elsewhere [23]. For multi-block analyses, different block orders were tested. Every model was subjected to double-cross validation (DCV), an intensive and robust internal validation strategy consisting of two nested cross-validation loops. The inner loop is used for model training and optimization, and the outer loop for model validation [24]. For DCV, the following parameters were used: 10 segments for the inner loop, 20 segments for the outer loop, and 50 repetitions. Models were developed using in-house written functions in MATLAB R2021a (The MathWorks Inc, USA).

4.4.2.8.3. Evaluation of predictive performance

Models' predictive performance was evaluated by means of their sensitivity (Se), specificity (Sp) and non-error rate (NER). For the best models, the area under the receiver operating characteristic curve (AUC-ROC) was also determined. The mathematical definition of these parameters can be found elsewhere [25,26]. In general terms, the sensitivity and the specificity denote the ability to correctly

classify GDM and NGT subjects, respectively; the NER reflects the ability to correctly classify both GDM and NGT subjects; and the AUC-ROC represents the overall predictive performance of the model in a graphical manner. Every value is presented as the average ± the standard deviation of 50 repetitions in DCV.

4.4.2.8.4. Variable importance and selection

For each model, variable importance in projection (VIP) scores were obtained. Variables with average VIP scores larger than 1 were considered as relevant for model performance [24]. This information was used for variable selection in the multi-block models, and for biochemical interpretation in the final models.

4.4.3. Results

4.4.3.1. First trimester cohort

4.4.3.1.1. Description of the first trimester cohort

To characterize this cohort, classical statistical techniques were used. **Table 4.4-1** displays 28 medical variables, and compares their behavior in NGT and GDM pregnancies. In this cohort, the prevalence of GDM was 18.3%. Only two parameters are statistically different between the two groups: history of GDM in a

prior pregnancy, and family history of diabetes mellitus (DM). Both are more frequent in the GDM group, than in the NGT group.

Table 4.4-1. Medical variables in the first trimester cohort.

Variable	Unit	NGT (n=67)	GDM (n=15)	p value	All (n=82)
Age	years	30 ± 5	32 ± 7	0.394 NS	31 ± 6
BMI	Kg/m^2	27.6 (23.3-31.2)	29.7 (26.6-31.6)	0.051 NS	28.0 (24.1-31.5)
Supplement consumption	%	64.2 (43/67)	53.3 (8/15)	0.557 NS	62.2 (51/82)
Hyperemesis	%	26.9 (18/67)	26.7 (4/15)	>0.999 NS	26.8 (22/82)
Vaginal bleeding	%	9.0 (6/67)	13.3 (2/15)	0.634 NS	9.8 (8/82)
Drug use before pregnancy	%				
Cigarette		34.3 (23/67)	53.3 (8/15)	0.239 NS	37.8 (31/82)
Alcohol		53.7 (36/67)	60.0 (9/15)	0.777 NS	54.9 (45/82)
Other drugs		13.4 (9/67)	13.3 (2/15)	>0.999 NS	13.4 (11/82)
Prior pregnancy issues	%				
GDM		1.5 (1/67)	33.3 (5/15)	<0.001 ***	7.3 (6/82)
Hypertensive disorder		4.5 (3/67)	6.7 (1/15)	0.562 NS	4.9 (4/82)
Preterm birth		4.5 (3/67)	6.7 (1/15)	0.562 NS	4.9 (4/82)
Other		10.4 (7/67)	6.7 (1/15)	>0.999 NS	9.8 (8/82)
Prior non-viable pregnancy	%	20.9 (14/67)	20.0 (3/15)	>0.999 NS	20.7 (17/82)
Fertility problems	%	14.9 (10/67)	6.7 (1/15)	0.679 NS	13.4 (11/82)
PCOS	%	25.4 (17/67)	13.3 (2/15)	0.501 NS	23.2 (19/82)
First period age	years	13 (12-14)	12 (11-13)	0.078 NS	13 (12-13)
Last period month	%			0.202 NS	
January		7.5 (5/67)	6.7 (1/15)		7.3 (6/82)
February		6.0 (4/67)	20.0 (3/15)		8.5 (7/82)
March		7.5 (5/67)	0.0 (0/15)		6.1 (5/82)
April		3.0 (2/67)	6.7 (1/15)		3.7 (3/82)
May		13.4 (9/67)	20.0 (3/15)		14.6 (12/82)
June		10.4 (7/67)	13.3 (2/15)		11.0 (9/82)
July		9.0 (6/67)	13.3 (2/15)		9.8 (8/82)
August		7.5 (5/67)	0.0 (0/15)		6.1 (5/82)
September		6.0 (4/67)	0.0 (0/15)		4.9 (4/82)
October		13.4 (9/67)	6.7 (1/15)		12.2 (10/82)
November		10.4 (7/67)	13.3 (2/15)		11.0 (9/82)
December		6.0 (4/67)	0.0 (0/15)		4.9 (4/82)
Personal morbid history	%				
Insulin resistance		3.0 (2/67)	6.7 (1/15)	0.459 NS	3.7 (3/82)
Thyroid disfunction		4.5 (3/67)	6.7 (1/15)	0.562 NS	4.9 (4/82)
Asthma		6.0 (4/67)	0.0 (0/15)	>0.999 NS	4.9 (4/82)
Other		10.4 (7/67)	20.0 (3/15)	0.380 NS	12.2 (10/82)
Family morbid history	%				

Insulin resistance or prediabetes	3.0 (2/67)	6.7 (1/15)	0.459	NS	3.7 (3/82)
DM	32.8 (22/67)	66.7 (10/15)	0.020	*	39.0 (32/82)
Hypertension	41.8 (28/67)	60.0 (9/15)	0.255	NS	45.1 (37/82)
Hypothyroidism	17.9 (12/67)	33.3 (5/15)	0.287	NS	20.7 (17/82)
Hyperthyroidism	1.5 (1/67)	13.3 (2/15)	0.085	NS	3.7 (3/82)
Asthma	7.5 (5/67)	0.0 (0/15)	0.579	NS	6.1 (5/82)
Other	16.4 (11/67)	13.3 (2/15)	>0.999	NS	15.9 (13/82)

NGT: normal glucose tolerance; GDM: gestational diabetes mellitus; BMI: body mass index; PCOS: polycystic ovary syndrome; DM: diabetes mellitus; NS: not significant.

4.4.3.1.2. Prediction of GDM with first trimester serum NIR spectral data

To predict GDM using the biochemical information that serum samples hold, NIR spectra were acquired. **Figure 4.4-1** shows NIR spectra from first trimester NGT and GDM sera. The spectral traces behavior depend on the wavenumber range. In particular, signal sequential noise varies with wavenumber. There is a high-noise region between 10500 and 7600 cm^{-1}, a varying-noise region between 7600 and 5100 cm^{-1}, and a low-noise region between 5100 and 4000 cm^{-1}. Since different spectral ranges present different noise characteristics, they may require different pretreatments before ML analyses. Therefore, NIR spectra were divided in three ranges, according to their sequential noise features: range 1, from 10500 to 7600 cm^{-1} (R1); range 2, from 7600 to 5100 cm^{-1} (R2) and range 3, from 5100 to 4000 cm^{-1} (R3). Posterior analyses considered the three spectral regions, as well as the full range, from 10500 to 4000 cm^{-1} (Full).

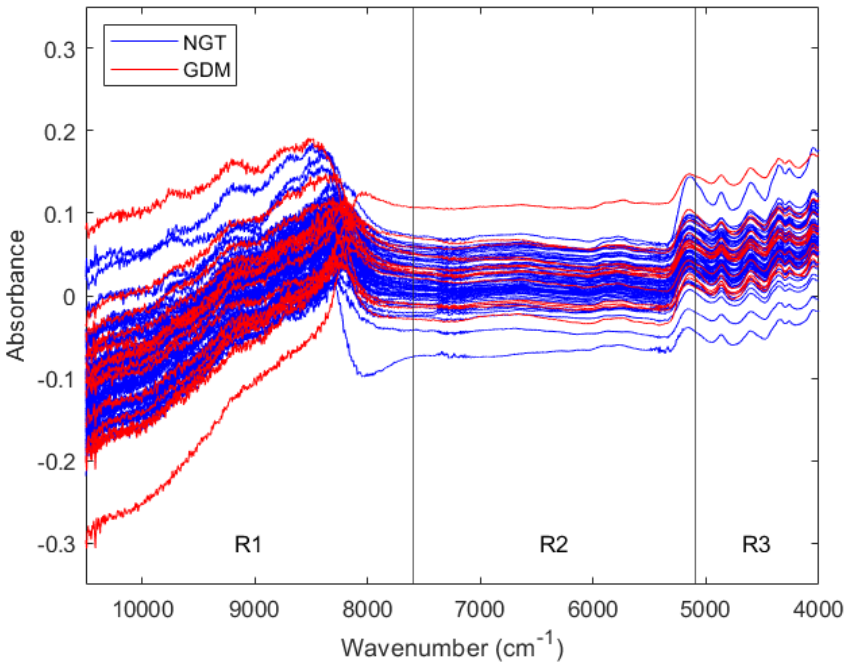


Figure 4.4-1. NIR spectra from first trimester serum samples. Spectra from NGT and GDM pregnant women are colored in blue and red, respectively. Four spectral ranges are considered: full, from 10500 to 4000 cm^{-1}; range 1, from 10500 to 7600 cm^{-1}; range 2, from 7600 to 5100 cm^{-1}; and range 3, from 5100 to 4000 cm^{-1}. *NIR: near-infrared; NGT: normal glucose tolerance; GDM: gestational diabetes mellitus; R1: range 1; R2: range 2; R3: range 3.*

First trimester NIR spectral data was used to develop different single-block predictive models for GDM. For every spectral region, Full, R1, R2 and R3, 80 combinations of pretreatments were tested (**Tables from 4.4-S1 to 4.4-S4**). **Table 4.4-2** presents the characteristics of the best models, i.e. the ones with the highest NER in DCV, for each spectral range. The NIR region with the best predictive performance is R1, with a NER of 0.6321 ± 0.0489. This value is moderately higher than the obtained with the Full spectral range, of 0.5726 ± 0.0410.

To assess if the latter models could be improved, NIR Full and NIR R1 data were combined with the 28 aforementioned medical variables. Different multi-block models were trained and validated (**Table 4.4-S5**). The addition of medical data do not improve the overall predictive performance of the original models, neither in comparison to the models based in NIR spectra only, nor in comparison to a model based in medical data only (NER of 0.6133 ± 0.0298 in DCV). None of the multi-block models outperform the best single-block model, obtained with NIR R1 data only. The simplification of the multi-block models through variable selection do not improve its predictive performance either. As in single-block analyses, NIR R1-based multi-block models tend to show moderately higher NERs than NIR Full-based multi-block models.

Table 4.4-2. Predictive performance of the best ML models using NIR spectral data from first trimester serum samples.

Range [a]	Model number	Pretreatment	Sensitivity		Specificity		Non-error rate	
			Av	StD	Av	StD	Av	StD
Full	80	SM (W=23) + N + MC	0.6946	0.0456	0.4507	0.0681	0.5726	0.0410
R1	2	N+MC	0.6722	0.0361	0.5920	0.0910	0.6321	0.0489
R2	75	SM (W=3) + N + MC	0.5678	0.0322	0.6480	0.1035	0.6079	0.0542
R3	38	SM (W=23) + MC	0.5931	0.0346	0.5133	0.0811	0.5532	0.0441

[a] Full: 10500-4000 cm^{-1}; R1: 10500-7600 cm^{-1}; R2: 7600-5100 cm^{-1}; R3: 5100-4000 cm^{-1}.

R1: range 1; R2: range 2; R3: range 3; Av: average; StD: standard deviation; SM: smoothing; W: width; N: normalization; MC: mean centering.

Figure 4.4-2 presents the overall predictive performance of the best ML model obtained using NIR spectra from first trimester serum samples. It corresponds to the NIR R1 spectral range (10500-7600 cm^{-1}) with pretreatment by normalization and mean centering. It predicts GDM with a DCV AUC-ROC of 0.5768 ± 0.0635. Moreover, the most relevant spectral intervals for the performance of this model, i.e. those mainly composed of variables with VIP scores larger than 1, are 10500-9828 cm^{-1} and 8826-7858 cm^{-1} (data not shown).

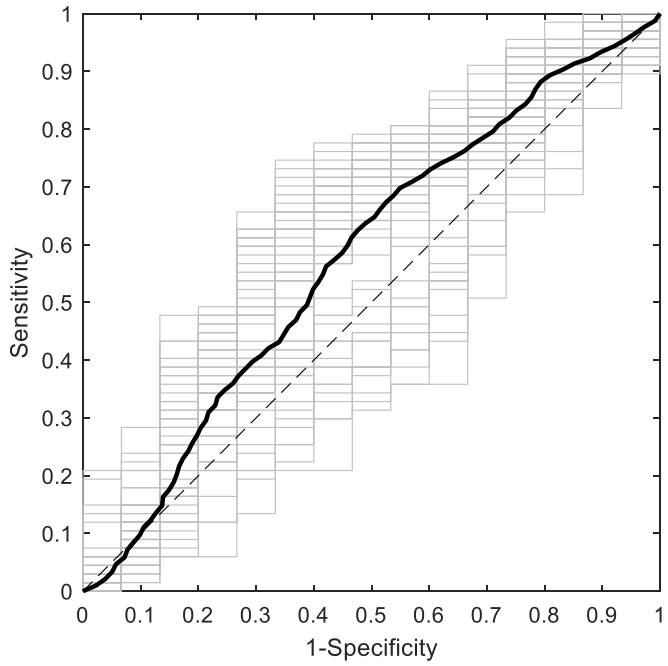


Figure 4.4-2. ROC curve of the best predictive model for the first trimester cohort. The model was trained with NIR spectra (R1, 10500-7600 cm^{-1}) from first trimester serum samples, after pretreatment by normalization and mean centering. The average and standard deviation of 50 DCV repetitions are colored in black and gray, respectively. *ROC: receiver operating characteristic; NIR: near-infrared; R1: range 1; DCV: double cross-validation.*

4.4.3.2. Second trimester cohort

4.4.3.2.1. Description of the second trimester cohort

To characterize this cohort, classical statistical analyses were performed. **Table 4.4-3** presents the same 28 medical parameters considered for the first trimester cohort, and compares their behavior in NGT and GDM subjects. In this cohort, the prevalence of GDM was 17.0%. There are only two variables that statistically differ between the two groups: BMI and history of GDM in a prior pregnancy. The former is higher and the latter is more frequent in GDM pregnancies than in NGT pregnancies.

Table 4.4-3. Medical variables in the second trimester cohort.

Variable	Unit	NGT (n=39)	GDM (n=8)	p value		All (n=47)
Age	years	29 ± 5	30 ± 7	0.606	NS	29 ± 5
BMI	Kg/m^2	27.0 ± 4.7	31.3 ± 6.5	0.034	*	27.7 ± 5.2
Supplement consumption	%	64.1 (25/39)	62.5 (5/8)	>0.999	NS	63.8 (30/47)
Hyperemesis	%	33.3 (13/39)	25.0 (2/8)	>0.999	NS	31.9 (15/47)
Vaginal bleeding	%	5.1 (2/39)	25.0 (2/8)	0.129	NS	8.5 (4/47)
Drug use before pregnancy	%					
Cigarette		33.3 (13/39)	37.5 (3/8)	>0.999	NS	34.0 (16/47)
Alcohol		61.5 (24/39)	50.0 (4/8)	0.697	NS	59.6 (28/47)
Other drugs		25.6 (10/39)	0.0 (0/8)	0.174	NS	21.3 (10/47)
Prior pregnancy issues	%					
GDM		0.0 (0/39)	37.5 (3/8)	0.004	**	6.4 (3/47)
Hypertensive disorder		7.7 (3/39)	0.0 (0/8)	>0.999	NS	6.4 (3/47)
Preterm birth		5.1 (2/39)	12.5 (1/8)	0.436	NS	6.4 (3/47)
Other		7.7 (3/39)	0.0 (0/8)	>0.999	NS	6.4 (3/47)
Prior non-viable pregnancy	%	17.9 (7/39)	12.5 (1/8)	>0.999	NS	17.0 (8/47)
Fertility problems	%	17.9 (7/39)	0.0 (0/8)	0.329	NS	14.9 (7/47)
PCOS	%	25.6 (10/39)	12.5 (1/8)	0.659	NS	23.4 (11/47)
First period age	years	13 (12-14)	12 (12-13)	0.058	NS	13 (12-14)

Last period month	%			0.729	NS	
January		2.6 (1/39)	0.0 (0/8)			2.1 (1/47)
February		7.7 (3/39)	25.0 (2/8)			10.6 (5/47)
March		12.8 (5/39)	0.0 (0/8)			10.6 (5/47)
April		5.1 (2/39)	12.5 (1/8)			6.4 (3/47)
May		12.8 (5/39)	0.0 (0/8)			10.6 (5/47)
June		10.3 (4/39)	12.5 (1/8)			10.6 (5/47)
July		15.4 (6/39)	25.0 (2/8)			17.0 (8/47)
August		12.8 (5/39)	0.0 (0/8)			10.6 (5/47)
September		2.6 (1/39)	12.5 (1/8)			4.3 (2/47)
October		10.3 (4/39)	12.5 (1/8)			10.6 (5/47)
November		5.1 (2/39)	0.0 (0/8)			4.3 (2/47)
December		2.6 (1/39)	0.0 (0/8)			2.1 (1/47)
Personal morbid history	%					
Insulin resistance		5.1 (2/39)	0.0 (0/8)	>0.999	NS	4.3 (2/47)
Thyroid disfunction		10.3 (4/39)	0.0 (0/8)	>0.999	NS	8.5 (4/47)
Asthma		7.7 (3/39)	0.0 (0/8)	>0.999	NS	6.4 (3/47)
Other		10.3 (4/39)	37.5 (3/8)	0.084	NS	14.9 (7/47)
Family morbid history	%					
Insulin resistance or prediabetes		7.7 (3/39)	12.5 (1/8)	0.539	NS	8.5 (4/47)
DM		35.9 (14/39)	62.5 (5/8)	0.240	NS	40.4 (19/47)
Hypertension		48.7 (19/39)	62.5 (5/8)	0.701	NS	51.1 (24/47)
Hypothyroidism		17.9 (7/39)	25.0 (2/8)	0.639	NS	19.1 (9/47)
Hyperthyroidism		5.1 (2/39)	12.5 (1/8)	0.436	NS	6.4 (3/47)
Asthma		10.3 (4/39)	0.0 (0/8)	>0.999	NS	8.5 (4/47)
Other		12.8 (5/39)	12.5 (1/8)	>0.999	NS	12.8 (6/47)

NGT: normal glucose tolerance; GDM: gestational diabetes mellitus; BMI: body mass index; PCOS: polycystic ovary syndrome; DM: diabetes mellitus; NS: not significant.

4.4.3.2.2. Prediction of GDM with second trimester serum NIR spectral data

To predict GDM taking advantage of the biochemistry of sera, NIR spectra were recorded. **Figure 4.4-3** exhibits NIR spectra from second trimester NGT and GDM serum samples. Due to their sequential noise behavior, NIR spectra were divided in the same three regions considered for the first trimester cohort: R1 (10500-7600 cm^{-1}), R2 (7600-5100 cm^{-1}) and R3 (5100- 4000 cm^{-1}). Subsequent analyses

considered both the three NIR regions, and the Full spectral range (10500 to 4000 cm^{-1}).

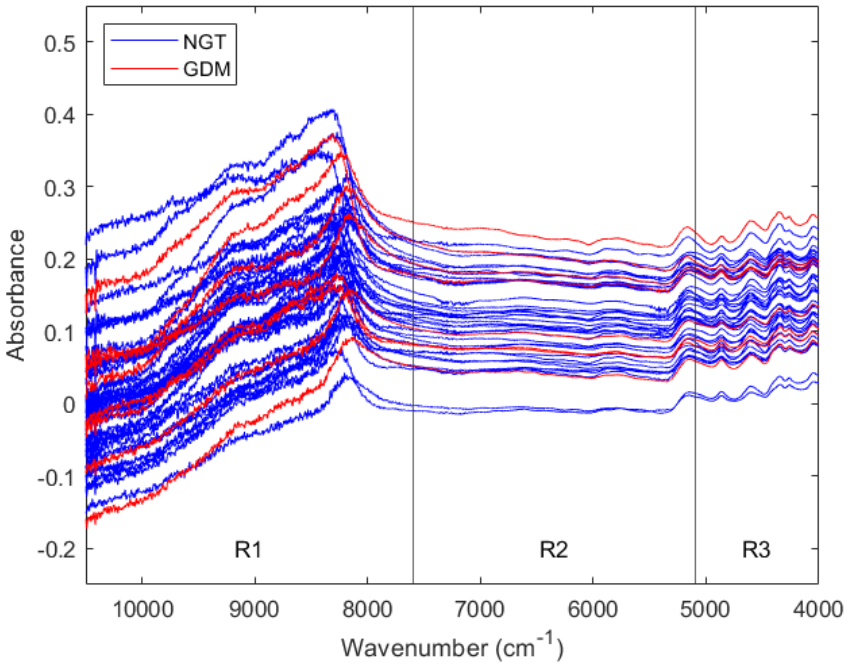


Figure 4.4-3. NIR spectra from second trimester serum samples. Spectra from NGT and GDM pregnant women are colored in blue and red, respectively. Four spectral ranges are considered: full, from 10500 to 4000 cm^{-1}; range 1, from 10500 to 7600 cm^{-1}; range 2, from 7600 to 5100 cm^{-1}; and range 3, from 5100 to 4000 cm^{-1}. *NIR: near-infrared; NGT: normal glucose tolerance; GDM: gestational diabetes mellitus; R1: range 1; R2: range 2; R3: range 3.*

Second trimester NIR spectral data was used to train and validate different single-block models for GDM prediction. For each spectral range, Full, R1, R2 and R3, 80 combinations of pretreatments were assessed (**Tables from 4.4-S6 to 4.4-S9**). **Table 4.4-4** displays the features of the best models, i.e. the ones with the

highest NER in DCV, for every NIR region. The range with the greatest predictive power is R3, with a NER of 0.7894 ± 0.0431. This performance is much better than the obtained with the Full NIR range, of 0.4642 ± 0.0321.

To evaluate if the performance of the latter predictive models could be enhanced, NIR Full and NIR R3 were combined with the 28 medical parameters mentioned above. Different multi-block models were developed (**Table 4.4-S10**). The combination of NIR Full with medical data improves the overall performance in comparison to the Full NIR single-block model. Likewise, the combination of NIR R3 with medical data increases the overall predictive power in comparison to a model based on medical data only (NER of 0.6115 ± 0.0467 in DCV). Nevertheless, the addition of medical data do not improve the predictive performance in comparison to the best single-block model, obtained with NIR R3 data only. The simplification of the multi-block models by means of variable selection do not outperform its predictive performance either. In line with what happens in single block-analyses, NIR R3-based multi-block models tend to present higher NERs than NIR Full-based multi-block models.

Table 4.4-4. Predictive performance of the best ML models using NIR spectral data from second trimester serum samples.

Range [a]	Model number	Pretreatment	Sensitivity		Specificity		Non-error rate	
			Av	StD	Av	StD	Av	StD
Full	30	2D (W=15) + N + MC	0.8133	0.0324	0.1150	0.0556	0.4642	0.0321
R1	4	WLS+N+MC	0.8754	0.0414	0.1625	0.1218	0.5189	0.0643

R2	27	2D (W=3) + N + MC	0.6821	0.0288	0.3875	0.1191	0.5348	0.0613
R3	12	1D (W=15) + MC	0.8713	0.0361	0.7075	0.0783	0.7894	0.0431

[a] Full: 10500-4000 cm^{-1}; R1: 10500-7600 cm^{-1}; R2: 7600-5100 cm^{-1}; R3: 5100-4000 cm^{-1}.

R1: range 1; R2: range 2; R3: range 3; Av: average; StD: standard deviation; 2D: second derivative; W: width; N: normalization; MC: mean centering; WLS: weighted least squares; 1D: first derivative.

Figure 4.4-4 displays the overall predictive performance of the best ML model obtained employing NIR spectra from second trimester sera. It corresponds to the NIR R3 spectral region (5100-4000 cm^{-1}) with pretreatment by first derivative (width=15) and mean centering. It predicts GDM with an AUC-ROC of 0.8836 ± 0.0259 in DCV. Furthermore, the most relevant spectral intervals for the performance of this model, i.e. those mainly composed of variables with VIP scores larger than 1, are 5028-4856 cm^{-1}, 4764-4702 cm^{-1}, 4492-4442 cm^{-1}, 4392-4364 cm^{-1}, 4302-4268 cm^{-1}, 4206-4176 cm^{-1} and 4096-4000 cm^{-1} (data not shown).

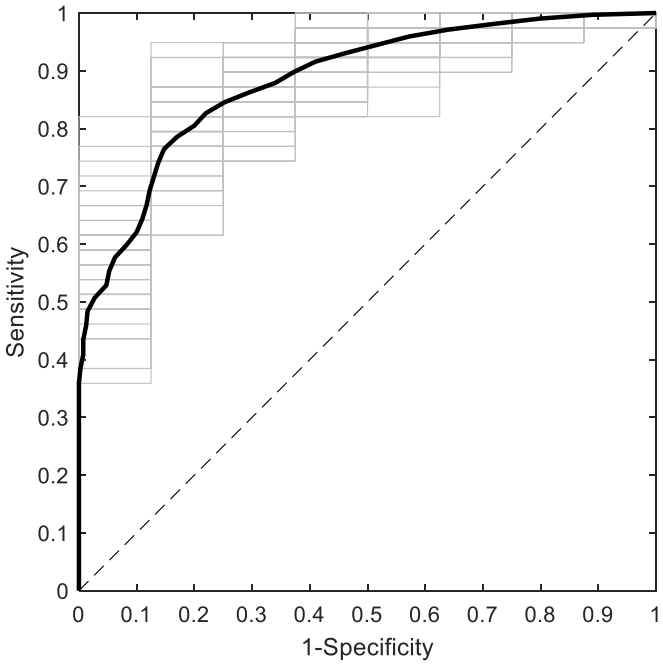


Figure 4.4-4. ROC curve of the best predictive model for the second trimester cohort. The model was trained with NIR spectra (R3, 5100-4000 cm^{-1}) from second trimester serum samples, after pretreatment by first derivative (width=15) and mean centering. The average and standard deviation of 50 DCV repetitions are colored in black and gray, respectively. *ROC: receiver operating characteristic; NIR: near-infrared; R3: range 3; DCV: double cross-validation.*

4.4.4. Discussion

This work shows that ML modeling with NIR spectra from first trimester sera leads to a moderate performance for the prediction of GDM. It also shows that modeling with NIR data from second trimester samples results in a high predictive power for GDM. In both cases, the entire method takes only 32 minutes, considering both sample preparation and data acquisition by NIR spectroscopy. These

findings suggest that the second trimester NIR spectroscopy-based method could be used as an alternative screening tool for GDM.

4.4.4.1. The addition of medical data does not improve the predictive performance of NIR data-based models

The prevalence of GDM in the study cohorts was 18.3% and 17.0% (**Table 4.4-1** and **Table 4.4-3**), higher than that reported for Chilean population in 2015, of 13.0% [27]. This behavior is consistent with the fact that the prevalence of GDM is increasing both in Chile [27] and worldwide [28]. The medical variables that differed between GDM and NGT groups were history of GDM in a prior pregnancy, family history of DM and BMI (**Table 4.4-1** and **Table 4.4-3**). Higher frequencies or levels of these variables were observed in pregnant women with GDM. This observation makes sense, since they are known risk factors for GDM both in Chilean [27] and global population [1]. The statistical behavior of these risk factors was not exactly the same in the first and the second trimester cohorts, however, in a multivariate scale the two cohorts behaved similarly. In fact, full medical data allowed to predict GDM with a very similar overall performance in the first and the second trimester cohorts, with DCV-NERs of 0.6133 ± 0.0298 and 0.6115 ± 0.0467, respectively. Moreover, the addition of medical data with multi-block techniques did not improve the performance of NIR data-based predictive models

in any of the cohorts, even when all statistically relevant variables were part of the best multi-block models (**Table 4.4-S5** and **Table 4.4-S10**).

Even though multi-block analysis is often associated with an increased predictive power compared to single-block analysis, this is not always the case [29]. One strategy that can improve predictive performance is to apply variable selection [24,26,30], however in this study it did not have that effect. Another strategy is to use multi-block techniques of a higher data fusion level [29]. Multi-block analysis can be related to low-, mid- or high-level data fusion (LLDF, MLDF and HLDF, respectively). LLDF techniques work directly on the original data blocks; MLDF techniques operate on features extracted from each data block; and HLDF techniques fuse the outcome of models built from each data block [31]. The multi-block technique applied here, SO-PLS-LDA, corresponds to MLDF, since information is sequentially extracted from the different blocks to construct the model [23]. Therefore, the application of a HLDF technique may increase the predictive power of the models presented here. The use of such technique should be considered carefully, since HLDF modelling is more complex, time-consuming and more difficult to interpret [31].

4.4.4.2. NIR data-based prediction has advantages over medical data-prediction

NIR data-based models performed as well as or better than medical data-based models. In fact, the best model built on NIR spectral data from first trimester serum samples (**Table 4.4-2**) showed an overall performance similar to that obtained with medical data in the same cohort, with NERs of 0.6321 ± 0.0489 and 0.6133 ± 0.0298, respectively. Likewise, the best model built on NIR spectral data from second trimester serum samples (**Table 4.4-4**) presented a much higher predictive power than the one obtained with medical data in the same cohort, with NERs of 0.7894 ± 0.0431 and 0.6115 ± 0.0467, respectively.

It is important to mention that all the medical parameters considered here are clinical. Clinical variables involve anthropometrical measurements, demographical parameters, and personal or family morbid history data. In general, models involving this type of variables are associated with a moderate performance for the prediction of GDM [14]. Furthermore, this kind of information is generally obtained by means of self-report questionnaires and, therefore, are subjected to bias. In contrast, NIR spectral data is obtained through the objective analysis of biological samples. Hence, besides presenting a similar or higher predictive power than medical data-based models, NIR data-based models are less subjective.

4.4.4.3. NIR spectral data pretreatment is essential to maximize predictive power

In the two study cohorts, particular spectral regions achieved a better predictive performance than full spectral ranges. In the case of first trimester serum samples, R1 presented a NER of 0.6321 ± 0.0489, moderately higher than the obtained with the Full range, of 0.5726 ± 0.0410 (**Table 4.4-2**). In the case of second trimester serum samples, R3 exhibited a NER of 0.7894 ± 0.0431, much higher than the obtained with the Full range, of 0.4642 ± 0.0321 (**Table 4.4-4**). This tendency was maintained when medical data was added (**Table 4.4-S5** and **Table 4.4-S10**).

It is likely that spectral segmentation allowed to better optimize data pretreatment. Pretreatment operations are used to remove chemically irrelevant sources of variation in the data, reducing the contribution of signals that are not related to the property being predicted, and improving the performance of both qualitative and quantitative analyses [21]. Typical pretreatment for infrared (IR) data includes selecting the optimal wavenumber range, and correcting different spectral alterations, such as random and systematic noise, light scattering, baseline shift, among others [21,22]. The effect of pretreatment on predictive power is highly data-dependent [22]. Therefore, it is important to optimize it depending on the spectral features that need to be corrected in each particular case. Here, signal

sequential noise varied between NIR spectral regions (**Figure 4.4-1** and **Figure 4.4-3**), suggesting that they may require different pretreatments. Indeed, the optimal pretreatment for each spectral region was different, both between short spectral regions and compared to the Full NIR range (**Table 4.4-2** and **Table 4.4-4**).

4.4.4.4. Predictive performance in the first and the second trimester is related to biochemical changes occurring throughout GDM

The optimal NIR ranges differed between trimesters. For first and second trimester sera, the spectral ranges with higher predictive power were 10500-7600 cm^{-1} (**Table 4.4-2**) and 5100-4000 cm^{-1} (**Table 4.4-4**), respectively. They were associated with AUC-ROCs of 0.5768 ± 0.0635 (**Figure 4.4-2**) and 0.8836 ± 0.0259 (**Figure 4.4-4**), respectively. This difference in optimal spectral range and associated predictive performance, might be related to the biochemical changes that underly the development of GDM.

The biochemical interpretation of NIR spectra is a challenging task, since the bands observed in NIR spectra are mainly due to overtones and combination bands of fundamental vibrational modes [15]. However, it is possible to tentatively relate spectral patterns to particular biomolecules [32–34]. In the first trimester best model, two NIR intervals stood out, 10500-9828 cm^{-1} and 8826-7858 cm^{-1},

whereas in the second trimester case, six spectral intervals did so, 5028-4856 cm^{-1}, 4764-4702 cm^{-1}, 4492-4442 cm^{-1}, 4392-4364 cm^{-1}, 4302-4268 cm^{-1}, 4206-4176 cm^{-1} and 4096-4000 cm^{-1}. These first and second trimester spectral intervals involve vibrations of various chemical bonds, among which there are some that have been associated with carbohydrates, lipids and proteins (**Table 4.4-S11** and **Table 4.4-S12**). Therefore, these three biomolecules would be altered in GDM, in both trimesters of pregnancy.

Based on the tentative assignments made, the potential biochemical differences between the two trimesters are not evident. However, there is a key difference between them in GDM. The hyperglycemia state that characterizes GDM manifests only in the late second trimester or in the early third trimester of pregnancy [1]. In other words, while glycemia is not altered in the first trimester, it is altered in the second trimester. Interestingly, the optimal spectral range for predicting GDM in the second trimester (5100-4000 cm^{-1}) has been identified as relevant for quantifying glucose in serum samples. Indeed, Goodarzi and Saeys found that 2100-2300 nm (4762-4348 cm^{-1}) was the most important NIR region for glucose quantification in human serum. They discussed that this was consistent with previous studies, which had identified 2000-2500 nm (5000-4000 cm^{-1}) as the most informative wavelength range for glucose measurement [35]. Their result is in line with the tentative assignments performed here, in which wavenumbers near 4762-4348 cm^{-1} were related to carbohydrates (**Table 4.4-**

S12). In consequence, it is very likely that the best second trimester model achieves a better predictive performance than the best first trimester counterpart, because it accounts for biochemical changes that become evident only when GDM is fully established.

4.4.4.5. NIR data-based prediction has advantages over other instrumental data-based prediction

The NIR-based models presented here allowed to predict GDM in 32 minutes, considering sample preparation and spectral acquisition of each instrumental replicate. This is fast compared to other instrumental-based methods reported in literature.

There are studies applying LC-MS- or GC-MS-based methods to predict GDM at different stages of pregnancy. Their predictive performance varies, e.g. with AUC-ROCs of 0.7075 [36], 0.745-0.797 [37] and 0.729-0.906 [38] before GDM diagnosis; and 0.7800 [36], 0.745-0.828 [37] and 0.83-0.90 [39] at the time of GDM diagnosis. Even though some of these methods achieve a high predictive power, they are very time-consuming. For instance, the LC-MS metabolomics strategy of Zhang et al. [36] takes approximately 1.5 hours, with a sample preparation step of at least 15 minutes, and two LC-MS runs of 30 minutes each. Likewise, the LC-MS proteomics approach of Guo et al. [38] includes a sample

preparation process that takes more than 4 hours. Similarly, the GC-MS metabolomics methods of Raczkowska et al. [37] and Dudzik et al. [39] require a sample preparation procedure of more than 16 hours.

Other articles use NMR metabolomics-based methods to predict GDM at different points in pregnancy [40–44]. Their predictive power is variable, but tends to be better at the time of GDM diagnosis than earlier, e.g. with AUC-ROCs of 0.62 and 0.59, respectively, in the study of McBride et al. [40] and NERs of 69.5-88.5% and 63.5-82.5%, respectively, in the study of Pinto et al. [41]. Papers presenting this kind of strategy do not usually mention details about time of analysis in the methodology section. However, sample preparation of biological fluids for this type of analysis usually takes 1-1.5 hours, while NMR data acquisition typically takes 4-5 minutes per sample [45].

There are also works employing PCR-based methods for GDM prediction. Some authors have based their methods on single nucleotide polymorphisms. Such methods are valid to predict GDM at any point of life. Their predictive power is moderate, e.g. with an accuracy of 0.531-0.552 [46] and an AUC-ROC of 0.7694 [47]. Moreover, they are associated with long analysis times. For example, the strategy of Yu et al. [46] consists of the extraction of genomic DNA, and its analysis by PCR-restriction fragment length polymorphism (PCR-RFLP). The PCR-RFLP protocol alone takes more than 2 hours. Likewise, the method of

Zulueta et al. [47] consists of genomic DNA extraction, and its analysis by iPLEX-PCR, using the MassARRAY system from Agena Bioscience. According to the manufacturer, the entire workflow for iPLEX MassARRAY PCR takes 8 hours [48]. Some other authors have based their GDM predictive methods on micro-RNAs. Their predictive performance varies, e.g. with AUC-ROCs of 60.0-66.9% [49] before GDM diagnosis, and AUC-ROCs of 0.74-0.92 [50] at the time of GDM diagnosis. Although some of these methods reach a high predictive power, they involve long times of analysis. For instance, the approaches of Zhao et al. [49] and Cao et al. [50] consist of RNA extraction, reverse transcription and a TaqMan-based quantitative PCR (TaqMan-qPCR) that takes, alone, about 1 hour. Furthermore, sample preparation for TaqMan-qPCR usually takes more than 1.5 hours [51].

Even though the best first trimester method presented here is simple and fast, it showed a moderate performance for the prediction of GDM (AUC-ROC of 0.5768 ± 0.0635, **Figure 4.4-2**) compared to that reported in literature with more time-consuming methods. This NIR data-based method could be improved by modifying sample preparation, e.g. by removing from sera the high concentration proteins that might be interfering with the analysis of lower concentration biomolecules, which could be important to differentiate the two study groups. This would increase the time of analysis, however it could be adjusted, for example, by reducing the drying time. The simplicity and rapidity of this method, coupled

with an improved predictive power, would make it ideal for the early detection of GDM. On the other hand, the best second trimester method presented here exhibited a very high predictive power (AUC-ROC of 0.8836 ± 0.0259, **Figure 4.4-4**), similar or better than that reported in literature with much slower methods. This predictive performance, together with its simplicity and rapidity, makes it an excellent alternative screening method for GDM.

4.4.4.6. The proposed strategy has advantages over other IR-based strategies

Before this work, NIR spectroscopy had never been assessed as a diagnostic support tool for GDM. There is only one study applying IR spectroscopy for the prediction of this pregnancy disease, that of Bernardes-Oliveira et al. Their method consists of the analysis of plasma samples with attenuated total reflection Fourier-transform mid-IR spectroscopy, and is able to predict GDM with an accuracy of 100% [52]. The limitation of this strategy is that plasma samples were collected in a very wide time range, 9-39 or 12-38 weeks of pregnancy for the control or the GDM group, respectively. Therefore, even though their study showed the great potential of IR spectroscopy to differentiate subjects with and without GDM, it was not designed to predict GDM at particular stages of pregnancy. In contrast, in the present study serum samples were collected in the

first and the second trimester of pregnancy, allowing us to evaluate the capability of IR spectroscopy to predict GDM both before and at the time of diagnosis.

4.4.4.7. Strengths of this study

To our knowledge, this is the first study reporting NIR spectra-based methods for the prediction of GDM, either as early detection or alternative screening tools. They are simpler and faster than other strategies proposed in literature to predict this pregnancy disease. Moreover, the best second trimester model achieved a highly competitive predictive power compared to methods from literature, making it ideal as an alternative screening tool for GDM. In addition, NIR data pretreatment was performed in an exhaustive and systematic manner, enabling to maximize the predictive power for both first and second trimester sera. Finally, every model was subjected to DCV, allowing to obtain reliable results despite of a limited sample size.

4.4.4.8. Limitations of this study

The sample size is small. Models constructed on a limited number of samples are prone to overfitting. Even though DCV was used to minimize this effect, future external validation studies are needed to confirm the effectiveness of the developed methods to predict GDM in different populations. In addition, the

predictive models for GDM presented here are restricted to the Chilean diagnostic criteria. Further studies should be performed to evaluate the performance of NIR spectra-based methods for the prediction of GDM under other diagnostic criteria. Finally, this is not a longitudinal study, so the comparison between first and second trimester results should be made with caution.

4.4.5. Conclusion

In this work, NIR spectroscopy of serum samples was evaluated for the prediction of GDM at different stages of pregnancy. NIR data-based predictive models were methodically optimized and robustly validated. The developed methods are simple and fast, and have a great potential to be applied as clinical decision support tools in medical practice. Even though the first trimester approach should still be improved to be applied as an early detection tool for GDM, the second trimester strategy presents characteristics that make it suitable to be used as an alternative screening tool to the OGTT at the time of GDM diagnosis, e.g. a high predictive power for GDM, simplicity and rapidity. Further studies are needed to confirm these findings in other populations, and under different GDM diagnostic criteria.

Acknowledgments

The authors would like to thank the volunteers who participated in this work. Thanks also to the health staff of CESFAM Victor Manuel Fernández, CESFAM Santa Sabina, and CESFAM Tucapel; and Joanna Tereszczuk from CMA Bio-Bío, for their technical support.

Funding

This study was financed by Agencia Nacional de Investigación y Desarrollo via PhD scholarship 21190736 (DM), FONDECYT 11181153 (JA) and FOVI 210057 (EG-G); and Universidad de Concepción/Ministerio de Educación through Proyecto UCO 1866 (DM).

4.4.6. References

1. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. Nat Rev Dis Prim. 2019;5: 47. doi:10.1038/s41572-019-0098-8

2. Sweeting A, Wong J, Murphy HR, Ross GP. A Clinical Update on Gestational Diabetes Mellitus. Endocr Rev. 2022;43: 763–793. doi:10.1210/endrev/bnac003

3. Tsakiridis I, Giouleka S, Mamopoulos A, Kourtis A, Athanasiadis A, Filopoulou D, et al. Diagnosis and Management of Gestational Diabetes Mellitus: An Overview of National and International Guidelines. Obstet Gynecol Surv. 2021;76: 367–381. doi:10.1097/OGX.0000000000000899

4. Agarwal MM, Punnose J, Dhatt GS. Gestational diabetes: Problems associated with the oral glucose tolerance test. Diabetes Res Clin Pract. 2004;63: 73–74. doi:10.1016/j.diabres.2003.08.005

5. Lachmann EH, Fox RA, Dennison RA, Usher-Smith JA, Meek CL, Aiken CE. Barriers to completing oral glucose tolerance testing in women at risk of gestational diabetes. Diabet Med. 2020;37: 1482–1489. doi:10.1111/dme.14292

6. Hanna FWF, Peters JR. Screening for gestational diabetes; past, present and future. Diabet Med. 2002;19: 351–358. doi:10.1046/j.1464-5491.2002.00684.x

7. Huhn EA, Rossi SW, Hoesli I, Göbl CS. Controversies in Screening and Diagnostic Criteria for Gestational Diabetes in Early and Late Pregnancy. Front Endocrinol (Lausanne). 2018;9: 5–12. doi:10.3389/fendo.2018.00696

8. Bonongwe P, Lindow SW, Coetzee EJ. Reproducibility of a 75G oral glucose tolerance test in pregnant women. J Perinat Med. 2015;43: 333–338. doi:10.1515/jpm-2014-0208

9. Munang YN, Noubiap JJ, Danwang C, Sama JD, Azabji-Kenfack M, Mbanya JC, et al. Reproducibility of the 75 g oral glucose tolerance test for the diagnosis of gestational diabetes mellitus in a sub-Saharan African population. BMC Res Notes. 2017;10: 1–6. doi:10.1186/s13104-017-2944-7

10. Sovio U, Murphy HR, Smith GCS. Accelerated fetal growth prior to diagnosis of gestational diabetes mellitus: A prospective cohort study of nulliparous women. Diabetes Care. 2016;39: 982–987. doi:10.2337/dc16-0160

11. Venkataraman H, Ram U, Craik S, Arungunasekaran A, Seshadri S, Saravanan P. Increased fetal adiposity prior to diagnosis of gestational diabetes in South Asians: more evidence for the ‘thin–fat’ baby. Diabetologia. 2017;60: 399–405. doi:10.1007/s00125-016-4166-2

12. Yovera L, Zaharia M, Jachymski T, Velicu-Scraba O, Coronel C, de Paco Matallana C, et al. Impact of gestational diabetes mellitus on fetal cardiac morphology and function: cohort comparison of second- and third-trimester fetuses. Ultrasound Obstet Gynecol. 2021;57: 607–613. doi:10.1002/uog.22148

13. Goecks J, Jalili V, Heiser LM, Gray JW. How Machine Learning Will Transform Biomedicine. Cell. 2020;181: 92–101. doi:10.1016/j.cell.2020.03.022

14. Mennickent D, Rodríguez A, Farías-Jofré M, Araya J, Guzmán-Gutiérrez E. Machine learning-based models for gestational diabetes mellitus prediction before 24–28 weeks of pregnancy: A review. Artif Intell Med. 2022;132: 102378. doi:10.1016/j.artmed.2022.102378

15. Ozaki Y, Huck C. Near-Infrared Spectroscopy. Springer; 2021. doi:10.1007/978-981-15-8648-4

16. Bunaciu AA, Fleschin Ş, Hoang VD, Aboul-Enein HY. Vibrational Spectroscopy in Body Fluids Analysis. Crit Rev Anal Chem. 2017;47: 67–75. doi:10.1080/10408347.2016.1209104

17. Balan V, Mihai CT, Cojocaru FD, Uritu CM, Dodi G, Botezat D, et al. Vibrational spectroscopy fingerprinting in medicine: From molecular to clinical practice. Materials (Basel). 2019;12: 1–40. doi:10.3390/ma12182884

18. Sakudo A. Near-infrared spectroscopy for medical applications: Current status and future perspectives. Clin Chim Acta. 2016;455: 181–188. doi:10.1016/j.cca.2016.02.009

19. Villena Gonzales W, Mobashsher A, Abbosh A. The Progress of Glucose Monitoring—A Review of Invasive to Minimally and Non-Invasive Techniques, Devices and Sensors. Sensors. 2019;19: 800. doi:10.3390/s19040800

20. Ministerio de Salud. Guía Perinatal. 2015.

21. Morais CLM, Lima KMG, Singh M, Martin FL. Tutorial: multivariate classification for vibrational spectroscopy in biological samples. Nat Protoc. 2020;15: 2143–2162. doi:10.1038/s41596-020-0322-8

22. Jiao Y, Li Z, Chen X, Fei S. Preprocessing methods for near-infrared spectrum calibration. J Chemom. 2020;34: 1–19. doi:10.1002/cem.3306

23. Biancolillo A, Måge I, Næs T. Combining SO-PLS and linear discriminant analysis for multi-block classification. Chemom Intell Lab Syst. 2015;141: 58–67. doi:10.1016/j.chemolab.2014.12.001

24. Cocchi M, Biancolillo A, Marini F. Chemometric Methods for Classification and Feature Selection. Comprehensive Analytical Chemistry. Elsevier; 2018. pp. 265–299. doi:10.1016/bs.coac.2018.08.006

25. Ballabio D, Grisoni F, Todeschini R. Multivariate comparison of classification performance measures. Chemom Intell Lab Syst. 2018;174: 33–44. doi:10.1016/j.chemolab.2017.12.004

26. Mennickent D, Ortega-Contreras B, Gutiérrez-Vega S, Castro E, Rodríguez A, Araya J, et al. Evaluation of first and second trimester maternal thyroid profile on the prediction of gestational diabetes mellitus and post load glycemia. PLoS One. 2023;18: e0280513. doi:10.1371/journal.pone.0280513

27. Garmendia ML, Mondschein S, Montiel B, Kusanovic JP. Trends and predictors of gestational diabetes mellitus in Chile. Int J Gynecol Obstet. 2020;148: 210–218. doi:10.1002/ijgo.13023

28. International Diabetes Federation. IDF Diabetes Atlas. 2019.

29. Azcarate SM, Ríos-Reina R, Amigo JM, Goicoechea HC. Data handling in data fusion: Methodologies and applications. TrAC - Trends Anal Chem. 2021;143. doi:10.1016/j.trac.2021.116355

30. Post MJ, van der Putten P, van Rijn JN. Does Feature Selection Improve Classification? A Large Scale Experiment in OpenML. Lecture Notes in Computer Science. Springer; 2016. pp. 158–170. doi:10.1007/978-3-319-46349-0_14

31. Cocchi M. Data Fusion Methodology and Applications. Elsevier; 2019.

32. Workman Jr. J, Weyer L. Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy. CRC Press; 2012. doi:10.1201/b11894

33. Beć KB, Grabska J, Huck CW. Near-Infrared Spectroscopy in Bio-Applications. Molecules. 2020;25: 2948. doi:10.3390/molecules25122948

34. Golic M, Walsh K, Lawson P. Short-Wavelength Near-Infrared Spectra of Sucrose, Glucose, and Fructose with Respect to Sugar Concentration and Temperature. Appl Spectrosc. 2003;57: 139–145. doi:10.1366/000370203321535033

35. Goodarzi M, Saeys W. Selection of the most informative near infrared spectroscopy wavebands for continuous glucose monitoring in human serum. Talanta. 2016;146: 155–165. doi:10.1016/j.talanta.2015.08.033

36. Zhang H, Zhao Y, Zhao D, Chen X, Khan NU, Liu X, et al. Potential biomarkers identified in plasma of patients with gestational diabetes mellitus. Metabolomics. 2021;17: 1–12. doi:10.1007/s11306-021-01851-x

37. Raczkowska BA, Mojsak P, Rojo D, Telejko B, Paczkowska–Abdulsalam M, Hryniewicka J, et al. Gas Chromatography–Mass Spectroscopy-Based Metabolomics Analysis Reveals Potential Biochemical Markers for Diagnosis of Gestational Diabetes Mellitus. Front Pharmacol. 2021;12: 1–9. doi:10.3389/fphar.2021.770240

38. Guo Y, Han Z, Guo L, Liu Y, Li G, Li H, et al. Identification of urinary biomarkers for the prediction of gestational diabetes mellitus in early second trimester of young gravidae based on iTRAQ quantitative proteomics. Endocr J. 2018;65: 727–735. doi:10.1507/endocrj.EJ17-0471

39. Dudzik D, Zorawski M, Skotnicki M, Zarzycki W, García A, Angulo S, et al. GC–MS based Gestational Diabetes Mellitus longitudinal study: Identification of 2-and 3-hydroxybutyrate as potential prognostic biomarkers. J Pharm Biomed Anal. 2017;144: 90–98. doi:10.1016/j.jpba.2017.02.056

40. McBride N, Yousefi P, White SL, Poston L, Farrar D, Sattar N, et al. Do nuclear magnetic resonance (NMR)-based metabolomics improve the prediction of pregnancy-related disorders? Findings from a UK birth cohort with independent validation. BMC Med. 2020;18: 366. doi:10.1186/s12916-020-01819-z

41. Pinto J, Almeida LM, Martins AS, Duarte D, Barros AS, Galhano E, et al. Prediction of gestational diabetes through NMR metabolomics of maternal blood. J Proteome Res. 2015;14: 2696–2706. doi:10.1021/acs.jproteome.5b00260

42. Mokkala K, Vahlberg T, Pellonperä O, Houttu N, Koivuniemi E, Laitinen K. Distinct Metabolic Profile in Early Pregnancy of Overweight and Obese Women Developing Gestational Diabetes. J Nutr. 2020;150: 31–37. doi:10.1093/jn/nxz220

43. Diaz SO, Pinto J, Graça G, Duarte IF, Barros AS, Galhano E, et al. Metabolic biomarkers of prenatal disorders: An exploratory NMR

metabonomics study of second trimester maternal urine and blood plasma. J Proteome Res. 2011;10: 3732–3742. doi:10.1021/pr200352m

44. Piras C, Neri I, Pintus R, Noto A, Petrella E, Monari F, et al. First trimester metabolomics 1H-NMR study of the urinary profile predicts gestational diabetes mellitus development in obese women. J Matern Neonatal Med. 2021;0: 1–9. doi:10.1080/14767058.2021.1970133

45. Beckonert O, Keun HC, Ebbels TMD, Bundy J, Holmes E, Lindon JC, et al. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. Nat Protoc. 2007;2: 2692–2703. doi:10.1038/nprot.2007.376

46. Yu XY, Song LP, Zheng HT, Wei SD, Wen XL, Huang B, et al. Association between functional genetic variants in retinoid X receptor-α/γ and the risk of gestational diabetes mellitus in a southern Chinese population. Biosci Rep. 2021;41: 1–10. doi:10.1042/BSR20211338

47. Zulueta M, Gallardo-Rincón H, Martinez-Juarez LA, Lomelin-Gascon J, Ortega-Montiel J, Montoya A, et al. Development and validation of a multivariable genotype-informed gestational diabetes prediction algorithm for clinical use in the Mexican population: insights into susceptibility mechanisms. BMJ open diabetes Res care. 2023;11. doi:10.1136/bmjdrc-2022-003046

48. Agena Bioscience I. Single Nucleotide Polymorphism Detection with the iPLEX® Assay and the MassARRAY® System. 2015. Available: http://agenabio.com/wp-content/uploads/2015/07/51-20061R1.0-iPLEX-Application-Note_WEB.pdf

49. Zhao C, Dong J, Jiang T, Shi Z, Yu B, Zhu Y, et al. Early Second-Trimester Serum MiRNA Profiling Predicts Gestational Diabetes Mellitus. PLoS One. 2011;6: e23925. doi:10.1371/journal.pone.0023925

50. Cao YL, Jia YJ, Xing BH, Shi DD, Dong XJ. Plasma microRNA-16-5p, -17-5p and -20a-5p: Novel diagnostic biomarkers for gestational diabetes mellitus. J Obstet Gynaecol Res. 2017;43: 974–981. doi:10.1111/jog.13317

51. Kroh EM, Parkin RK, Mitchell PS, Tewari M. Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). Methods. 2010;50: 298–301. doi:10.1016/j.ymeth.2010.01.032

52. Bernardes-Oliveira E, de Freitas DLD, de Morais C de LM, Cornetta M da C de M, Camargo JD de AS, de Lima KMG, et al. Spectrochemical differentiation in gestational diabetes mellitus based on attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy and multivariate analysis. Sci Rep. 2020;10: 19259. doi:10.1038/s41598-020-75539-y

4.4.7. Supporting information

Table 4.4-S1. Predictive performance of models using NIR spectral data from first trimester sera (Full, 10500-4000 cm^{-1}).

Model number	Dataset	Detail	Pretreatment	Width	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
1	NIR1T	Full	MC		0.6185	0.0459	0.3253	0.1091	0.4719	0.0592
2			N+MC		0.6490	0.0445	0.4347	0.0886	0.5418	0.0496
3			WLS+MC		0.7143	0.0372	0.3013	0.0906	0.5078	0.0490
4			WLS+N+MC		0.6970	0.0314	0.3093	0.0891	0.5032	0.0472
5			SNV + MC		0.6749	0.0396	0.3493	0.0948	0.5121	0.0514
6			SNV + WLS + MC		0.7000	0.0433	0.1627	0.0810	0.4313	0.0459
7			SNV + N + MC		0.6803	0.0448	0.3333	0.0808	0.5068	0.0462
8			SNV + WLS + N + MC		0.6970	0.0314	0.3093	0.0891	0.5032	0.0472
9			1D + MC	3	0.8212	0.0262	0.0827	0.0370	0.4519	0.0227
10				7	0.8140	0.0222	0.1080	0.0444	0.4610	0.0248
11				11	0.8006	0.0326	0.1213	0.0440	0.4610	0.0274
12				15	0.7266	0.0317	0.1760	0.0657	0.4513	0.0364
13				19	0.7036	0.0471	0.1787	0.0767	0.4411	0.0450
14				23	0.6519	0.0528	0.2133	0.0738	0.4326	0.0454
15			1D + N + MC	3	0.8239	0.0268	0.0720	0.0183	0.4479	0.0162
16				7	0.8236	0.0287	0.1120	0.0475	0.4678	0.0278
17				11	0.7824	0.0333	0.1027	0.0430	0.4425	0.0272
18				15	0.7481	0.0310	0.1640	0.0649	0.4560	0.0360
19				19	0.7296	0.0404	0.1680	0.0777	0.4488	0.0438
20				23	0.6782	0.0529	0.1960	0.0743	0.4371	0.0456
21			2D + MC	3	0.7854	0.0263	0.1147	0.0405	0.4500	0.0241
22				7	0.7821	0.0252	0.0400	0.0426	0.4110	0.0248

23		11	0.8027	0.0277	0.1000	0.0431	0.4513	0.0256
24		15	0.8131	0.0294	0.1440	0.0492	0.4786	0.0287
25		19	0.7463	0.0321	0.1347	0.0436	0.4405	0.0271
26		23	0.7579	0.0291	0.1213	0.0613	0.4396	0.0339
27	2D + N + MC	3	0.7860	0.0296	0.1040	0.0385	0.4450	0.0243
28		7	0.7737	0.0352	0.0333	0.0387	0.4035	0.0262
29		11	0.8227	0.0290	0.1160	0.0568	0.4693	0.0319
30		15	0.8224	0.0194	0.1560	0.0439	0.4892	0.0240
31		19	0.7484	0.0302	0.1293	0.0283	0.4388	0.0207
32		23	0.7582	0.0316	0.1307	0.0485	0.4444	0.0289
33	SM + MC	3	0.6158	0.0449	0.3267	0.0964	0.4712	0.0532
34		7	0.6143	0.0430	0.3400	0.0992	0.4772	0.0541
35		11	0.6131	0.0462	0.3360	0.1017	0.4746	0.0558
36		15	0.6322	0.0448	0.2880	0.1154	0.4601	0.0619
37		19	0.6278	0.0454	0.2960	0.0925	0.4619	0.0515
38		23	0.6239	0.0522	0.3000	0.0955	0.4619	0.0544
39	SM + SNV + MC	3	0.6934	0.0454	0.3453	0.0669	0.5194	0.0404
40		7	0.6946	0.0503	0.3187	0.0789	0.5066	0.0468
41		11	0.6964	0.0404	0.3320	0.0731	0.5142	0.0418
42		15	0.7179	0.0477	0.3373	0.0692	0.5276	0.0420
43		19	0.7197	0.0504	0.3480	0.0811	0.5339	0.0478
44		23	0.7072	0.0529	0.3387	0.0784	0.5229	0.0473
45	SM + SNV + WLS + MC	3	0.7012	0.0554	0.1667	0.0635	0.4339	0.0421
46		7	0.7322	0.0435	0.2000	0.0700	0.4661	0.0412
47		11	0.7439	0.0485	0.1947	0.0806	0.4693	0.0470
48		15	0.7528	0.0411	0.2453	0.0824	0.4991	0.0461
49		19	0.7540	0.0409	0.2373	0.0605	0.4957	0.0365
50		23	0.7645	0.0421	0.2227	0.0733	0.4936	0.0423
51	SM + SNV + N + MC	3	0.6815	0.0458	0.3307	0.0785	0.5061	0.0454

52		7	0.7060	0.0551	0.3387	0.0748	0.5223	0.0465
53		11	0.7051	0.0524	0.3293	0.0779	0.5172	0.0469
54		15	0.7030	0.0461	0.3480	0.0865	0.5255	0.0490
55		19	0.7101	0.0507	0.3520	0.0700	0.5311	0.0432
56		23	0.7248	0.0468	0.3400	0.0777	0.5324	0.0453
57	SM + SNV + WLS + N + MC	3	0.6934	0.0386	0.2973	0.0906	0.4954	0.0492
58		7	0.6958	0.0382	0.3027	0.1001	0.4992	0.0536
59		11	0.6964	0.0385	0.2853	0.0884	0.4909	0.0482
60		15	0.7099	0.0422	0.2573	0.0863	0.4836	0.0480
61		19	0.7152	0.0420	0.2600	0.0945	0.4876	0.0517
62		23	0.7119	0.0416	0.2760	0.0863	0.4940	0.0479
63	SM + WLS + MC	3	0.7131	0.0391	0.3200	0.0774	0.5166	0.0433
64		7	0.7200	0.0347	0.2947	0.0904	0.5073	0.0484
65		11	0.7164	0.0378	0.2853	0.0894	0.5009	0.0485
66		15	0.7182	0.0433	0.2533	0.0893	0.4858	0.0497
67		19	0.7322	0.0491	0.2493	0.0979	0.4908	0.0548
68		23	0.7278	0.0445	0.2573	0.0785	0.4925	0.0451
69	SM + WLS + N + MC	3	0.6976	0.0408	0.3227	0.1038	0.5101	0.0557
70		7	0.6875	0.0439	0.3093	0.0997	0.4984	0.0545
71		11	0.6967	0.0343	0.3160	0.0795	0.5064	0.0433
72		15	0.7122	0.0421	0.2680	0.0791	0.4901	0.0448
73		19	0.7158	0.0432	0.2733	0.0875	0.4946	0.0488
74		23	0.7045	0.0412	0.2773	0.0705	0.4909	0.0408
75	SM + N + MC	3	0.6451	0.0408	0.4440	0.0826	0.5445	0.0461
76		7	0.6681	0.0488	0.4333	0.0875	0.5507	0.0501
77		11	0.6824	0.0443	0.4427	0.0747	0.5625	0.0434
78		15	0.6687	0.0559	0.4347	0.0765	0.5517	0.0474
79		19	0.6955	0.0461	0.4440	0.0733	0.5698	0.0433
80		23	0.6946	0.0456	0.4507	0.0681	0.5726	0.0410

NIR: near-infrared; 1T: first trimester; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation; MC: mean centering; N: normalization; WLS: weighted least squares; SNV: standard normal variate; 1D: first derivative; 2D: second derivative; SM: smoothing.

Table 4.4-S2. Predictive performance of models using NIR spectral data from first trimester sera (range 1, 10500-7600 cm^{-1}).

Model number	Dataset	Detail	Pretreatment	Width	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
1	NIR1T	R1	MC		0.6245	0.0328	0.4320	0.1010	0.5282	0.0531
2			N+MC		0.6722	0.0361	0.5920	0.0910	0.6321	0.0489
3			WLS+MC		0.7373	0.0549	0.1480	0.0753	0.4427	0.0466
4			WLS+N+MC		0.7499	0.0395	0.3067	0.0883	0.5283	0.0484
5			SNV + MC		0.6800	0.0406	0.4133	0.0738	0.5467	0.0421
6			SNV + WLS + MC		0.7579	0.0458	0.1320	0.0768	0.4450	0.0447
7			SNV + N + MC		0.6764	0.0346	0.4120	0.0793	0.5442	0.0433
8			SNV + WLS + N + MC		0.7433	0.0325	0.3840	0.0732	0.5636	0.0401
9			1D + MC	3	0.8146	0.0272	0.0987	0.0387	0.4566	0.0236
10				7	0.8239	0.0289	0.1147	0.0382	0.4693	0.0240
11				11	0.7857	0.0306	0.1240	0.0426	0.4548	0.0262
12				15	0.7203	0.0274	0.1987	0.0579	0.4595	0.0320
13				19	0.7110	0.0437	0.2120	0.0816	0.4615	0.0463
14				23	0.6555	0.0543	0.2227	0.0837	0.4391	0.0499
15			1D + N + MC	3	0.8358	0.0273	0.0907	0.0399	0.4632	0.0242
16				7	0.8352	0.0226	0.1160	0.0400	0.4756	0.0230
17				11	0.7839	0.0282	0.1267	0.0387	0.4553	0.0239
18				15	0.7439	0.0355	0.2093	0.0646	0.4766	0.0369
19				19	0.7269	0.0372	0.2067	0.0765	0.4668	0.0425
20				23	0.6493	0.0518	0.2253	0.0736	0.4373	0.0450
21			2D + MC	3	0.7985	0.0214	0.1173	0.0394	0.4579	0.0224
22				7	0.7761	0.0305	0.0653	0.0369	0.4207	0.0239
23				11	0.8030	0.0250	0.1120	0.0546	0.4575	0.0300
24				15	0.8158	0.0274	0.1520	0.0357	0.4839	0.0225
25				19	0.7403	0.0362	0.1427	0.0426	0.4415	0.0280
26				23	0.7612	0.0342	0.1600	0.0602	0.4606	0.0346

27	2D + N + MC	3	0.7872	0.0252	0.1000	0.0387	0.4436	0.0231
28		7	0.7872	0.0300	0.0453	0.0367	0.4162	0.0237
29		11	0.8257	0.0236	0.1120	0.0475	0.4688	0.0265
30		15	0.8278	0.0270	0.1600	0.0467	0.4939	0.0270
31		19	0.7460	0.0299	0.1333	0.0269	0.4397	0.0201
32		23	0.7627	0.0306	0.1387	0.0536	0.4507	0.0309
33	SM + MC	3	0.6260	0.0423	0.3933	0.0936	0.5097	0.0513
34		7	0.6164	0.0384	0.4147	0.1153	0.5155	0.0608
35		11	0.6301	0.0381	0.3920	0.1132	0.5111	0.0597
36		15	0.6251	0.0314	0.4067	0.0983	0.5159	0.0516
37		19	0.6257	0.0326	0.4213	0.1047	0.5235	0.0548
38		23	0.6218	0.0395	0.3747	0.1142	0.4982	0.0604
39	SM + SNV + MC	3	0.6943	0.0343	0.4253	0.0736	0.5598	0.0406
40		7	0.6851	0.0356	0.4253	0.0818	0.5552	0.0446
41		11	0.6797	0.0388	0.4027	0.0785	0.5412	0.0438
42		15	0.6946	0.0507	0.3653	0.0822	0.5300	0.0483
43		19	0.7000	0.0495	0.3613	0.0842	0.5307	0.0488
44		23	0.7099	0.0460	0.3453	0.0920	0.5276	0.0515
45	SM + SNV + WLS + MC	3	0.7522	0.0577	0.1347	0.0966	0.4435	0.0563
46		7	0.7609	0.0442	0.1640	0.0788	0.4624	0.0452
47		11	0.7779	0.0403	0.1627	0.0740	0.4703	0.0421
48		15	0.7872	0.0405	0.1760	0.0722	0.4816	0.0414
49		19	0.7961	0.0426	0.1893	0.0651	0.4927	0.0389
50		23	0.7878	0.0382	0.2147	0.0811	0.5012	0.0448
51	SM + SNV + N + MC	3	0.6872	0.0395	0.4173	0.0891	0.5522	0.0488
52		7	0.6839	0.0399	0.4160	0.0814	0.5499	0.0453
53		11	0.6890	0.0423	0.3987	0.0731	0.5438	0.0422
54		15	0.7045	0.0426	0.3507	0.0912	0.5276	0.0503
55		19	0.7152	0.0474	0.3427	0.0713	0.5289	0.0428

56		23	0.7087	0.0542	0.3693	0.0764	0.5390	0.0469
57	SM + SNV + WLS + N + MC	3	0.7543	0.0371	0.3613	0.0751	0.5578	0.0419
58		7	0.7469	0.0344	0.3453	0.0870	0.5461	0.0468
59		11	0.7597	0.0405	0.3520	0.1008	0.5559	0.0543
60		15	0.7481	0.0402	0.3253	0.0968	0.5367	0.0524
61		19	0.7522	0.0445	0.2800	0.0873	0.5161	0.0490
62		23	0.7549	0.0385	0.2507	0.0889	0.5028	0.0484
63	SM + WLS + MC	3	0.7475	0.0633	0.1320	0.0731	0.4397	0.0484
64		7	0.7349	0.0725	0.1360	0.0785	0.4355	0.0534
65		11	0.7699	0.0533	0.1347	0.0731	0.4523	0.0452
66		15	0.7678	0.0579	0.1613	0.0809	0.4645	0.0497
67		19	0.8000	0.0431	0.1867	0.0725	0.4933	0.0422
68		23	0.7979	0.0505	0.1747	0.0818	0.4863	0.0481
69	SM + WLS + N + MC	3	0.7522	0.0295	0.3653	0.0606	0.5588	0.0337
70		7	0.7433	0.0361	0.3453	0.0920	0.5443	0.0494
71		11	0.7501	0.0374	0.3387	0.0902	0.5444	0.0488
72		15	0.7561	0.0361	0.3093	0.0911	0.5327	0.0490
73		19	0.7499	0.0410	0.2840	0.0941	0.5169	0.0513
74		23	0.7460	0.0414	0.2813	0.0822	0.5137	0.0460
75	SM + N + MC	3	0.6752	0.0341	0.5653	0.1176	0.6203	0.0612
76		7	0.6663	0.0382	0.5840	0.1048	0.6251	0.0558
77		11	0.6746	0.0415	0.5293	0.0937	0.6020	0.0512
78		15	0.6761	0.0375	0.5480	0.1019	0.6121	0.0543
79		19	0.6842	0.0431	0.5040	0.0935	0.5941	0.0515
80		23	0.6988	0.0455	0.4987	0.0854	0.5987	0.0484

NIR: near-infrared; 1T: first trimester; R1: range 1; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation; MC: mean centering; N: normalization; WLS: weighted least squares; SNV: standard normal variate; 1D: first derivative; 2D: second derivative; SM: smoothing.

Table 4.4-S3. Predictive performance of models using NIR spectral data from first trimester sera (range 2, 7600-5100 cm^{-1}).

Model number	Dataset	Detail	Pretreatment	Width	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
1	NIR1T	R2	MC		0.6063	0.0364	0.5240	0.0819	0.5651	0.0448
2			N+MC		0.5725	0.0312	0.6120	0.1083	0.5923	0.0564
3			WLS+MC		0.6872	0.0402	0.3213	0.0770	0.5042	0.0434
4			WLS+N+MC		0.6675	0.0638	0.1493	0.0780	0.4084	0.0504
5			SNV + MC		0.7313	0.0280	0.4187	0.0674	0.5750	0.0365
6			SNV + WLS + MC		0.6812	0.0424	0.3213	0.0880	0.5013	0.0488
7			SNV + N + MC		0.7293	0.0353	0.4293	0.0752	0.5793	0.0415
8			SNV + WLS + N + MC		0.7015	0.0439	0.3320	0.0994	0.5167	0.0543
9			1D + MC	3	0.6027	0.0403	0.2800	0.1387	0.4413	0.0722
10				7	0.7069	0.0507	0.0920	0.0630	0.3994	0.0404
11				11	0.6666	0.0479	0.1227	0.0651	0.3946	0.0404
12				15	0.6791	0.0377	0.1680	0.0765	0.4236	0.0426
13				19	0.6776	0.0429	0.1800	0.0716	0.4288	0.0417
14				23	0.6812	0.0435	0.1893	0.0801	0.4353	0.0456
15			1D + N + MC	3	0.6188	0.0354	0.3573	0.1304	0.4881	0.0676
16				7	0.6107	0.0650	0.1893	0.0917	0.4000	0.0562
17				11	0.6460	0.0566	0.1387	0.0748	0.3923	0.0469
18				15	0.6430	0.0613	0.1413	0.0682	0.3922	0.0458
19				19	0.6301	0.0456	0.1480	0.0822	0.3891	0.0470
20				23	0.6460	0.0612	0.1853	0.0592	0.4157	0.0426
21			2D + MC	3	0.5331	0.0186	0.6507	0.0825	0.5919	0.0423
22				7	0.6767	0.0597	0.1573	0.0911	0.4170	0.0545
23				11	0.7185	0.0650	0.1093	0.0599	0.4139	0.0442
24				15	0.5722	0.0758	0.1013	0.0833	0.3368	0.0563
25				19	0.6272	0.0664	0.0693	0.0659	0.3482	0.0468
26				23	0.6463	0.0615	0.1093	0.0817	0.3778	0.0511

27	2D + N + MC	3	0.7433	0.0394	0.1840	0.1074	0.4636	0.0572
28		7	0.7006	0.0649	0.1107	0.0654	0.4056	0.0461
29		11	0.6797	0.0552	0.1427	0.0797	0.4112	0.0485
30		15	0.6209	0.0346	0.3027	0.0982	0.4618	0.0521
31		19	0.6391	0.0434	0.2147	0.0974	0.4269	0.0533
32		23	0.6200	0.0659	0.0893	0.0837	0.3547	0.0533
33	SM + MC	3	0.6242	0.0317	0.5253	0.0746	0.5748	0.0405
34		7	0.6134	0.0268	0.5267	0.0854	0.5700	0.0448
35		11	0.6236	0.0277	0.5040	0.0832	0.5638	0.0438
36		15	0.6209	0.0426	0.5333	0.0913	0.5771	0.0504
37		19	0.6191	0.0361	0.5347	0.0878	0.5769	0.0475
38		23	0.6146	0.0310	0.5267	0.0800	0.5706	0.0429
39	SM + SNV + MC	3	0.7218	0.0306	0.4293	0.0689	0.5756	0.0377
40		7	0.7352	0.0327	0.4480	0.0603	0.5916	0.0343
41		11	0.7296	0.0279	0.4440	0.0640	0.5868	0.0349
42		15	0.7158	0.0374	0.4293	0.0787	0.5726	0.0436
43		19	0.7352	0.0283	0.4493	0.0657	0.5923	0.0358
44		23	0.7263	0.0339	0.4533	0.0673	0.5898	0.0377
45	SM + SNV + WLS + MC	3	0.6818	0.0384	0.3440	0.0730	0.5129	0.0412
46		7	0.6731	0.0461	0.2853	0.0831	0.4792	0.0475
47		11	0.6779	0.0485	0.3080	0.0872	0.4930	0.0499
48		15	0.6848	0.0447	0.2880	0.0947	0.4864	0.0524
49		19	0.6854	0.0460	0.3067	0.0893	0.4960	0.0502
50		23	0.6851	0.0488	0.2720	0.0658	0.4785	0.0410
51	SM + SNV + N + MC	3	0.7310	0.0323	0.4213	0.0743	0.5762	0.0405
52		7	0.7310	0.0325	0.4360	0.0663	0.5835	0.0369
53		11	0.7301	0.0313	0.4413	0.0570	0.5857	0.0325
54		15	0.7197	0.0277	0.4440	0.0781	0.5819	0.0414
55		19	0.7230	0.0384	0.4373	0.0648	0.5802	0.0377

56		23	0.7191	0.0337	0.4360	0.0728	0.5776	0.0401
57	SM + SNV + WLS + N + MC	3	0.6937	0.0420	0.3160	0.0931	0.5049	0.0511
58		7	0.6958	0.0395	0.3373	0.0947	0.5166	0.0513
59		11	0.6943	0.0410	0.3133	0.0777	0.5038	0.0439
60		15	0.6949	0.0424	0.3467	0.0883	0.5208	0.0490
61		19	0.6922	0.0387	0.3427	0.0883	0.5175	0.0482
62		23	0.6890	0.0310	0.3827	0.0969	0.5358	0.0509
63	SM + WLS + MC	3	0.6967	0.0331	0.3347	0.0836	0.5157	0.0449
64		7	0.6815	0.0383	0.3373	0.0730	0.5094	0.0412
65		11	0.6878	0.0436	0.3053	0.0884	0.4965	0.0493
66		15	0.6985	0.0402	0.3187	0.0906	0.5086	0.0496
67		19	0.6910	0.0473	0.3160	0.0902	0.5035	0.0509
68		23	0.6928	0.0400	0.3347	0.0857	0.5138	0.0473
69	SM + WLS + N + MC	3	0.7015	0.0399	0.3373	0.0743	0.5194	0.0422
70		7	0.7012	0.0375	0.3227	0.0927	0.5119	0.0500
71		11	0.6982	0.0495	0.3147	0.0797	0.5064	0.0469
72		15	0.6991	0.0361	0.3373	0.0743	0.5182	0.0413
73		19	0.6940	0.0348	0.3387	0.0784	0.5163	0.0429
74		23	0.6913	0.0412	0.3627	0.1079	0.5270	0.0577
75	SM + N + MC	3	0.5678	0.0322	0.6480	0.1035	0.6079	0.0542
76		7	0.5621	0.0406	0.6213	0.0985	0.5917	0.0533
77		11	0.5675	0.0269	0.6320	0.1028	0.5997	0.0531
78		15	0.5612	0.0273	0.6333	0.1071	0.5973	0.0553
79		19	0.5728	0.0330	0.6293	0.0991	0.6011	0.0522
80		23	0.5633	0.0322	0.6373	0.1045	0.6003	0.0547

NIR: near-infrared; 1T: first trimester; R2: range 2; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation; MC: mean centering; N: normalization; WLS: weighted least squares; SNV: standard normal variate; 1D: first derivative; 2D: second derivative; SM: smoothing.

Table 4.4-S4. Predictive performance of models using NIR spectral data from first trimester sera (range 3, 5100-4000 cm^{-1}).

Model number	Dataset	Detail	Pretreatment	Width	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
1	NIR1T	R3	MC		0.5961	0.0295	0.4907	0.0870	0.5434	0.0459
2			N+MC		0.6167	0.0565	0.1387	0.0671	0.3777	0.0439
3			WLS+MC		0.6051	0.0452	0.1627	0.0727	0.3839	0.0428
4			WLS+N+MC		0.6606	0.0307	0.2733	0.0896	0.4670	0.0474
5			SNV + MC		0.6278	0.0351	0.1400	0.0844	0.3839	0.0457
6			SNV + WLS + MC		0.6140	0.0417	0.1200	0.0883	0.3670	0.0488
7			SNV + N + MC		0.6260	0.0407	0.1307	0.0749	0.3783	0.0426
8			SNV + WLS + N + MC		0.6081	0.0519	0.1787	0.0888	0.3934	0.0514
9			1D + MC	3	0.7137	0.0375	0.1147	0.0539	0.4142	0.0328
10				7	0.6394	0.0467	0.1333	0.0797	0.3864	0.0462
11				11	0.6340	0.0407	0.1200	0.0660	0.3770	0.0388
12				15	0.6185	0.0348	0.1240	0.0841	0.3713	0.0455
13				19	0.6310	0.0374	0.1520	0.0884	0.3915	0.0480
14				23	0.6391	0.0401	0.1893	0.0717	0.4142	0.0411
15			1D + N + MC	3	0.7188	0.0339	0.1333	0.0673	0.4261	0.0377
16				7	0.6588	0.0363	0.1400	0.0777	0.3994	0.0429
17				11	0.6439	0.0334	0.1333	0.0785	0.3886	0.0427
18				15	0.6484	0.0413	0.1093	0.0931	0.3788	0.0509
19				19	0.6472	0.0386	0.1747	0.0807	0.4109	0.0447
20				23	0.6337	0.0434	0.1573	0.0828	0.3955	0.0467
21			2D + MC	3	0.7493	0.0261	0.0533	0.0426	0.4013	0.0250
22				7	0.7388	0.0309	0.1240	0.0660	0.4314	0.0364
23				11	0.7445	0.0321	0.1440	0.0412	0.4442	0.0261
24				15	0.7272	0.0353	0.1947	0.0585	0.4609	0.0341
25				19	0.6734	0.0320	0.1600	0.0750	0.4167	0.0408
26				23	0.6672	0.0451	0.1947	0.0658	0.4309	0.0399

27	2D + N + MC	3	0.7591	0.0292	0.0467	0.0431	0.4029	0.0260
28		7	0.7436	0.0264	0.1040	0.0489	0.4238	0.0278
29		11	0.7385	0.0251	0.1453	0.0480	0.4419	0.0271
30		15	0.7104	0.0350	0.1853	0.0729	0.4479	0.0404
31		19	0.6857	0.0320	0.2187	0.0762	0.4522	0.0413
32		23	0.6743	0.0399	0.1947	0.0772	0.4345	0.0434
33	SM + MC	3	0.6018	0.0331	0.4987	0.0677	0.5502	0.0377
34		7	0.5878	0.0340	0.4813	0.0886	0.5345	0.0475
35		11	0.5791	0.0411	0.4773	0.1105	0.5282	0.0590
36		15	0.5866	0.0357	0.5147	0.0809	0.5506	0.0442
37		19	0.5955	0.0358	0.4973	0.0788	0.5464	0.0433
38		23	0.5931	0.0346	0.5133	0.0811	0.5532	0.0441
39	SM + SNV + MC	3	0.6242	0.0356	0.1573	0.0697	0.3908	0.0391
40		7	0.6272	0.0362	0.1600	0.0797	0.3936	0.0438
41		11	0.6272	0.0317	0.1173	0.0720	0.3722	0.0393
42		15	0.6254	0.0377	0.1413	0.0826	0.3834	0.0454
43		19	0.6230	0.0395	0.1347	0.0802	0.3788	0.0447
44		23	0.6251	0.0451	0.1467	0.0841	0.3859	0.0477
45	SM + SNV + WLS + MC	3	0.6233	0.0356	0.1627	0.0854	0.3930	0.0462
46		7	0.6173	0.0441	0.1467	0.0713	0.3820	0.0419
47		11	0.6084	0.0445	0.1627	0.0821	0.3855	0.0467
48		15	0.6242	0.0399	0.1387	0.0671	0.3814	0.0390
49		19	0.6266	0.0437	0.1693	0.0885	0.3980	0.0494
50		23	0.6128	0.0456	0.1893	0.0834	0.4011	0.0475
51	SM + SNV + N + MC	3	0.6322	0.0388	0.1573	0.0722	0.3948	0.0410
52		7	0.6101	0.0462	0.1360	0.0712	0.3731	0.0424
53		11	0.6310	0.0371	0.1427	0.0738	0.3869	0.0413
54		15	0.6072	0.0410	0.1427	0.0725	0.3749	0.0417
55		19	0.6221	0.0447	0.1587	0.0761	0.3904	0.0441

56		23	0.6119	0.0407	0.1440	0.0665	0.3780	0.0390
57	SM + SNV + WLS + N + MC	3	0.6173	0.0453	0.1760	0.0828	0.3967	0.0472
58		7	0.6152	0.0487	0.1853	0.0916	0.4003	0.0519
59		11	0.6051	0.0386	0.1800	0.0677	0.3925	0.0390
60		15	0.6063	0.0422	0.2093	0.0852	0.4078	0.0475
61		19	0.6203	0.0401	0.1827	0.0735	0.4015	0.0419
62		23	0.6164	0.0442	0.1813	0.0884	0.3989	0.0494
63	SM + WLS + MC	3	0.5848	0.0552	0.1493	0.0640	0.3671	0.0422
64		7	0.5958	0.0569	0.1653	0.0788	0.3806	0.0486
65		11	0.6036	0.0414	0.1867	0.0738	0.3951	0.0423
66		15	0.6039	0.0447	0.1880	0.0849	0.3959	0.0480
67		19	0.6131	0.0360	0.1933	0.0777	0.4032	0.0428
68		23	0.5961	0.0405	0.2000	0.0913	0.3981	0.0500
69	SM + WLS + N + MC	3	0.6191	0.0383	0.1600	0.0862	0.3896	0.0472
70		7	0.6176	0.0444	0.1853	0.0833	0.4015	0.0472
71		11	0.6075	0.0436	0.1973	0.0737	0.4024	0.0428
72		15	0.6167	0.0476	0.1920	0.0890	0.4044	0.0505
73		19	0.6182	0.0357	0.1880	0.0900	0.4031	0.0484
74		23	0.6221	0.0301	0.2013	0.0928	0.4117	0.0488
75	SM + N + MC	3	0.6179	0.0464	0.1493	0.0757	0.3836	0.0444
76		7	0.6152	0.0482	0.1560	0.0733	0.3856	0.0439
77		11	0.6134	0.0657	0.1760	0.0860	0.3947	0.0541
78		15	0.6093	0.0540	0.1733	0.0830	0.3913	0.0495
79		19	0.6245	0.0436	0.1440	0.0897	0.3842	0.0499
80		23	0.6078	0.0417	0.1947	0.0970	0.4012	0.0528

NIR: near-infrared; 1T: first trimester; R3: range 3; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation; MC: mean centering; N: normalization; WLS: weighted least squares; SNV: standard normal variate; 1D: first derivative; 2D: second derivative; SM: smoothing.

Table 4.4-S5. Predictive performance of models combining NIR spectral data of first trimester sera with medical parameters.

NIR 1T range[a]	Plus	Block order[b]	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
Full			0.6946	0.0456	0.4507	0.0681	0.5726	0.0410
Full	Medical	1-2	0.7355	0.0479	0.3147	0.0863	0.5251	0.0493
		2-1	0.7773	0.0432	0.3360	0.0883	0.5567	0.0491
Full simp[c]	Medical simp[c]	1-2	0.7525	0.0564	0.4227	0.1171	0.5876	0.0650
		2-1	0.7728	0.0702	0.4453	0.0947	0.6091	0.0589
R1			0.6722	0.0361	0.5920	0.0910	0.6321	0.0489
R1	Medical	1-2	0.7048	0.0425	0.3827	0.1059	0.5437	0.0570
		2-1	0.7227	0.0483	0.3827	0.0931	0.5527	0.0525
R1 simp[c]	Medical simp[c]	1-2	0.7328	0.0682	0.5027	0.1251	0.6178	0.0712
		2-1	0.7513	0.0560	0.5013	0.1080	0.6263[d]	0.0608[d]

[a] Full: 10500-4000 cm^{-1}; R1: 10500-7600 cm^{-1}.

[b] NIR blocks were pretreated using the conditions identified as best in single-block analysis. Medical blocks were pretreated by autoscaling.

[c] Data obtained after variable selection, using the variable importance information from the original equivalent multi-block model.

[d] Best multi-block model. The NIR block consists of 41% of the original R1 variables (591 wavenumbers out of 1450). The medical block consists of 5 variables, history of GDM in a prior pregnancy, family history of DM, BMI, family history of hyperthyroidism and first period age. Its DCV-AUC-ROC is 0.5219 ± 0.0922.

NIR: near-infrared; 1T: first trimester; R1: range 1; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation.

Table 4.4-S6. Predictive performance of models using NIR spectral data from second trimester sera (Full, 10500-4000 cm^{-1}).

Model number	Dataset	Detail	Pretreatment	Width	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
1	NIR2T	Full	MC		0.6528	0.1013	0.0425	0.0598	0.3477	0.0588
2			N+MC		0.6590	0.0645	0.0475	0.0709	0.3532	0.0480
3			WLS+MC		0.5769	0.0498	0.1625	0.1018	0.3697	0.0567
4			WLS+N+MC		0.5210	0.0927	0.1000	0.1211	0.3105	0.0763
5			SNV + MC		0.5738	0.0538	0.2150	0.1337	0.3944	0.0721
6			SNV + WLS + MC		0.5067	0.0868	0.0775	0.1007	0.2921	0.0665
7			SNV + N + MC		0.5728	0.0557	0.1900	0.1191	0.3814	0.0657
8			SNV + WLS + N + MC		0.5210	0.0776	0.0600	0.0808	0.2905	0.0560
9			1D + MC	3	0.7908	0.0317	0.0000	0.0000	0.3954	0.0158
10				7	0.7436	0.0377	0.0400	0.0589	0.3918	0.0350
11				11	0.7651	0.0312	0.0525	0.0623	0.4088	0.0349
12				15	0.7615	0.0329	0.0600	0.0631	0.4108	0.0356
13				19	0.7815	0.0360	0.0675	0.0629	0.4245	0.0362
14				23	0.8149	0.0389	0.0800	0.0606	0.4474	0.0360
15			1D + N + MC	3	0.7974	0.0285	0.0000	0.0000	0.3987	0.0142
16				7	0.7728	0.0284	0.0725	0.0623	0.4227	0.0342
17				11	0.7733	0.0333	0.0900	0.0567	0.4317	0.0329
18				15	0.7708	0.0368	0.0525	0.0623	0.4116	0.0362
19				19	0.7559	0.0388	0.0100	0.0343	0.3829	0.0259
20				23	0.7826	0.0422	0.0450	0.0606	0.4138	0.0369
21			2D + MC	3	0.8677	0.0321	0.0050	0.0247	0.4363	0.0203
22				7	0.7487	0.0284	0.0000	0.0000	0.3744	0.0142
23				11	0.7267	0.0330	0.0000	0.0000	0.3633	0.0165
24				15	0.8036	0.0386	0.0700	0.0676	0.4368	0.0389
25				19	0.7836	0.0328	0.0750	0.0619	0.4293	0.0350
26				23	0.7687	0.0338	0.0425	0.0598	0.4056	0.0343

27	2D + N + MC	3	0.8815	0.0219	0.0000	0.0000	0.4408	0.0110
28		7	0.7559	0.0333	0.0000	0.0000	0.3779	0.0166
29		11	0.7333	0.0332	0.0000	0.0000	0.3667	0.0166
30		15	0.8133	0.0324	0.1150	0.0556	0.4642	0.0321
31		19	0.7887	0.0361	0.0800	0.0606	0.4344	0.0353
32		23	0.7600	0.0376	0.0325	0.0554	0.3963	0.0335
33	SM + MC	3	0.6390	0.0766	0.0425	0.0649	0.3407	0.0502
34		7	0.6415	0.0831	0.0325	0.0554	0.3370	0.0499
35		11	0.6641	0.0694	0.0300	0.0595	0.3471	0.0457
36		15	0.6708	0.0642	0.0450	0.0657	0.3579	0.0459
37		19	0.6846	0.0798	0.0650	0.0808	0.3748	0.0568
38		23	0.6764	0.0667	0.0550	0.0805	0.3657	0.0523
39	SM + SNV + MC	3	0.5718	0.0668	0.2100	0.1055	0.3909	0.0624
40		7	0.5718	0.0467	0.1825	0.1190	0.3771	0.0639
41		11	0.5872	0.0554	0.1475	0.1000	0.3673	0.0572
42		15	0.6000	0.0515	0.1250	0.1101	0.3625	0.0608
43		19	0.6190	0.0720	0.1100	0.1175	0.3645	0.0689
44		23	0.6123	0.0694	0.1500	0.1237	0.3812	0.0709
45	SM + SNV + WLS + MC	3	0.5015	0.0905	0.0700	0.0881	0.2858	0.0631
46		7	0.4903	0.0926	0.0825	0.0897	0.2864	0.0645
47		11	0.5374	0.0804	0.0525	0.0841	0.2950	0.0582
48		15	0.5246	0.0762	0.0675	0.0918	0.2961	0.0596
49		19	0.5472	0.0812	0.0550	0.0721	0.3011	0.0543
50		23	0.5308	0.0887	0.0425	0.0598	0.2866	0.0535
51	SM + SNV + N + MC	3	0.5728	0.0557	0.1900	0.1191	0.3814	0.0657
52		7	0.5713	0.0486	0.1725	0.1208	0.3719	0.0651
53		11	0.5821	0.0513	0.1450	0.1140	0.3635	0.0625
54		15	0.6103	0.0695	0.1300	0.1100	0.3701	0.0650
55		19	0.6164	0.0604	0.1100	0.1119	0.3632	0.0636

56		23	0.6318	0.0706	0.0950	0.0781	0.3634	0.0526
57	SM + SNV + WLS + N + MC	3	0.4995	0.0848	0.0800	0.1003	0.2897	0.0657
58		7	0.5149	0.0766	0.0750	0.0875	0.2949	0.0581
59		11	0.5410	0.1073	0.0550	0.0844	0.2980	0.0683
60		15	0.5072	0.0953	0.0725	0.0841	0.2898	0.0635
61		19	0.5200	0.0770	0.0675	0.0766	0.2938	0.0543
62		23	0.5267	0.0875	0.0850	0.0960	0.3058	0.0649
63	SM + WLS + MC	3	0.5846	0.0647	0.1500	0.1071	0.3673	0.0626
64		7	0.5687	0.0673	0.1575	0.0938	0.3631	0.0577
65		11	0.5959	0.0537	0.1750	0.1101	0.3854	0.0612
66		15	0.5692	0.0667	0.1550	0.0996	0.3621	0.0600
67		19	0.5759	0.0616	0.1400	0.0933	0.3579	0.0559
68		23	0.5769	0.0458	0.1450	0.0923	0.3610	0.0515
69	SM + WLS + N + MC	3	0.5087	0.0820	0.1025	0.0968	0.3056	0.0634
70		7	0.5267	0.1044	0.0925	0.0972	0.3096	0.0713
71		11	0.5359	0.1001	0.0650	0.0919	0.3004	0.0679
72		15	0.5159	0.0860	0.0775	0.0833	0.2967	0.0599
73		19	0.5344	0.0810	0.0850	0.0926	0.3097	0.0615
74		23	0.5190	0.0978	0.0775	0.0974	0.2982	0.0690
75	SM + N + MC	3	0.6508	0.0729	0.0925	0.1066	0.3716	0.0646
76		7	0.6605	0.0808	0.0800	0.0866	0.3703	0.0592
77		11	0.6764	0.0752	0.0750	0.0911	0.3757	0.0591
78		15	0.6974	0.0705	0.0925	0.0904	0.3950	0.0573
79		19	0.7236	0.0744	0.1125	0.0809	0.4180	0.0549
80		23	0.7456	0.0655	0.1125	0.1079	0.4291	0.0631

NIR: near-infrared; 2T: second trimester; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation; MC: mean centering; N: normalization; WLS: weighted least squares; SNV: standard normal variate; 1D: first derivative; 2D: second derivative; SM: smoothing.

Table 4.4-S7. Predictive performance of models using NIR spectral data from second trimester sera (range 1, 10500-7600 cm^{-1}).

Model number	Dataset	Detail	Pretreatment	Width	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
1	NIR2T	R1	MC		0.6682	0.0916	0.0325	0.0609	0.3504	0.0550
2			N+MC		0.6267	0.0714	0.0250	0.0505	0.3258	0.0437
3			WLS+MC		0.7918	0.0451	0.0200	0.0527	0.4059	0.0347
4			WLS+N+MC		0.8754	0.0414	0.1625	0.1218	0.5189	0.0643
5			SNV + MC		0.6805	0.0965	0.0475	0.0663	0.3640	0.0585
6			SNV + WLS + MC		0.6021	0.0711	0.1350	0.1180	0.3685	0.0689
7			SNV + N + MC		0.6503	0.1085	0.0300	0.0539	0.3401	0.0606
8			SNV + WLS + N + MC		0.6862	0.0656	0.1000	0.1071	0.3931	0.0628
9			1D + MC	3	0.7785	0.0365	0.0000	0.0000	0.3892	0.0183
10				7	0.7559	0.0360	0.0375	0.0631	0.3967	0.0363
11				11	0.7646	0.0376	0.0625	0.0631	0.4136	0.0367
12				15	0.7508	0.0324	0.0400	0.0589	0.3954	0.0336
13				19	0.7656	0.0284	0.0675	0.0629	0.4166	0.0345
14				23	0.7795	0.0449	0.0525	0.0623	0.4160	0.0384
15			1D + N + MC	3	0.7928	0.0319	0.0000	0.0000	0.3964	0.0159
16				7	0.7703	0.0362	0.0650	0.0631	0.4176	0.0364
17				11	0.7749	0.0367	0.0900	0.0567	0.4324	0.0338
18				15	0.7533	0.0366	0.0425	0.0598	0.3979	0.0351
19				19	0.7354	0.0465	0.0100	0.0343	0.3727	0.0289
20				23	0.7492	0.0452	0.0300	0.0539	0.3896	0.0352
21			2D + MC	3	0.8103	0.0351	0.0000	0.0000	0.4051	0.0176
22				7	0.7472	0.0302	0.0000	0.0000	0.3736	0.0151
23				11	0.7359	0.0371	0.0025	0.0177	0.3692	0.0205
24				15	0.8236	0.0372	0.0725	0.0672	0.4480	0.0384
25				19	0.7867	0.0295	0.0500	0.0619	0.4183	0.0343
26				23	0.7646	0.0365	0.0625	0.0631	0.4136	0.0365

27	2D + N + MC	3	0.8492	0.0305	0.0000	0.0000	0.4246	0.0152
28		7	0.7533	0.0323	0.0000	0.0000	0.3767	0.0162
29		11	0.7400	0.0384	0.0125	0.0379	0.3763	0.0270
30		15	0.8154	0.0366	0.1125	0.0725	0.4639	0.0406
31		19	0.7867	0.0321	0.0625	0.0680	0.4246	0.0376
32		23	0.7621	0.0298	0.0350	0.0567	0.3985	0.0320
33	SM + MC	3	0.6795	0.0972	0.0150	0.0482	0.3472	0.0543
34		7	0.6821	0.0881	0.0325	0.0554	0.3573	0.0520
35		11	0.6872	0.0777	0.0225	0.0485	0.3548	0.0458
36		15	0.6349	0.0866	0.0225	0.0485	0.3287	0.0496
37		19	0.6395	0.0892	0.0225	0.0485	0.3310	0.0508
38		23	0.6046	0.0902	0.0150	0.0410	0.3098	0.0495
39	SM + SNV + MC	3	0.6744	0.1102	0.0425	0.0598	0.3584	0.0627
40		7	0.6600	0.1153	0.0325	0.0554	0.3463	0.0640
41		11	0.6400	0.0915	0.0300	0.0595	0.3350	0.0546
42		15	0.5985	0.1008	0.0450	0.0747	0.3217	0.0627
43		19	0.5487	0.1031	0.0325	0.0554	0.2906	0.0585
44		23	0.5385	0.1105	0.0150	0.0410	0.2767	0.0589
45	SM + SNV + WLS + MC	3	0.6021	0.0821	0.1250	0.0911	0.3635	0.0613
46		7	0.6087	0.0829	0.1150	0.1207	0.3619	0.0732
47		11	0.5687	0.0764	0.1300	0.1335	0.3494	0.0769
48		15	0.5877	0.0695	0.1775	0.1290	0.3826	0.0733
49		19	0.5836	0.0679	0.2075	0.1303	0.3955	0.0734
50		23	0.5877	0.0647	0.2000	0.1312	0.3938	0.0731
51	SM + SNV + N + MC	3	0.6810	0.1007	0.0400	0.0589	0.3605	0.0583
52		7	0.6595	0.1011	0.0325	0.0609	0.3460	0.0590
53		11	0.6436	0.1146	0.0450	0.0657	0.3443	0.0660
54		15	0.5938	0.1052	0.0400	0.0641	0.3169	0.0616
55		19	0.5595	0.1257	0.0325	0.0609	0.2960	0.0698

56		23	0.5195	0.1116	0.0225	0.0547	0.2710	0.0621
57	SM + SNV + WLS + N + MC	3	0.6764	0.0691	0.1050	0.1082	0.3907	0.0642
58		7	0.6646	0.0669	0.1125	0.1079	0.3886	0.0635
59		11	0.6513	0.0639	0.1550	0.1117	0.4031	0.0643
60		15	0.6169	0.0666	0.1400	0.1303	0.3785	0.0732
61		19	0.6262	0.0521	0.2000	0.1406	0.4131	0.0750
62		23	0.6118	0.0505	0.2725	0.1280	0.4421	0.0688
63	SM + WLS + MC	3	0.8000	0.0541	0.0225	0.0547	0.4113	0.0385
64		7	0.7841	0.0602	0.0250	0.0505	0.4046	0.0393
65		11	0.7785	0.0423	0.0275	0.0523	0.4030	0.0336
66		15	0.7477	0.0664	0.0275	0.0581	0.3876	0.0441
67		19	0.7010	0.0644	0.0250	0.0505	0.3630	0.0409
68		23	0.6472	0.0858	0.0325	0.0706	0.3398	0.0556
69	SM + WLS + N + MC	3	0.6969	0.0659	0.1050	0.1082	0.4010	0.0633
70		7	0.6723	0.0594	0.1000	0.1101	0.3862	0.0625
71		11	0.6349	0.0619	0.1450	0.1021	0.3899	0.0597
72		15	0.6328	0.0585	0.1475	0.0968	0.3902	0.0565
73		19	0.6087	0.0540	0.1925	0.1455	0.4006	0.0776
74		23	0.6067	0.0482	0.1875	0.1218	0.3971	0.0655
75	SM + N + MC	3	0.6123	0.0769	0.0250	0.0505	0.3187	0.0460
76		7	0.6092	0.0933	0.0200	0.0527	0.3146	0.0536
77		11	0.6133	0.0830	0.0450	0.0657	0.3292	0.0529
78		15	0.6108	0.0702	0.0125	0.0379	0.3116	0.0399
79		19	0.5800	0.0910	0.0250	0.0505	0.3025	0.0521
80		23	0.5708	0.0833	0.0250	0.0565	0.2979	0.0503

NIR: near-infrared; 2T: second trimester; R1: range 1; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation; MC: mean centering; N: normalization; WLS: weighted least squares; SNV: standard normal variate; 1D: first derivative; 2D: second derivative; SM: smoothing.

Table 4.4-S8. Predictive performance of models using NIR spectral data from second trimester sera (range 2, 7600-5100 cm^{-1}).

Model number	Dataset	Detail	Pretreatment	Width	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
1	NIR2T	R2	MC		0.6692	0.0602	0.2550	0.1183	0.4621	0.0664
2			N+MC		0.5974	0.0758	0.2900	0.1503	0.4437	0.0842
3			WLS+MC		0.7605	0.0410	0.1325	0.1084	0.4465	0.0579
4			WLS+N+MC		0.8267	0.0426	0.2200	0.0930	0.5233	0.0511
5			SNV + MC		0.6200	0.0540	0.2825	0.1308	0.4513	0.0707
6			SNV + WLS + MC		0.6554	0.0600	0.1550	0.0930	0.4052	0.0553
7			SNV + N + MC		0.6241	0.0527	0.2750	0.1360	0.4496	0.0729
8			SNV + WLS + N + MC		0.7656	0.0577	0.1850	0.0919	0.4753	0.0542
9			1D + MC	3	0.5190	0.0638	0.3175	0.1365	0.4182	0.0753
10				7	0.7477	0.0519	0.1525	0.1109	0.4501	0.0612
11				11	0.7985	0.0466	0.1550	0.1088	0.4767	0.0592
12				15	0.7615	0.0425	0.1650	0.1169	0.4633	0.0622
13				19	0.7574	0.0452	0.1875	0.1049	0.4725	0.0571
14				23	0.7626	0.0528	0.2200	0.1145	0.4913	0.0630
15			1D + N + MC	3	0.6364	0.0592	0.1975	0.1132	0.4170	0.0639
16				7	0.7662	0.0386	0.1550	0.1028	0.4606	0.0549
17				11	0.7892	0.0431	0.2225	0.1080	0.5059	0.0581
18				15	0.7769	0.0352	0.2600	0.1125	0.5185	0.0589
19				19	0.7697	0.0438	0.2550	0.1128	0.5124	0.0605
20				23	0.7595	0.0543	0.2400	0.0869	0.4997	0.0512
21			2D + MC	3	0.5159	0.0314	0.5400	0.1113	0.5279	0.0578
22				7	0.4944	0.0740	0.1975	0.1290	0.3459	0.0744
23				11	0.6605	0.0642	0.1125	0.1108	0.3865	0.0640
24				15	0.6610	0.0688	0.1875	0.1018	0.4243	0.0614
25				19	0.7538	0.0683	0.1400	0.0999	0.4469	0.0605
26				23	0.7595	0.0538	0.1250	0.0911	0.4422	0.0529

27	2D + N + MC	3	0.6821	0.0288	0.3875	0.1191	0.5348	0.0613
28		7	0.7282	0.0489	0.1025	0.0968	0.4154	0.0542
29		11	0.7595	0.0523	0.2050	0.0937	0.4822	0.0536
30		15	0.7282	0.0494	0.1800	0.0950	0.4541	0.0536
31		19	0.7995	0.0442	0.1825	0.0952	0.4910	0.0525
32		23	0.8246	0.0356	0.2025	0.0941	0.5136	0.0503
33	SM + MC	3	0.6554	0.0577	0.1950	0.1105	0.4252	0.0624
34		7	0.6610	0.0519	0.2175	0.1124	0.4393	0.0619
35		11	0.6518	0.0510	0.2450	0.1100	0.4484	0.0606
36		15	0.6349	0.0686	0.2275	0.1000	0.4312	0.0607
37		19	0.6262	0.0536	0.2250	0.1010	0.4256	0.0572
38		23	0.6415	0.0537	0.2275	0.1121	0.4345	0.0621
39	SM + SNV + MC	3	0.6221	0.0581	0.2700	0.1523	0.4460	0.0815
40		7	0.6303	0.0572	0.2750	0.1406	0.4526	0.0759
41		11	0.6359	0.0543	0.2525	0.1394	0.4442	0.0748
42		15	0.6185	0.0550	0.2925	0.1327	0.4555	0.0718
43		19	0.6287	0.0575	0.2775	0.1219	0.4531	0.0674
44		23	0.6323	0.0527	0.2750	0.1237	0.4537	0.0672
45	SM + SNV + WLS + MC	3	0.6374	0.0480	0.1800	0.1076	0.4087	0.0589
46		7	0.6513	0.0579	0.1675	0.1201	0.4094	0.0667
47		11	0.6277	0.0536	0.1525	0.0955	0.3901	0.0548
48		15	0.6395	0.0549	0.1600	0.1131	0.3997	0.0628
49		19	0.6159	0.0711	0.1575	0.1179	0.3867	0.0689
50		23	0.6221	0.0595	0.1550	0.1145	0.3885	0.0645
51	SM + SNV + N + MC	3	0.6159	0.0493	0.2725	0.1255	0.4442	0.0674
52		7	0.6297	0.0635	0.2800	0.1373	0.4549	0.0756
53		11	0.6390	0.0677	0.2825	0.1283	0.4607	0.0725
54		15	0.6287	0.0563	0.2400	0.1402	0.4344	0.0756
55		19	0.6262	0.0595	0.2850	0.1102	0.4556	0.0626

56		23	0.6174	0.0460	0.2550	0.1359	0.4362	0.0717
57	SM + SNV + WLS + N + MC	3	0.7544	0.0518	0.1950	0.0805	0.4747	0.0479
58		7	0.7359	0.0516	0.1675	0.0965	0.4517	0.0547
59		11	0.7328	0.0531	0.1500	0.0875	0.4414	0.0512
60		15	0.7097	0.0476	0.1400	0.0862	0.4249	0.0492
61		19	0.7097	0.0479	0.1625	0.0986	0.4361	0.0548
62		23	0.7159	0.0480	0.2000	0.1185	0.4579	0.0639
63	SM + WLS + MC	3	0.7533	0.0520	0.1425	0.0911	0.4479	0.0525
64		7	0.7523	0.0413	0.1625	0.1018	0.4574	0.0549
65		11	0.7338	0.0480	0.1650	0.0854	0.4494	0.0490
66		15	0.7374	0.0590	0.1375	0.0919	0.4375	0.0546
67		19	0.7251	0.0572	0.1625	0.0809	0.4438	0.0495
68		23	0.7164	0.0612	0.1525	0.0921	0.4345	0.0553
69	SM + WLS + N + MC	3	0.7728	0.0536	0.1875	0.0809	0.4802	0.0485
70		7	0.7477	0.0447	0.1650	0.0854	0.4563	0.0482
71		11	0.7287	0.0568	0.1750	0.0838	0.4519	0.0506
72		15	0.7159	0.0502	0.1375	0.1049	0.4267	0.0581
73		19	0.7154	0.0392	0.1625	0.1136	0.4389	0.0601
74		23	0.7164	0.0561	0.1750	0.1071	0.4457	0.0605
75	SM + N + MC	3	0.6005	0.0558	0.3175	0.1293	0.4590	0.0704
76		7	0.6349	0.0631	0.2775	0.1390	0.4562	0.0763
77		11	0.6015	0.0845	0.3150	0.1389	0.4583	0.0813
78		15	0.5974	0.0692	0.3075	0.1433	0.4525	0.0796
79		19	0.5928	0.0741	0.3025	0.1496	0.4477	0.0835
80		23	0.5928	0.0629	0.3050	0.1560	0.4489	0.0841

NIR: near-infrared; 2T: second trimester; R2: range 2; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation; MC: mean centering; N: normalization; WLS: weighted least squares; SNV: standard normal variate; 1D: first derivative; 2D: second derivative; SM: smoothing.

Table 4.4-S9. Predictive performance of models using NIR spectral data from second trimester sera (range 3, 5100-4000 cm^{-1}).

Model number	Dataset	Detail	Pretreatment	Width	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
1	NIR2T	R3	MC		0.7805	0.0421	0.6375	0.1318	0.7090	0.0692
2			N+MC		0.7626	0.0483	0.4550	0.0937	0.6088	0.0527
3			WLS+MC		0.8108	0.0401	0.5600	0.1217	0.6854	0.0641
4			WLS+N+MC		0.7497	0.0403	0.4425	0.1048	0.5961	0.0561
5			SNV + MC		0.8133	0.0411	0.7075	0.0998	0.7604	0.0540
6			SNV + WLS + MC		0.7913	0.0427	0.6125	0.1049	0.7019	0.0566
7			SNV + N + MC		0.7908	0.0467	0.7025	0.1126	0.7466	0.0610
8			SNV + WLS + N + MC		0.8426	0.0351	0.4925	0.1222	0.6675	0.0636
9			1D + MC	3	0.7215	0.0577	0.1200	0.0944	0.4208	0.0553
10				7	0.7918	0.0465	0.3875	0.1736	0.5896	0.0899
11				11	0.8415	0.0350	0.5925	0.1179	0.7170	0.0615
12				15	0.8713	0.0361	0.7075	0.0783	0.7894	0.0431
13				19	0.7815	0.0526	0.5500	0.1383	0.6658	0.0740
14				23	0.7523	0.0490	0.4975	0.1300	0.6249	0.0695
15			1D + N + MC	3	0.7949	0.0460	0.1600	0.1042	0.4774	0.0570
16				7	0.8082	0.0422	0.3850	0.1533	0.5966	0.0795
17				11	0.8395	0.0451	0.7200	0.1226	0.7797	0.0653
18				15	0.8415	0.0445	0.7200	0.1301	0.7808	0.0688
19				19	0.8026	0.0425	0.7100	0.1113	0.7563	0.0596
20				23	0.7723	0.0403	0.6700	0.1064	0.7212	0.0569
21			2D + MC	3	0.7497	0.0368	0.0000	0.0000	0.3749	0.0184
22				7	0.8431	0.0416	0.0000	0.0000	0.4215	0.0208
23				11	0.6733	0.0553	0.0075	0.0300	0.3404	0.0314
24				15	0.7221	0.0428	0.1625	0.1108	0.4423	0.0594
25				19	0.7446	0.0460	0.1350	0.1067	0.4398	0.0581
26				23	0.7436	0.0455	0.2225	0.0921	0.4830	0.0513

27	2D + N + MC	3	0.7585	0.0302	0.0000	0.0000	0.3792	0.0151
28		7	0.8441	0.0397	0.0100	0.0343	0.4271	0.0262
29		11	0.6923	0.0515	0.0100	0.0343	0.3512	0.0309
30		15	0.7528	0.0394	0.1825	0.1293	0.4677	0.0676
31		19	0.7744	0.0374	0.1350	0.1067	0.4547	0.0565
32		23	0.7641	0.0455	0.2325	0.1237	0.4983	0.0659
33	SM + MC	3	0.7856	0.0442	0.6200	0.1262	0.7028	0.0668
34		7	0.7708	0.0450	0.6750	0.1010	0.7229	0.0553
35		11	0.7605	0.0444	0.6500	0.1129	0.7053	0.0607
36		15	0.7682	0.0433	0.6475	0.0934	0.7079	0.0515
37		19	0.7636	0.0455	0.6400	0.1061	0.7018	0.0577
38		23	0.7759	0.0417	0.6575	0.1095	0.7167	0.0586
39	SM + SNV + MC	3	0.8072	0.0458	0.6925	0.1017	0.7498	0.0558
40		7	0.8015	0.0401	0.6825	0.0952	0.7420	0.0516
41		11	0.8062	0.0472	0.6800	0.1189	0.7431	0.0640
42		15	0.7964	0.0432	0.6950	0.1105	0.7457	0.0593
43		19	0.8103	0.0427	0.6875	0.1108	0.7489	0.0594
44		23	0.8041	0.0477	0.7100	0.1142	0.7571	0.0619
45	SM + SNV + WLS + MC	3	0.7810	0.0337	0.5950	0.1373	0.6880	0.0707
46		7	0.7938	0.0388	0.6425	0.1101	0.7182	0.0584
47		11	0.7862	0.0410	0.6250	0.1010	0.7056	0.0545
48		15	0.7846	0.0405	0.6525	0.1138	0.7186	0.0604
49		19	0.7897	0.0411	0.6475	0.1092	0.7186	0.0583
50		23	0.8010	0.0462	0.6500	0.1041	0.7255	0.0570
51	SM + SNV + N + MC	3	0.8026	0.0360	0.7100	0.0960	0.7563	0.0512
52		7	0.8046	0.0424	0.7075	0.1089	0.7561	0.0584
53		11	0.7954	0.0453	0.7150	0.1042	0.7552	0.0568
54		15	0.7954	0.0416	0.6650	0.1055	0.7302	0.0567
55		19	0.8092	0.0385	0.7125	0.1164	0.7609	0.0613

56		23	0.8164	0.0428	0.6825	0.1217	0.7495	0.0645
57	SM + SNV + WLS + N + MC	3	0.8364	0.0404	0.4975	0.1115	0.6670	0.0593
58		7	0.8436	0.0409	0.5000	0.1010	0.6718	0.0545
59		11	0.8456	0.0326	0.5300	0.1199	0.6878	0.0621
60		15	0.8374	0.0465	0.5100	0.1491	0.6737	0.0781
61		19	0.8282	0.0455	0.5225	0.1399	0.6754	0.0736
62		23	0.8431	0.0432	0.5350	0.1186	0.6890	0.0631
63	SM + WLS + MC	3	0.8010	0.0410	0.5850	0.1223	0.6930	0.0645
64		7	0.8159	0.0488	0.5675	0.1135	0.6917	0.0618
65		11	0.8082	0.0458	0.5650	0.1018	0.6866	0.0558
66		15	0.8159	0.0407	0.5750	0.0945	0.6954	0.0514
67		19	0.8159	0.0436	0.5725	0.1239	0.6942	0.0657
68		23	0.8056	0.0531	0.5825	0.1174	0.6941	0.0644
69	SM + WLS + N + MC	3	0.8328	0.0392	0.4900	0.1207	0.6614	0.0634
70		7	0.8226	0.0550	0.4975	0.0961	0.6600	0.0554
71		11	0.8328	0.0402	0.5150	0.1375	0.6739	0.0716
72		15	0.8405	0.0367	0.5175	0.1042	0.6790	0.0552
73		19	0.8369	0.0318	0.5325	0.1004	0.6847	0.0527
74		23	0.8369	0.0376	0.5300	0.1173	0.6835	0.0616
75	SM + N + MC	3	0.7692	0.0521	0.4475	0.0981	0.6084	0.0555
76		7	0.7692	0.0443	0.4650	0.1158	0.6171	0.0620
77		11	0.7487	0.0508	0.4775	0.1000	0.6131	0.0561
78		15	0.7426	0.0494	0.4850	0.1147	0.6138	0.0625
79		19	0.7508	0.0500	0.4750	0.1101	0.6129	0.0604
80		23	0.7549	0.0565	0.4725	0.1109	0.6137	0.0622

NIR: near-infrared; 2T: second trimester; R3: range 3; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation; MC: mean centering; N: normalization; WLS: weighted least squares; SNV: standard normal variate; 1D: first derivative; 2D: second derivative; SM: smoothing.

Table 4.4-S10. Predictive performance of models combining NIR spectral data of second trimester sera with medical parameters.

NIR 2T range[a]	Plus	Block order[b]	Se (Av)	Se (StD)	Sp (Av)	Sp (StD)	NER (Av)	NER (StD)
Full			0.8133	0.0324	0.1150	0.0556	0.4642	0.0321
Full	Medical	1-2	0.8718	0.0398	0.3175	0.1048	0.5946	0.0560
Full	Medical	2-1	0.8703	0.0360	0.2950	0.0747	0.5826	0.0415
Full simp[c]	Medical simp[c]	1-2	0.6759	0.0915	0.5125	0.1478	0.5942	0.0869
Full simp[c]	Medical simp[c]	2-1	0.7600	0.0727	0.5275	0.1457	0.6438	0.0814
R3			0.8713	0.0361	0.7075	0.0783	0.7894	0.0431
R3	Medical	1-2	0.8667	0.0319	0.5750	0.1617	0.7208[d]	0.0824[d]
R3	Medical	2-1	0.8487	0.0511	0.4875	0.1641	0.6681	0.0860
R3 simp[c]	Medical simp[c]	1-2	0.6492	0.0713	0.4625	0.1562	0.5559	0.0858
R3 simp[c]	Medical simp[c]	2-1	0.7195	0.0748	0.4750	0.1263	0.5972	0.0734

[a] Full: 10500-4000 cm^{-1}; R3: 5100-4000 cm^{-1}.

[b] NIR blocks were pretreated using the conditions identified as best in single-block analysis. Medical blocks were pretreated by autoscaling.

[c] Data obtained after variable selection, using the variable importance information from the original equivalent multi-block model.

[d] Best multi-block model. Its DCV-AUC-ROC is 0.8126 ± 0.0696.

NIR: near-infrared; 2T: second trimester; R3: range 3; Se: sensitivity; Sp: specificity; NER: non-error rate; Av: average; StD: standard deviation.

Table 4.4-S11. Molecular vibrations and biomolecules that have been associated with the most relevant spectral intervals of the best first trimester model.

Spectral interval	Tentative molecular vibration	Theoretical wavenumber	Potential biomolecule	Reference
10500-9828 cm^{-1}	Second overtone of O-H stretching	9911, 10288-10300 cm^{-1}	Carbohydrates	[32,34]
	Second overtone of N-H stretching	10277-9804 cm^{-1}	Proteins	[32]
8826-7858 cm^{-1}	Second overtone of C-H stretching	8250-8696 cm^{-1}	Carbohydrates and lipids	[32,33]

Table 4.4-S12. Molecular vibrations and biomolecules that have been associated with the most relevant spectral intervals of the best second trimester model.

Spectral interval	Tentative molecular vibration	Theoretical wavenumber	Potential biomolecule	Reference
5028-4856 cm^{-1}	Combination of N-H stretching + C-N stretching + N-H bending	4925, 4975, 5025 cm^{-1}	Proteins	[32]
	Second overtone of C=O stretching	4926 cm^{-1}	Proteins	[32]
	Combination of N-H stretching + N-H bending	4880-4902 cm^{-1}	Proteins	[32,33]
	Combination of N-H bending + C-N stretching + N-H bending	4878 cm^{-1}	Proteins	[32]
	Combination of N-H stretching + C=O stretching	4866 cm^{-1}	Proteins	[32,33]
	Combination of second overtone of N-H bending + N-H stretching	4854 cm^{-1}	Proteins	[32]
	Combination band of N-H stretching + N-H bending	4850 cm^{-1}	Proteins	[32]
4764-4702 cm^{-1}	Third overtone of C=O-O stretching	4762 cm^{-1}	Carbohydrates	[32]
	Combination of O-H bending + C-O stretching	4762 cm^{-1}	Carbohydrates	[32]
4492-4442 cm^{-1}	Combination bands of C-H	4250-4760 cm^{-1}	Carbohydrates	[32,34]
4392-4364 cm^{-1}	Combination of C-H stretching + CH_2 deformation	4386 cm^{-1}	Carbohydrates	[32]
4302-4268 cm^{-1}	Combination of C-H stretching + CH_2 deformation	4283, 4292 cm^{-1}	Carbohydrates	[32]
4206-4176 cm^{-1}	Combination band of C-H stretching + C-H stretching	4202 cm^{-1}	Lipids	[32]
4096-4000 cm^{-1}	Combination bands of C-H	4049 cm^{-1}	Lipids	[32]
	First overtone of C-N-C stretching	4049 cm^{-1}	Proteins	[32]
	Combination of C-H stretching + C-C stretching + C-O-C stretching	4000 cm^{-1}	Carbohydrates	[32]

5. CONCLUSIÓN

En esta investigación, se desarrollaron distintos métodos multivariados para predecir DG en el primer trimestre del embarazo (**Figura 5-1**). El mejor modelo desarrollado en base a datos gineco-obstétricos logró un AUC de 0,867 en validación interna. Por otro lado, el mejor modelo generado en base a datos espectrales de suero consiguió un AUC de 0,577 en validación interna, y un tiempo total de análisis de 32 minutos. De manera similar, el mejor modelo desarrollado en base a la combinación de datos gineco-obstétricos y espectrales de suero, obtuvo un AUC de 0,522 en validación interna, y un tiempo total de análisis de 32 minutos.

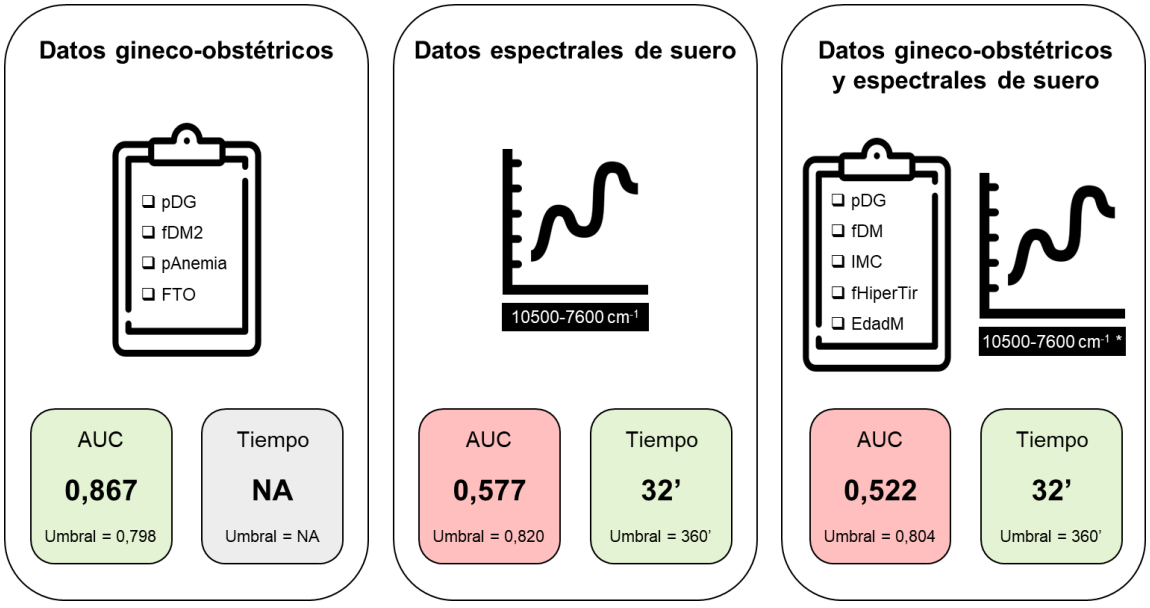


Figura 5-1. Métodos multivariados desarrollados para predecir DG en el primer trimestre del embarazo. Para cada objetivo específico de este trabajo (rectángulo redondeado grande), se presenta el mejor modelo desarrollado. Para cada modelo, se muestra el tipo de variables utilizadas, en general (título) y en detalle (íconos). Igualmente, se expone el poder predictivo en términos de AUC (rectángulo redondeado pequeño izquierdo) y el tiempo de análisis en minutos (rectángulo redondeado pequeño derecho). El poder predictivo y el tiempo de análisis se contrastan respecto al umbral establecido en base a literatura. Existen tres situaciones: mejora respecto al umbral (verde), no mejora respecto al umbral (rojo), y no aplica (gris). *pDG: antecedente previo de DG; fDM2: historia familiar de DM2; pAnemia: historia personal de anemia; FTO: genotipo FTO; NA: no aplica; fDM: historia familiar de DM; IMC: índice de masa corporal; fHiperTir: historia familiar de hipertiroidismo; EdadM: edad menarquia. * El modelo se construyó con un 41% de las variables del rango espectral presentado.*

5.1. Conclusión parcial del primer objetivo

El método generado a partir de datos gineco-obstétricos es el único que logra superar el umbral de poder predictivo establecido en base a literatura internacional. De manera interesante, este también supera el desempeño de modelos nacionales [5,90,91]. Para su desarrollo, fue fundamental aplicar técnicas multivariadas distintas a LR, la más utilizada en el área [47]. En particular, el mejor poder predictivo se obtuvo con la técnica de análisis discriminante por mínimos cuadrados parciales (PLS-DA), cuyo uso en literatura suele estar restringido a estudios de metabolómica [92]. De igual manera, fue esencial considerar predictores distintos a los clásicos, para potenciar su desempeño. En efecto, el modelo final combina cuatro parámetros, antecedente previo de DG, historia familiar de DM2, historia personal de anemia y genotipo FTO; entre los cuales solo los dos primeros son factores de riesgo clásicos para DG [2,5]. Es más, fue muy importante utilizar predictores de origen bioquímico, además de aquellos de origen clínico. En efecto, el genotipo FTO contribuyó de manera

sustancial al poder predictivo del modelo final. También fue muy relevante aplicar selección de variables. Esto permitió simplificar el modelo para su potencial aplicación clínica, como se ha sugerido en literatura [46,93]; y también mejorar su poder predictivo. Tanto por la accesibilidad a los predictores, como por su interpretabilidad, los modelos que se basan en datos gineco-obstétricos son los más factibles de implementar en la realidad clínica. Por su desempeño y su simpleza, el método desarrollado es una excelente alternativa para predecir DG en el primer trimestre del embarazo.

5.2. Conclusión parcial del segundo objetivo

El método generado a partir de datos espectrales de suero no logra mejorar el poder predictivo respecto al umbral definido en base a literatura; sin embargo reduce sustancialmente el tiempo de análisis. Para su desarrollo, se utilizó espectroscopía NIR. La principal ventaja de esta técnica instrumental son sus propiedades analíticas complementarias, tales como simpleza, rapidez, y bajo costo [61]. Esto se suma a la objetividad propia de analizar líquidos biológicos con técnicas instrumentales, que contrastan con la subjetividad propia de recolectar información clínica a través de formularios de autorreporte. Una ventaja adicional de la espectroscopía NIR es que permite identificar cambios bioquímicos asociados a una condición fisiopatológica de interés [94], lo cual contribuye a la comprensión de los mecanismos que la subyacen. En este caso,

permitió detectar alteraciones espectrales que se asocian a carbohidratos, lípidos y proteínas, en el suero de mujeres con DG. En contraste, la principal desventaja de la espectroscopía NIR es su baja sensibilidad analítica [95]. Esta característica podría explicar parcialmente el bajo poder predictivo del método desarrollado. Es probable que en el primer trimestre de gestación las diferencias bioquímicas del suero de mujeres con y sin DG sean muy sutiles, y el método presentado no sea capaz de detectarlas. Es importante señalar que este es el primer estudio que utiliza espectroscopía NIR para predecir DG y, por lo tanto, existen aspectos que se pueden mejorar. Por ejemplo, se podría modificar el proceso de preparación de muestra, con el fin de remover analitos de alta concentración que puedan estar interfiriendo con la detección de analitos de menor concentración, como se pretende al precipitar proteínas en estudios de metabolómica [96]. Esta modificación, seguida de una apropiada optimización del pretratamiento de los datos, podría mejorar el desempeño de método actual. De igual manera, se podrían evaluar técnicas multivariadas distintas a PLS-DA, lo que también podría aumentar el poder predictivo. Por sus propiedades analíticas complementarias, este método tiene un gran potencial para ser aplicado en la detección temprana de DG. Sin embargo, su acercamiento a la realidad clínica queda supeditado a la mejora de su poder predictivo.

5.3. Conclusión parcial del tercer objetivo

El método generado a partir de datos gineco-obstétricos y espectrales de suero no logra mejorar el poder predictivo; sin embargo disminuye sustancialmente el tiempo de análisis respecto al umbral definido en base a literatura. Para su desarrollo, se utilizaron parámetros de origen clínico, espectroscopía NIR, análisis multivariado del tipo *multi-block* y selección de variables. El uso de espectroscopía NIR es el que permite lograr un tiempo de análisis tan reducido. En general, tanto la aplicación de técnicas *multi-block* como el uso de selección de variables se asocian a un incremento en el poder predictivo [26,97]; sin embargo, en muchos casos no es así [28,98]. En este caso, ninguna de las dos estrategias permitió mejorar el desempeño respecto al obtenido a partir de los bloques de data clínica y NIR por sí solos. El bajo poder predictivo de los bloques individuales podría explicar parcialmente el modesto desempeño del modelo final. Es importante señalar que este es el primer estudio que utiliza técnicas multivariadas del tipo *multi-block* para la predicción de DG y, por lo tanto, existen aspectos que se pueden mejorar. De acuerdo a lo discutido anteriormente para los bloques individuales, se podrían incorporar otras variables, particularmente, parámetros bioquímicos; y optimizar el proceso de preparación de muestra previo a la adquisición espectral. De igual manera, se podrían evaluar técnicas *multi-block* de mayor complejidad, que muchas veces se asocian a un mayor poder predictivo [28]. La implementación de estas modificaciones podría mejorar el

desempeño de método actual. Por las propiedades analíticas complementarias de la técnica instrumental en que se basa, este método tiene un gran potencial para ser utilizado en la detección temprana de DG. Sin embargo, su acercamiento a la práctica clínica queda condicionado a la mejora de su poder predictivo.

5.4. Conclusión final

El análisis multivariado ha sido ampliamente utilizado para resolver problemáticas en obstetricia y ginecología, y particularmente para predecir DG en etapas tempranas del embarazo. Los métodos de detección temprana para DG desarrollados en este trabajo permitieron mejorar el poder predictivo y el tiempo de análisis respecto a los métodos reportados en literatura, lo cual confirma la hipótesis de esta investigación.

De acuerdo con lo anterior, la conclusión final de esta tesis es que las estrategias basadas en el análisis multivariado de datos gineco-obstétricos y/o espectrales de suero del primer trimestre de gestación mejoran la predicción de DG en etapas tempranas del embarazo, ya sea incrementando el poder predictivo, o reduciendo los tiempos de análisis.

Los resultados de este trabajo demuestran la importancia de innovar para generar mejoras respecto a lo estudios que componen el estado del arte. En

efecto, la estrategia analítica utilizada permitió lograr progresos respecto a publicaciones nacionales e internacionales. Todos los modelos presentados tienen un gran potencial para ser utilizados en la predicción de DG, siendo el basado en data gineco-obstétrica el que está más cercano a una posible aplicación clínica en embarazos de primer trimestre.

Los resultados de esta investigación también demuestran que el análisis multivariado puede ser utilizado exitosamente con otros fines, tales como predecir la glicemia post carga propiamente tal como una manera de predecir DG bajo cualquier criterio diagnóstico; y predecir DG en el segundo trimestre del embarazo como un método de screening alternativo a la PTGO. El análisis de data gineco-obstétrica en el primer caso, y de data espectral de suero en el segundo, constituyeron tácticas particularmente novedosas que propiciaron mejoras importantes respecto al estado del arte.

Las proyecciones de este estudio incluyen explorar estrategias para mejorar el desempeño de los métodos basados en espectroscopía NIR y técnicas *multi-block*, y realizar estudios de validación externa con el fin de evaluar los modelos presentados en diferentes poblaciones. Esto permitirá acercar el trabajo realizado a la práctica clínica, desde donde podría aportar no solo a la predicción temprana de DG propiamente tal, sino también a la prevención de las consecuencias negativas que ésta conlleva para la salud materna y filial a corto y largo plazo.

6. REFERENCIAS

Las referencias enlistadas a continuación corresponden a los capítulos 1, 2, 3 y 5 de la presente tesis. Las referencias del capítulo 4 se enlistan al final de cada una de sus secciones (4.1, 4.2, 4.3 y 4.4).

1. ElSayed NA, Aleppo G, Aroda VR, Bannuru RR, Brown FM, Bruemmer D, et al. 2. Classification and Diagnosis of Diabetes: Standards of Care in Diabetes-2023. Diabetes Care. 2023;46: S19–S40. doi:10.2337/dc23-S002

2. McIntyre HD, Catalano P, Zhang C, Desoye G, Mathiesen ER, Damm P. Gestational diabetes mellitus. Nat Rev Dis Prim. 2019;5: 47. doi:10.1038/s41572-019-0098-8

3. Ministerio de Salud. Guía Perinatal. 2015.

4. International Diabetes Federation. Diabetes Atlas. 2021.

5. Garmendia ML, Mondschein S, Montiel B, Kusanovic JP. Trends and predictors of gestational diabetes mellitus in Chile. Int J Gynecol Obstet. 2020;148: 210–218. doi:10.1002/ijgo.13023

6. Plows JF, Stanley JL, Baker PN, Reynolds CM, Vickers MH. The pathophysiology of gestational diabetes mellitus. Int J Mol Sci. 2018;19: 1–21. doi:10.3390/ijms19113342

7. Farahvar S, Walfisch A, Sheiner E. Gestational diabetes risk factors and long-term consequences for both mother and offspring: a literature review. Expert Rev Endocrinol Metab. 2019;14: 63–74. doi:10.1080/17446651.2018.1476135

8. Metzger B, Lowe L, Dyer A, Trimble E, Chaovarindr U, Coustan D, et al. Hyperglycemia and Adverse Pregnancy Outcomes. N Engl J Med. 2008;358: 1991–2002. doi:10.1056/NEJMoa0707943

9. Lowe WL, Scholtens DM, Lowe LP, Kuang A, Nodzenski M, Talbot O, et al. Association of gestational diabetes with maternal disorders of glucose

metabolism and childhood adiposity. JAMA - J Am Med Assoc. 2018;320: 1005–1016. doi:10.1001/jama.2018.11628

10. Lowe WL, Scholtens DM, Kuang A, Linder B, Lawrence JM, Lebenthal Y, et al. Hyperglycemia and adverse Pregnancy Outcome follow-up study (HAPO FUS): Maternal gestational diabetes mellitus and childhood glucose metabolism. Diabetes Care. 2019;42: 372–380. doi:10.2337/dc18-1646

11. Sovio U, Murphy HR, Smith GCS. Accelerated fetal growth prior to diagnosis of gestational diabetes mellitus: A prospective cohort study of nulliparous women. Diabetes Care. 2016;39: 982–987. doi:10.2337/dc16-0160

12. Venkataraman H, Ram U, Craik S, Arungunasekaran A, Seshadri S, Saravanan P. Increased fetal adiposity prior to diagnosis of gestational diabetes in South Asians: more evidence for the 'thin–fat' baby. Diabetologia. 2017;60: 399–405. doi:10.1007/s00125-016-4166-2

13. Yovera L, Zaharia M, Jachymski T, Velicu-Scraba O, Coronel C, de Paco Matallana C, et al. Impact of gestational diabetes mellitus on fetal cardiac morphology and function: cohort comparison of second- and third-trimester fetuses. Ultrasound Obstet Gynecol. 2021;57: 607–613. doi:10.1002/uog.22148

14. Moran PS, Wuytack F, Turner M, Normand C, Brown S, Begley C, et al. Economic burden of maternal morbidity – A systematic review of cost-of-illness studies. PLoS One. 2020;15: 1–18. doi:10.1371/JOURNAL.PONE.0227377

15. Wexler DJ, Powe CE, Barbour LA, Buchanan T, Coustan DR, Corcoy R, et al. Research gaps in gestational diabetes mellitus: Executive summary of a national institute of diabetes and digestive and kidney diseases workshop. Obstet Gynecol. 2018;132: 496–505. doi:10.1097/AOG.0000000000002726

16. McIntyre D, Desoye G, Dunne F, Simeoni U, Visser GHA, Kapur A, et al. FIGO analysis of research priorities in hyperglycemia in pregnancy. Diabetes Res Clin Pract. 2018;145: 5–14. doi:10.1016/j.diabres.2018.03.026

17. Immanuel J, Simmons D. Screening and Treatment for Early-Onset Gestational Diabetes Mellitus: a Systematic Review and Meta-analysis. Curr Diab Rep. 2017;17: 1–11. doi:10.1007/s11892-017-0943-7

18. Powe CE. Early Pregnancy Biochemical Predictors of Gestational Diabetes Mellitus. Curr Diab Rep. 2017;17: 12. doi:10.1007/s11892-017-0834-y

19. Lorenzo-Almorós A, Hang T, Peiró C, Soriano-Guillén L, Egido J, Tuñón J, et al. Predictive and diagnostic biomarkers for gestational diabetes and its associated metabolic and cardiovascular diseases. Cardiovasc Diabetol. 2019;18: 140. doi:10.1186/s12933-019-0935-9

20. Gan WZ, Ramachandran V, Lim CSY, Koh RY. Omics-based biomarkers in the diagnosis of diabetes. J Basic Clin Physiol Pharmacol. 2020;31: 1–21. doi:10.1515/jbcpp-2019-0120

21. Bonongwe P, Lindow SW, Coetzee EJ. Reproducibility of a 75G oral glucose tolerance test in pregnant women. J Perinat Med. 2015;43: 333–338. doi:10.1515/jpm-2014-0208

22. Munang YN, Noubiap JJ, Danwang C, Sama JD, Azabji-Kenfack M, Mbanya JC, et al. Reproducibility of the 75 g oral glucose tolerance test for the diagnosis of gestational diabetes mellitus in a sub-Saharan African population. BMC Res Notes. 2017;10: 1–6. doi:10.1186/s13104-017-2944-7

23. Huhn EA, Rossi SW, Hoesli I, Göbl CS. Controversies in Screening and Diagnostic Criteria for Gestational Diabetes in Early and Late Pregnancy. Front Endocrinol (Lausanne). 2018;9: 5–12. doi:10.3389/fendo.2018.00696

24. National Institute of Standards and Technology. Elements of Multivariate Analysis. Available: https://www.itl.nist.gov/div898/handbook/pmc/section5/pmc54.htm

25. Amigo JM. Data Mining, Machine Learning, Deep Learning, Chemometrics. Definitions, Common Points and Trends (Spoiler Alert: VALIDATE your models!). Brazilian J Anal Chem. 2021;8: 45–61. doi:10.30744/brjac.2179-3425.AR-38-2021

26. Mishra P, Roger JM, Jouan-Rimbaud-Bouveresse D, Biancolillo A, Marini F, Nordon A, et al. Recent trends in multi-block data analysis in chemometrics for multi-source data integration. TrAC - Trends Anal Chem. 2021;137: 116206. doi:10.1016/j.trac.2021.116206

27. Olivieri AC. Analytical advantages of multivariate data processing. One, two, three, infinity? Anal Chem. 2008;80: 5713–5720. doi:10.1021/ac800692c

28. Azcarate SM, Ríos-Reina R, Amigo JM, Goicoechea HC. Data handling in data fusion: Methodologies and applications. TrAC - Trends Anal Chem. 2021;143. doi:10.1016/j.trac.2021.116355

29. Cleophas TJ, Zwinderman AH. Machine learning in medicine. Machine Learning in Medicine. 2013. doi:10.1007/978-94-007-5824-7

30. Mennickent D, Rodríguez A, Opazo MC, Riedel CA, Castro E, Eriz-Salinas A, et al. Machine learning applied in maternal and fetal health: a narrative review focused on pregnancy diseases and complications. Front Endocrinol (Lausanne). 2023;14: 1–22. doi:10.3389/fendo.2023.1130139

31. Goecks J, Jalili V, Heiser LM, Gray JW. How Machine Learning Will Transform Biomedicine. Cell. 2020;181: 92–101. doi:10.1016/j.cell.2020.03.022

32. Szymańska E. Modern data science for analytical chemical data – A comprehensive review. Anal Chim Acta. 2018;1028: 1–10. doi:10.1016/j.aca.2018.05.038

33. Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98: 691–698. doi:10.1136/heartjnl-2011-301247

34. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature XXII: How to use articles about clinical decision rules. J Am Med Assoc. 2000;284: 79–84. doi:10.1001/jama.284.1.79

35. Mennickent D, Rodríguez A, Farías-Jofré M, Araya J, Guzmán-Gutiérrez E. Machine learning-based models for gestational diabetes mellitus prediction before 24–28 weeks of pregnancy: A review. Artif Intell Med. 2022;132: 102378. doi:10.1016/j.artmed.2022.102378

36. Teede HJ, Harrison CL, Teh WT, Paul E, Allan CA. Gestational diabetes: Development of an early risk prediction tool to facilitate opportunities for prevention. Aust New Zeal J Obstet Gynaecol. 2011;51: 499–504. doi:10.1111/j.1479-828X.2011.01356.x

37. Balani J, Hyer S, Shehata H, Mohareb F. Visceral fat mass as a novel risk factor for predicting gestational diabetes in obese pregnant women. Obstet Med. 2018;11: 121–125. doi:10.1177/1753495X17754149

38. Wang C, Zhu W, Wei Y, Su R, Feng H, Lin L, et al. The Predictive Effects of Early Pregnancy Lipid Profiles and Fasting Glucose on the Risk of Gestational Diabetes Mellitus Stratified by Body Mass Index. J Diabetes Res. 2016;2016: 1–8. doi:10.1155/2016/3013567

39. Amini M, Kazemnejad A, Zayeri F, Montazeri A, Rasekhi A, Amirian A, et al. Diagnostic accuracy of maternal serum multiple marker screening for early detection of gestational diabetes mellitus in the absence of a gold standard test. BMC Pregnancy Childbirth. 2020;20: 375. doi:10.1186/s12884-020-03068-7

40. Qiu H, Yu HY, Wang LY, Yao Q, Wu SN, Yin C, et al. Electronic Health Record Driven Prediction for Gestational Diabetes Mellitus in Early Pregnancy. Sci Rep. 2017;7: 1–13. doi:10.1038/s41598-017-16665-y

41. Li G, Huang W, Zhang L, Tian Z, Zheng W, Wang T, et al. A prospective cohort study of early-pregnancy risk factors for gestational diabetes in polycystic ovarian syndrome. Diabetes Metab Res Rev. 2018;34: 3–9. doi:10.1002/dmrr.3003

42. Sweeting AN, Wong J, Appelblom H, Ross GP, Kouru H, Williams PF, et al. A first trimester prediction model for gestational diabetes utilizing aneuploidy and pre-eclampsia screening markers. J Matern Neonatal Med. 2018;31: 2122–2130. doi:10.1080/14767058.2017.1336759

43. Sweeting AN, Wong J, Appelblom H, Ross GP, Kouru H, Williams PF, et al. A Novel Early Pregnancy Risk Prediction Model for Gestational Diabetes Mellitus. Fetal Diagn Ther. 2019;45: 76–84. doi:10.1159/000486853

44. Sirico A, Lanzone A, Mappa I, Sarno L, Słodki M, Pitocco D, et al. The role of first trimester fetal heart rate in the prediction of gestational diabetes: A multicenter study. Eur J Obstet Gynecol Reprod Biol. 2019;243: 158–161. doi:10.1016/j.ejogrb.2019.10.019

45. Zheng T, Ye W, Wang X, Li X, Zhang J, Little J, et al. A simple model to predict risk of gestational diabetes mellitus from 8 to 20 weeks of gestation in Chinese women. BMC Pregnancy Childbirth. 2019;19: 252. doi:10.1186/s12884-019-2374-8

46. Artzi NS, Shilo S, Hadar E, Rossman H, Barbash-Hazan S, Ben-Haroush A, et al. Prediction of gestational diabetes based on nationwide electronic health records. Nat Med. 2020;26: 71–76. doi:10.1038/s41591-019-0724-8

47. Zhang Z, Yang L, Han W, Wu Y, Zhang L, Gao C, et al. Machine Learning

Prediction Models for Gestational Diabetes Mellitus: Meta-analysis. J Med Internet Res. 2022;24: e26634. doi:10.2196/26634

48. Kuo CH, Li HY. Diagnostic Strategies for Gestational Diabetes Mellitus: Review of Current Evidence. Curr Diab Rep. 2019;19: 11–16. doi:10.1007/s11892-019-1271-x

49. Yoffe L, Polsky A, Gilam A, Raff C, Mecacci F, Ognibene A, et al. Early diagnosis of gestational diabetes mellitus using circulating microRNAs. Eur J Endocrinol. 2019;181: 565–577. doi:10.1530/EJE-19-0206

50. Guo Z, Yang F, Zhang J, Zhang Z, Li K, Tian Q, et al. Whole-Genome Promoter Profiling of Plasma DNA Exhibits Diagnostic Value for Placenta-Origin Pregnancy Complications. Adv Sci. 2020;7: 1901819. doi:10.1002/advs.201901819

51. Zheng W, Xu Q, Huang W, Yan Q, Chen Y, Zhang L, et al. Gestational Diabetes Mellitus Is Associated with Reduced Dynamics of Gut Microbiota during the First Half of Pregnancy. mSystems. 2020;5: 1–13. doi:10.1128/msystems.00109-20

52. Ma S, You Y, Huang L, Long S, Zhang J, Guo C, et al. Alterations in Gut Microbiota of Gestational Diabetes Patients During the First Trimester of Pregnancy. Front Cell Infect Microbiol. 2020;10. doi:10.3389/fcimb.2020.00058

53. Sakurai K, Eguchi A, Watanabe M, Yamamoto M, Ishikawa K, Mori C. Exploration of predictive metabolic factors for gestational diabetes mellitus in Japanese women using metabolomic analysis. J Diabetes Investig. 2019;10: 513–520. doi:10.1111/jdi.12887

54. Leitner M, Fragner L, Danner S, Holeschofsky N, Leitner K, Tischler S, et al. Combined metabolomic analysis of plasma and urine reveals AHBA, tryptophan and serotonin metabolism as potential risk factors in Gestational Diabetes Mellitus (GDM). Front Mol Biosci. 2017;4. doi:10.3389/fmolb.2017.00084

55. Diaz SO, Pinto J, Graça G, Duarte IF, Barros AS, Galhano E, et al. Metabolic biomarkers of prenatal disorders: An exploratory NMR metabonomics study of second trimester maternal urine and blood plasma. J Proteome Res. 2011;10: 3732–3742. doi:10.1021/pr200352m

56. White SL, Lawlor DA, Briley AL, Godfrey KM, Nelson SM, Oteng-Ntim E, et al. Early antenatal prediction of gestational diabetes in obese women:

Development of prediction tools for targeted intervention. PLoS One. 2016;11: 1–17. doi:10.1371/journal.pone.0167846

57. Zhao D, Shen L, Wei Y, Xie J, Chen S, Liang Y, et al. Identification of candidate biomarkers for the prediction of gestational diabetes mellitus in the early stages of pregnancy using iTRAQ quantitative proteomics. Proteomics - Clin Appl. 2017;11: 7–8. doi:10.1002/prca.201600152

58. Ravnsborg T, Andersen LLT, Trabjerg ND, Rasmussen LM, Jensen DM, Overgaard M. First-trimester multimarker prediction of gestational diabetes mellitus using targeted mass spectrometry. Diabetologia. 2016;59: 970–979. doi:10.1007/s00125-016-3869-8

59. Larkin PJ. Introduction: Infrared and Raman Spectroscopy. Infrared and Raman Spectroscopy. Principles and Spectral Interpretation. Elsevier; 2018. pp.1–5. doi:10.1016/C2015-0-00806-1

60. Bunaciu AA, Fleschin Ş, Hoang VD, Aboul-Enein HY. Vibrational Spectroscopy in Body Fluids Analysis. Crit Rev Anal Chem. 2017;47: 67–75. doi:10.1080/10408347.2016.1209104

61. Balan V, Mihai CT, Cojocaru FD, Uritu CM, Dodi G, Botezat D, et al. Vibrational spectroscopy fingerprinting in medicine: From molecular to clinical practice. Materials (Basel). 2019;12: 1–40. doi:10.3390/ma12182884

62. Si L, Lin R, Jia Y, Jian W, Yu Q, Wang M, et al. Lactobacillus bulgaricus improves antioxidant capacity of black garlic in the prevention of gestational diabetes mellitus: A randomized control trial. Biosci Rep. 2019;39: 1–19. doi:10.1042/BSR20182254

63. Shapira D, Mandel D, Mimouni FB, Moran-Lev H, Marom R, Mangel L, et al. The effect of gestational diabetes mellitus on human milk macronutrients content. J Perinatol. 2019;39: 820–823. doi:10.1038/s41372-019-0362-5

64. Zhou X, Wang W, Zhong J, Yu W, Tang Y, Shu Q. Biomolecular characterization of placental tissues in gestational diabetes mellitus using Fourier transform infrared microspectroscopy. Spectrochim Acta - Part A Mol Biomol Spectrosc. 2022;270: 120794. doi:10.1016/j.saa.2021.120794

65. Bernardes-Oliveira E, de Freitas DLD, de Morais C de LM, Cornetta M da C de M, Camargo JD de AS, de Lima KMG, et al. Spectrochemical differentiation in gestational diabetes mellitus based on attenuated total reflection Fourier-transform infrared (ATR-FTIR) spectroscopy and

multivariate analysis. Sci Rep. 2020;10: 19259. doi:10.1038/s41598-020-75539-y

66. Dipla K, Triantafyllou A, Grigoriadou I, Kintiraki E, Triantafyllou GA, Poulios P, et al. Impairments in microvascular function and skeletal muscle oxygenation in women with gestational diabetes mellitus: links to cardiovascular disease risk factors. Diabetologia. 2017;60: 192–201. doi:10.1007/s00125-016-4129-7

67. Kintiraki E, Dipla K, Triantafyllou A, Koletsos N, Grigoriadou I, Poulakos P, et al. Blunted cerebral oxygenation during exercise in women with gestational diabetes mellitus: associations with macrovascular function and cardiovascular risk factors. Metabolism. 2018;83: 25–30. doi:10.1016/j.metabol.2018.01.009

68. Vounzoulaki E, Dipla K, Kintiraki E, Triantafyllou A, Grigoriadou I, Koletsos N, et al. Pregnancy and post-partum muscle and cerebral oxygenation during intermittent exercise in gestational diabetes: A pilot study. Eur J Obstet Gynecol Reprod Biol. 2019;232: 54–59. doi:10.1016/j.ejogrb.2018.11.012

69. Kim W, Bang A, Kim S, Lee GJ, Kim YH, Choi S. Adiponectin-targeted SERS immunoassay biosensing platform for early detection of gestational diabetes mellitus. Biosens Bioelectron. 2022;213: 114488. doi:10.1016/j.bios.2022.114488

70. Liu KZ, Dembinski TC, Mantsch HH. Rapid determination of fetal lung maturity from infrared spectra of amniotic fluid. Am J Obstet Gynecol. 1998;178: 234–241. doi:10.1016/S0002-9378(98)80006-9

71. Liu K-Z, Dembinski TC, Mantsch HH. Prediction of RDS from amniotic fluid analysis: a comparison of the prognostic value of TLC and infra-red spectroscopy. Prenat Diagn. 1998;18: 1267–1275. doi:10.1002/(SICI)1097-0223(199812)18:12<1267::AID-PD441>3.0.CO;2-7

72. Graça G, Moreira AS, Correia AJ V., Goodfellow BJ, Barros AS, Duarte IF, et al. Mid-infrared (MIR) metabolic fingerprinting of amniotic fluid: A possible avenue for early diagnosis of prenatal disorders? Anal Chim Acta. 2013;764: 24–31. doi:10.1016/j.aca.2012.12.023

73. Mukherjee R, Ray CD, Ray S, Dasgupta S, Chaudhury K. Altered metabolic profile in early and late onset preeclampsia: An FTIR spectroscopic study. Pregnancy Hypertens. 2014;4: 70–80. doi:10.1016/j.preghy.2013.11.005

74. Liu KZ, Ahmed MK, Dembinski TC, Mantsch HH. Prediction of fetal lung maturity from near-infrared spectra of amniotic fluid. Int J Gynecol Obstet. 1997;57: 161–168. doi:10.1016/S0020-7292(97)02892-0

75. Power KM, Sanchez-Galan JE, Luskey GW, Koski KG, Burns DH. Use of Near-Infrared Spectroscopic Analysis of Second Trimester Amniotic Fluid to Assess Preterm Births. J Pregnancy. 2011;2011: 1–6. doi:10.1155/2011/980985

76. Araujo-Andrade C, Pichardo-Molina JL, Barbosa-Sabanero G, Frausto-Reyes C, Torres-López A. Detection of the presence of antibodies against Toxoplasma gondii in human colostrum by Raman spectroscopy and principal component analysis. J Biomed Opt. 2007;12: 034006. doi:10.1117/1.2749740

77. Chen SJ, Zhang Y, Ye XP, Hu K, Zhu MF, Huang YY, et al. Study of the molecular variation in pre-eclampsia placenta based on micro-Raman spectroscopy. Arch Gynecol Obstet. 2014;290: 943–946. doi:10.1007/s00404-014-3282-9

78. Kim W, Lee SH, Kim JH, Ahn YJ, Kim YH, Yu JS, et al. Paper-Based Surface-Enhanced Raman Spectroscopy for Diagnosing Prenatal Diseases in Women. ACS Nano. 2018;12: 7100–7108. doi:10.1021/acsnano.8b02917

79. Kao YC, Han X, Lee YH, Lee HK, Phan-Quang GC, Lay CL, et al. Multiplex Surface-Enhanced Raman Scattering Identification and Quantification of Urine Metabolites in Patient Samples within 30 min. ACS Nano. 2020;14: 2542–2552. doi:10.1021/acsnano.0c00515

80. Scott DA, Renaud DE, Krishnasamy S, Meriç P, Buduneli N, Çetinkalp Ş, et al. Diabetes-related molecular signatures in infrared spectra of human saliva. Diabetol Metab Syndr. 2010;2: 1–9. doi:10.1186/1758-5996-2-48

81. Guang P, Huang W, Guo L, Yang X, Huang F, Yang M, et al. Blood-based FTIR-ATR spectroscopy coupled with extreme gradient boosting for the diagnosis of type 2 diabetes. Medicine (Baltimore). 2020;99: e19657. doi:10.1097/md.0000000000019657

82. Brown CD, Davis HT, Ediger MN, Fleming CM, Hull EL, Rohrscheib M. Clinical assessment of near-infrared spectroscopy for noninvasive diabetes screening. Diabetes Technol Ther. 2005;7: 456–466. doi:10.1089/dia.2005.7.456

83. Samant P, Agarwal R. Machine learning techniques for medical diagnosis of diabetes using iris images. Comput Methods Programs Biomed. 2018;157: 121–128. doi:10.1016/j.cmpb.2018.01.004

84. González-Solís JL, Villafan-Bernal JR, Martínez-Zérega BE, Sánchez-Enríquez S. Type 2 diabetes detection based on serum sample Raman spectroscopy. Lasers Med Sci. 2018;33: 1791–1797. doi:10.1007/s10103-018-2543-4

85. Guevara E, Torres-Galván JC, Ramírez-Elías MG, Luevano-Contreras C, González FJ. Use of Raman spectroscopy to screen diabetes mellitus with machine learning tools. Biomed Opt Express. 2018;9: 4998. doi:10.1364/boe.9.004998

86. Wang L, Liu GD, Mu X, Xiao H Bin, Qi C, Zhang SQ, et al. Red blood cells raman spectroscopy comparison of type two diabetes patients and rats. Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy Spectr Anal. 2015;35: 2776–2780. doi:10.3964/j.issn.1000-0593(2015)10-2776-05

87. Roman M, Kamińska A, Drożdż A, Platt M, Kuźniewski M, Małecki MT, et al. Raman spectral signatures of urinary extracellular vesicles from diabetic patients and hyperglycemic endothelial cells as potential biomarkers in diabetes. Nanomedicine Nanotechnology, Biol Med. 2019;17: 137–149. doi:10.1016/j.nano.2019.01.011

88. Wang L, Liu GD, Mu X, Xiao HB, Qi C, Zhang SQ, et al. Red Blood Cells Raman Spectroscopy Comparison of Type Two Diabetes Patients and Rats. Guang Pu Xue Yu Guang Pu Fen Xi. 2015;35: 2776–2780. doi:10.3964/j.issn.1000-0593(2015)10-2776-05

89. Mennickent D, Ortega-Contreras B, Gutiérrez-Vega S, Castro E, Rodríguez A, Araya J, et al. Evaluation of first and second trimester maternal thyroid profile on the prediction of gestational diabetes mellitus and post load glycemia. PLoS One. 2023;18: e0280513. doi:10.1371/journal.pone.0280513

90. Correa PJ, Venegas P, Palmeiro Y, Albers D, Rice G, Roa J, et al. First trimester prediction of gestational diabetes mellitus using plasma biomarkers: a case-control study. J Perinat Med. 2019;47: 161–168. doi:10.1515/jpm-2018-0120

91. Chaparro A, Zúñiga E, Varas-Godoy M, Albers D, Ramírez V, Hernández M, et al. Periodontitis and placental growth factor in oral fluids are early pregnancy predictors of gestational diabetes mellitus. J Periodontol.

2018;89: 1052–1060. doi:10.1002/JPER.17-0497

92. Gromski PS, Muhamadali H, Ellis DI, Xu Y, Correa E, Turner ML, et al. A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. Anal Chim Acta. 2015;879: 10–23. doi:10.1016/j.aca.2015.02.012

93. Wu Y-TT, Zhang C-JJ, Mol BW, Kawai A, Li C, Chen L, et al. Early Prediction of Gestational Diabetes Mellitus in the Chinese Population via Advanced Machine Learning. J Clin Endocrinol Metab. 2021;106: e1191–e1205. doi:10.1210/clinem/dgaa899

94. Beć KB, Grabska J, Huck CW. Near-Infrared Spectroscopy in Bio-Applications. Molecules. 2020;25: 2948. doi:10.3390/molecules25122948

95. Pasquini C. Near infrared spectroscopy: A mature analytical technique with new perspectives – A review. Anal Chim Acta. 2018;1026: 8–36. doi:10.1016/j.aca.2018.04.004

96. Dudzik D, Barbas-Bernardos C, García A, Barbas C. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. J Pharm Biomed Anal. 2018;147: 149–173. doi:10.1016/j.jpba.2017.07.044

97. Cocchi M, Biancolillo A, Marini F. Chemometric Methods for Classification and Feature Selection. Comprehensive Analytical Chemistry. Elsevier; 2018. pp. 265–299. doi:10.1016/bs.coac.2018.08.006

98. Post MJ, van der Putten P, van Rijn JN. Does Feature Selection Improve Classification? A Large Scale Experiment in OpenML. Lecture Notes in Computer Science. Springer; 2016. pp. 158–170. doi:10.1007/978-3-319-46349-0_14

7. ANEXO

7.1. Productividad durante el programa de doctorado

7.1.1. Participación en artículos científicos

7.1.1.1. Productos de tesis doctoral

(En preparación). Mennickent D, Marini F, Romero-Albornoz L, Araya J, Guzmán-Gutiérrez E. Near-infrared spectroscopy-based machine learning models as a simple and fast tool for the prediction of gestational diabetes mellitus at different stages of pregnancy.

Mayo 2023 (Publicado). Mennickent D, Rodríguez A, Riedel CA, Castro E, Eriz-Salinas A, Appel-Rubio J, Aguayo C, Opazo MC, Damiano AE, Guzmán-Gutiérrez E, Araya J. Machine learning applied in maternal and fetal health: A narrative review focused on pregnancy diseases and complications, Frontiers in Endocrinology, 14:1130139, https://doi.org/10.3389/fendo.2023.1130139

Enero 2023 (Publicado). Mennickent D, Ortega-Contreras B, Gutiérrez-Vega S, Castro E, Rodríguez A, Araya J, Guzmán-Gutiérrez E. Evaluation of first and second trimester maternal thyroid profile on the prediction of gestational diabetes

mellitus and post load glycemia, PLOS ONE, 18(1), e0280513, https://doi.org/10.1371/journal.pone.0280513

Agosto 2022 (Publicado). Mennickent D, Rodríguez A, Farías-Jofré M, Araya J, Guzmán-Gutiérrez E. Machine learning-based models for gestational diabetes mellitus prediction before 24-28 weeks of pregnancy: a review, Artificial Intelligence in Medicine, 132, 102378, https://doi.org/10.1016/j.artmed.2022.102378

7.1.1.2. Relacionadas con tema de tesis doctoral

(En preparación). Appel-Rubio J, Mennickent D, Rodríguez A, Damiano AE, Guzmán-Gutiérrez E, Araya J. Early prediction of gestational diabetes, preterm delivery and macrosomia using clinical and biochemical data and machine learning.

Abril 2022 (Publicado). Ortega-Contreras B, Armella A, Appel J, Mennickent D, Araya J, González M, Castro E, Obregon AM, Lamperti L, Gutiérrez J and Guzmán-Gutiérrez E. Pathophysiological role of genetic factors associated with gestational diabetes mellitus, Frontiers in Physiology, 13, 769924. https://doi.org/10.3389/fphys.2022.769924

Enero 2021 (Publicado). Araya J, Rodríguez A, Lagos-San Martín K, Mennickent D, Gutiérrez-Vega S, Ortega-Contreras B, Valderrama-Gutiérrez B, González M, Farías-Jofré M, Guzmán-Gutiérrez E. Maternal thyroid profile in first and second trimester of pregnancy is correlated with gestational diabetes mellitus through machine learning, Placenta, 103, 82–85. https://doi.org/10.1016/j.placenta.2020.10.015

Noviembre 2020 (Publicado). Gutiérrez-Vega S, Armella A, Mennickent D, Loyola M, Covarrubias A, Ortega-Contreras B, Escudero C, Gonzalez M, Alcalá M, Ramos MP, Viana M, Castro E, Leiva A, Guzmán-Gutiérrez E. High levels of maternal total tri-iodothyronine, and low levels of fetal free L-thyroxine and total tri-iodothyronine, are associated with altered deiodinase expression and activity in placenta with gestational diabetes mellitus, PLoS ONE, 15(11), e0242743. https://doi.org/10.1371/journal.pone.0242743

7.1.1.3. Relacionadas con el programa doctoral

(En preparación). Tiznado-Fariña MF, Álvez G, Fuentes-Ravanal F, Mennickent D, Rodríguez A, Guzmán-Gutiérrez E, Fuentealba J, Araya J. In-situ identification of different amyloid-β peptide conformational states in murine brain via hyperspectral imaging and machine learning.

(En preparación). Álvez G, Tiznado-Fariña MF, Fuentes-Ravanal F, Mennickent D, Rodríguez A, Guzmán-Gutiérrez E, Fuentealba J, Araya J. Recognition of spectral patterns associated with different amyloid-β peptide aggregation states and their correlation with cellular toxicity.

(En preparación). Nova D, Olivares L, Mennickent D, Sáez-Orellana F, Neira JY, Vallejos A, Radojkovic C, Bustamante L, Fuentealba J, Mardones C. Aqueous infusion of Berberis microphylla (Calafate) leaves, a new functional beverage: characterization by UHPLC-QTOF-MS/MS, GC-ESI-MS/MS, TXRF and in vitro bioactivity.

Enero 2023 (Publicado). Olivares-Caro L, Nova-Baza D, Radojkovic C, Bustamante L, Duran D, Mennickent D, Melin V, Contreras D, Pérez AJ, Mardones C. Berberis microphylla G. Forst intake reduces the cardiovascular disease plasmatic markers associated to a high fat diet in a mice model, Antioxidants, 12(2), 304, https://doi.org/10.3390/antiox12020304

Noviembre 2021 (Publicado). Mennickent D, Castillo RP, Araya J, Neira JY. Analytical performance of Compton/Rayleigh signal ratio by total reflection X-ray fluorescence (TXRF): A potential methodological tool for sample differentiation, X-Ray Spectrometry, 51(2), 142-150. https://doi.org/10.1002/xrs.3273

7.1.1.4. Productos de otras colaboraciones

(En preparación). Bliem L, Pipper C, León L, Mennickent D, Wagner D, Bodner C, Guzmán-Gutiérrez E, Stingl M, Untersmayr-Elsenhuber E, Sepúlveda N, Bertinat R, Westermeier F. ME/CFS patients stratified by sex and severity exhibit different steroid hormone profiles.

(En preparación). Muñoz HC, Araya J, Llanos AJ, Guzmán-Gutiérrez E, Herrera EA, Mennickent D, Reyes RV, Diaz M, Ebensperger G, Rodriguez AI. Identification of new cardiopulmonary variables in ovine neonatal pulmonary hypertension development by machine learning.

Octubre 2021 (Publicado). Godoy PA, Mennickent D, Cuchillo-Ibáñez I, Ramírez-Molina O, Silva-Grecchi T, Panes-Fernández J, Castro P, Sáez-Valero J, Fuentealba J. Increased P2X2 receptors induced by amyloid-β peptide participate in the neurotoxicity of Alzheimer's Disease, Biomedicine & Pharmacotherapy, 142, 111968. https://doi.org/10.1016/j.biopha.2021.111968

7.1.2. Presentaciones en congresos

7.1.2.1. Internacionales

Octubre 2022. III Autumn Meeting for Young Chemists in Biomedical Sciences. Nápoles, Italia. Presentación oral: Prediction of gestational diabetes mellitus with near-infrared-based chemometrics models. Mennickent D, Rodríguez A, Araya J, Guzmán-Gutiérrez E.

Mayo 2022. IX Latin American Society for Materno Fetal Interaction and Placenta Meeting. Modalidad en línea. Póster: Prediction of gestational diabetes mellitus with machine learning techniques: a comparison between near-infrared spectra and maternal data-based models. Mennickent D, Rodríguez A, Araya J, Guzmán-Gutiérrez E.

Diciembre 2021. 1º Encuentro Virtual de la Red Iberoamericana de Alteraciones Vasculares en trastornos del Embarazo: Fomentando nuestros vínculos. Modalidad en línea. Presentación oral: Variables maternas del primer trimestre del embarazo para la predicción de diabetes gestacional. Mennickent D, Rodríguez A, Araya J, Guzmán-Gutiérrez E.

Agosto 2021 – septiembre 2021. International Federation of Placenta Associations 2021 Meeting. Modalidad en línea. Póster: First trimester maternal variables as potential predictors for gestational diabetes mellitus. Mennickent D, Rodríguez A, Araya J, Guzmán-Gutiérrez E.

Julio 2021. Road to Chemometrics in Analytical Chemistry 2022 Conference. Modalidad en línea. Presentación oral: Chemometrics-based identification and assessment of maternal thyroid variables as a potential tool to improve gestational diabetes mellitus diagnosis. Mennickent D, Rodríguez A, Araya J, Guzmán-Gutiérrez E.

Mayo 2021. International Symposium on Reproductive Health 2021. Modalidad en línea. Presentación oral: Machine learning reveals maternal thyroid variables as a potential tool to improve the diagnosis of gestational diabetes mellitus. Mennickent D, Rodríguez A, Araya J, Guzmán-Gutiérrez E.

Agosto 2020. III Simposio Colombiano de Placenta e Interacción Materno-Fetal. Modalidad en línea. Póster: Variables gineco-obstétricas del primer trimestre de gestación para el potencial diagnóstico precoz de diabetes gestacional. Mennickent D, Araya J, Guzmán-Gutiérrez E.

7.1.2.2. Nacionales

Agosto 2022. XX Congreso Chileno de Química Clínica y Ciencias de Laboratorio. Santiago, Chile. Póster: Predicción de glicemia post carga en el segundo trimestre de gestación con datos de anamnesis y laboratorio del primer trimestre de embarazo. Mennickent D, Rodríguez A, Araya J, Guzmán-Gutiérrez E.

Agosto 2022. IV Vascular Health Research and Innovation Group Anual Meeting. Chillán, Chile. Póster: Prediction of gestational diabetes mellitus with machine learning techniques: a comparison between mid-infrared and medical data based-models. Mennickent D, Rodríguez A, Araya J, Guzmán-Gutiérrez E.

Enero 2020. XXXIII Jornadas Chilenas de Química. Puerto Varas, Chile. Póster: Comparación de métodos de extracción para análisis metabolómico no dirigido en plasma por UHPLC-ESI-QTOF-MS. Mennickent D, Olivares L, Bustamante L, Mardones C.

7.1.3. Recursos adjudicados durante el programa de doctorado

Marzo 2023. Extensión de Beca de Doctorado “COVID-19”. Agencia Nacional de Investigación y Desarrollo, Chile.

Septiembre 2022. Fondo “Apoyo a pasantía internacional de doctorado UCO 1866”. Universidad de Concepción y Ministerio de Educación, Chile.

Julio 2022. Beneficio complementario “Pasantía doctoral en el extranjero”. Agencia Nacional de Investigación y Desarrollo, Chile.

Septiembre 2021. Beneficio complementario “Gastos operacionales del proyecto de tesis doctoral”. Agencia Nacional de Investigación y Desarrollo, Chile.

Junio 2021. Fondo “Apoyo de asistencia a cursos de especialización de corta duración en sociedades científicas o instituciones académicas en el extranjero UCO 1866”. Universidad de Concepción y Ministerio de Educación, Chile.

Marzo 2019. Beca “Doctorado Nacional”, folio 21190736. Agencia Nacional de Investigación y Desarrollo, Chile.

7.2. Otras actividades relevantes durante el programa de doctorado

7.2.1. Formación internacional

Octubre 2022 – enero 2023. Pasantía en Departamento de Química, Facultad de Ciencias Matemáticas, Físicas y Naturales. Sapienza Università di Roma, Italia.

Junio 2021 – julio 2021. Curso de Postgrado “La placenta: Origen de las enfermedades crónicas del adulto”, de 80 horas teórico-prácticas, modalidad en línea. Facultad de Farmacia y Bioquímica, Universidad de Buenos Aires, Argentina.

Octubre 2020 – noviembre 2020. “International School of Chemometrics”, de 150 horas teórico-prácticas, modalidad en línea. Faculty of Science, University of Copenhagen, Dinamarca.

7.2.2. Participación en proyectos de investigación

Diciembre 2021 – abril 2023. Tesista Doctoral en Proyecto ANID FOVI 210057 “Fortalecimiento de una alianza de cooperación internacional chilena-argentina para el estudio de patologías del embarazo en la región del Bío-Bío utilizando inteligencia artificial”.

Marzo 2021 – a la fecha. Tesista Doctoral en Proyecto ANID FONDECYT 11181153 “Chemical mapping of different assemblies states of beta amyloid peptide in Alzheimer´s disease transgenic mouse brain using hyperspectral imaging and chemometrical analysis”.

Marzo 2020 – diciembre 2020. Tesista Doctoral en Proyecto ANID FONDECYT 11170710 “Maternal thyroid profile characterization during pregnancy: Association with gestational diabetes, altered metabolism thyroid in placenta and neonatal thyroid profile”.

Marzo 2019 – marzo 2020. Estudiante Doctoral en Proyecto ANID FONDECYT 1191276 “A combined analytical strategy based on metabolomics and targeted analysis using UHPLC-QTOF-MS, UHPLC-ESI-MS/MS and GC-EI-MS/MS for the study of health beneficial effects of calafate consumption”.

7.2.3. Participación en grupos de investigación

Agosto 2020 – a la fecha. Miembro del grupo de investigación “Machine Learning Applied in Biomedicine” (MLAB). Universidad de Concepción y Universidad del Bio-Bío, Chile.

7.2.4. Participación en docencia

Enero 2022. Charla “Predicción de diabetes gestacional en el primer trimestre del embarazo: una aplicación de machine learning sobre datos clínicos y bioquímicos maternos” en el Taller “Herramientas de machine learning aplicadas a

biomedicina” de la Escuela de Verano UdeC 2022. Universidad de Concepción, Chile.

Julio 2019 – Diciembre 2022. Colaboración docente en prácticos de Análisis Instrumental, Carreras de Bioingeniería y Biología Marina. Universidad de Concepción, Chile.

7.2.5. Participación en comunicación de la ciencia

Agosto 2022. Charla “Técnicas de machine learning para la predicción de diabetes gestacional” en el tercer seminario web del Proyecto ANID-FOVI 210057, modalidad en línea.

7.2.6. Participación en representación universitaria

Agosto 2020 – junio 2022. Representante estudiantil del programa de Doctorado en Ciencias y Tecnología Analítica. Universidad de Concepción, Chile.

Septiembre 2019 y 2022. Representante del programa de Doctorado en Ciencias y Tecnología Analítica en Feria de Postgrado. Universidad de Concepción, Chile.