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 , Lucas Romero-Albornoz , Sebastián Gutiérrez-Vega , Claudio Aguayo , Federico Marini , Enrique Guzmán-Gutiérrez * , Juan Araya *

Posted Date: 29 March 2024

doi: 10.20944/preprints202403.1737.v1

Keywords: Obstetrics and gynecology; Vibrational spectroscopy; Bioanalytical chemistry; Artificial intelligence

Article

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], Lucas Romero-Albornoz [2], Sebastián Gutiérrez-Vega [4], Claudio Aguayo [4], Federico Marini [5], Enrique Guzmán-Gutiérrez [3,4,*] and Juan Araya [2,3,*]

[1] Departamento de Ciencias Básicas y Morfología, Facultad de Medicina, Universidad Católica de la Santísima 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 Bioquímica Clínica e Inmunología, Facultad de Farmacia, Universidad de Concepción, Concepción, Chile
[5] Department of Chemistry, University of Rome La Sapienza, Rome, Italy.
* Correspondence: eguzman@udec.cl (E.G.-G.); jarayaq@udec.cl (J.A.); Tel.: +56 41 2203021 (E.G.-G.); +56 41 2203749 (J.A.)

Abstract: Gestational diabetes mellitus (GDM) is a hyperglycemia state that is typically diagnosed by an oral glucose tolerance test (OGTT), which is unpleasant, time-consuming, poorly reproducible, and tardy. The machine learning (ML) predictive models that have been proposed to improve GDM diagnosis are usually based on instrumental methods that take hours to get a result. Near-infrared (NIR) spectroscopy, a simple, fast and low-cost analytical technique has never been assessed for the prediction of GDM. 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. Serum samples from the first (before GDM diagnosis) and the second (at the time of GDM diagnosis) trimester of pregnancy were analyzed by NIR spectroscopy. Four spectral ranges were considered, and 80 pretreatments were tested for each. NIR data-based models were built with single- and multi-block ML techniques. Every model was subjected to double cross-validation. The best first and second trimester models got areas under the receiver operating characteristic curve of 0.5768 ± 0.0635 and 0.8836 ± 0.0259, respectively. This is the first study reporting NIR spectroscopy-based methods for the prediction of GDM. The developed methods allow to predict GDM from 10 µL of serum in only 32 minutes. They 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.

Keywords: Obstetrics and gynecology; vibrational spectroscopy; bioanalytical chemistry; artificial intelligence

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

 doi:10.20944/preprints202403.1737.v1

(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 12,500 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.

2. Materials and methods

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.

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.

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.

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.

 doi:10.20944/preprints202403.1737.v1

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.

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) (medical and NIR data are displayed in **Dataset S1A** and **Dataset S1B,** respectively). 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 (medical and NIR data are presented in **Dataset S2A** and **Dataset S2B,** respectively).

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).

2.8. ML Analyses

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.

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 sequential and 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).

 doi:10.20944/preprints202403.1737.v1

2.8.3. Evaluation of Predictive Performance

Models' predictive performance was evaluated by means of their specificity (Sp), sensitivity (Se) and non-error rate (NER). For the best models, the area under the receiver operating characteristic curve (AUROC) was also determined. The mathematical definition of these parameters can be found elsewhere [25,26]. In general terms, the specificity and the sensitivity denote the ability to correctly classify NGT and GDM subjects, respectively; the NER reflects the ability to correctly classify both NGT and GDM subjects; and the AUROC 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.

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.

3. Results

3.1. First Trimester Cohort

3.1.1. Description of the First Trimester Cohort

To characterize this cohort, classical statistical techniques were used. **Table 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 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/m²	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)

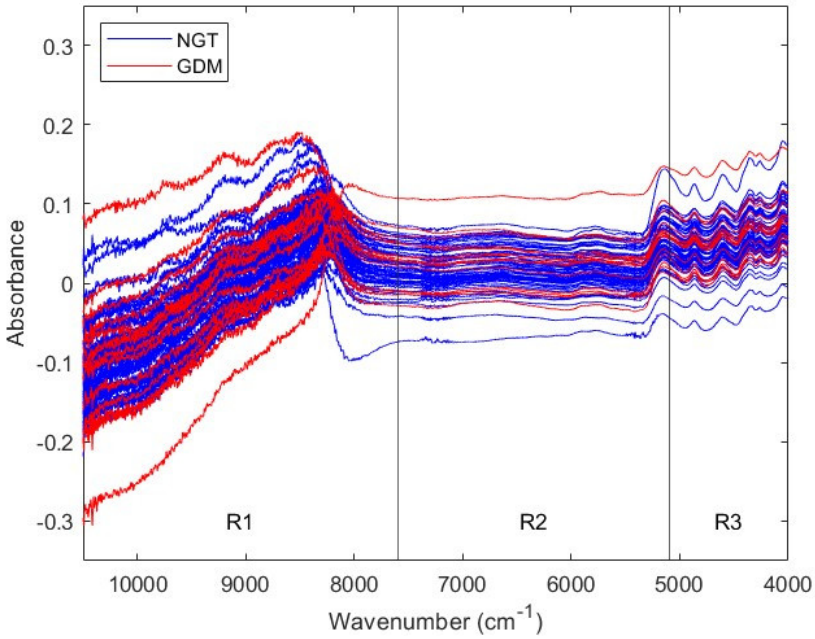
 doi:10.20944/preprints202403.1737.v1

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.

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. **Fig. 1** shows NIR spectra from first trimester NGT and GDM sera. The spectral traces behavior depends on the wavenumber range. In particular, signal sequential noise varies with wavenumber. There is a high-noise region between 10,500 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 10,500 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 10,500 to 4000 cm^{-1} (Full).



 doi:10.20944/preprints202403.1737.v1

Figure 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 10,500 to 4000 cm^{-1}; range 1, from 10,500 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 S1-4**). **Table 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.

Table 2. Predictive performance of the best ML models using NIR spectral data from first trimester serum samples.

Range [a]	Model number	Pretreatment	Sp		Se		NER	
			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}. Sp: specificity; Se: sensitivity; NER: non-error rate; Av: average; StD: standard deviation; R1: range 1; R2: range 2; R3: range 3; SM: smoothing; W: width; N: normalization; MC: mean centering.

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 S5**). The addition of medical data does 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.

Figure 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 AUROC 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).

 doi:10.20944/preprints202403.1737.v1

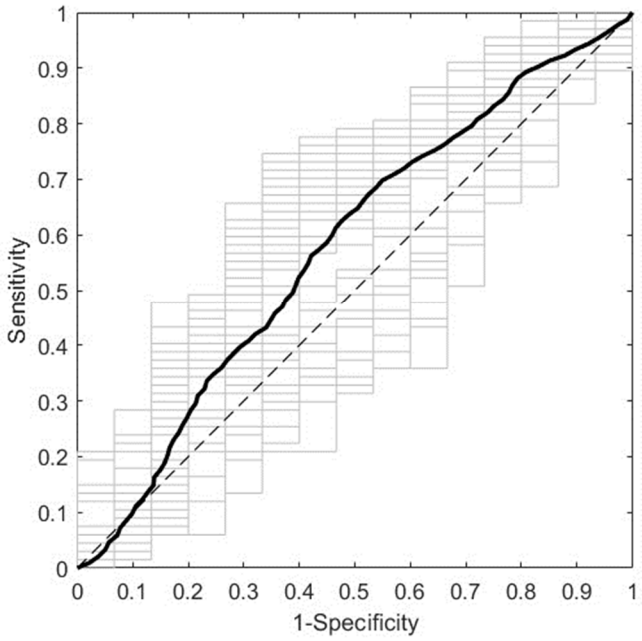


Figure 2. ROC curve of the best predictive model for the first trimester cohort. The model was trained with NIR spectra (R1, 10500-7600 cm^{-1}) from first trimester serum samples, after pretreatment by normalization and mean centering. The average and the individual curves of 50 DCV repetitions are colored in black and gray, respectively. *ROC: receiver operating characteristic; NIR: near-infrared; R1: range 1; DCV: double cross-validation.*

3.2. Second Trimester Cohort

3.2.1. Description of the Second Trimester Cohort

To characterize this cohort, classical statistical analyses were performed. **Table 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 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)

 doi:10.20944/preprints202403.1737.v1

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.

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. **Fig. 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}).

 doi:10.20944/preprints202403.1737.v1

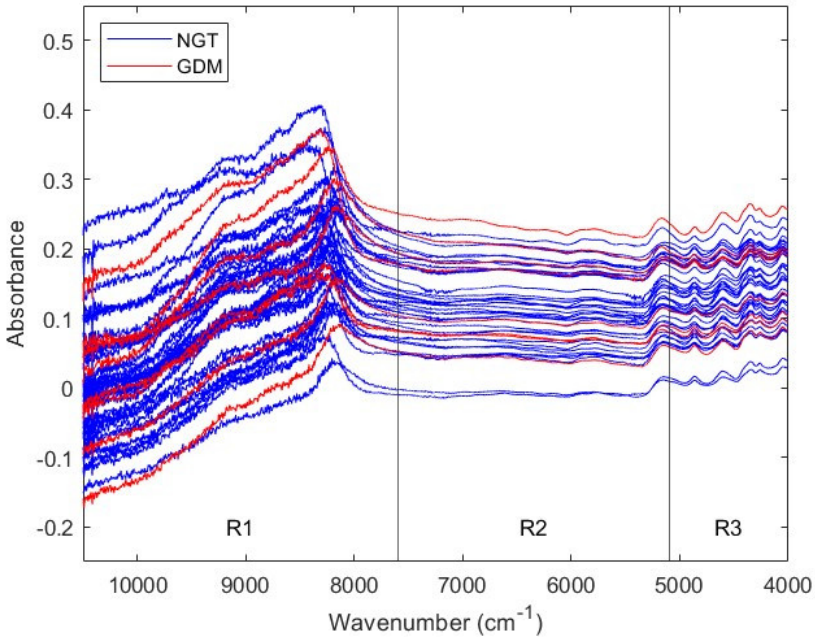


Figure 3. NIR spectra from second trimester serum samples. Spectra from NGT and GDM pregnant women are colored in blue and red, respectively. Four spectral ranges are considered: full, from 10,500 to 4000 cm^{-1}; range 1, from 10,500 to 7600 cm^{-1}; range 2, from 7600 to 5100 cm^{-1}; and range 3, 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 S6-9**). **Table 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.

Table 4. Predictive performance of the best ML models using NIR spectral data from second trimester serum samples.

Range [a]	Model number	Pretreatment	Sp		Se		NER	
			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}. Sp: specificity; Se: sensitivity; NER: non-error rate; Av: average; StD: standard deviation; R1: range 1; R2: range 2; R3: range 3; 2D: second derivative; W: width; N: normalization; MC: mean centering; WLS: weighted least squares; 1D: first derivative.

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 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

 doi:10.20944/preprints202403.1737.v1

R3 data only. The simplification of the multi-block models by means of variable selection does 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.

Figure 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 AUROC 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).

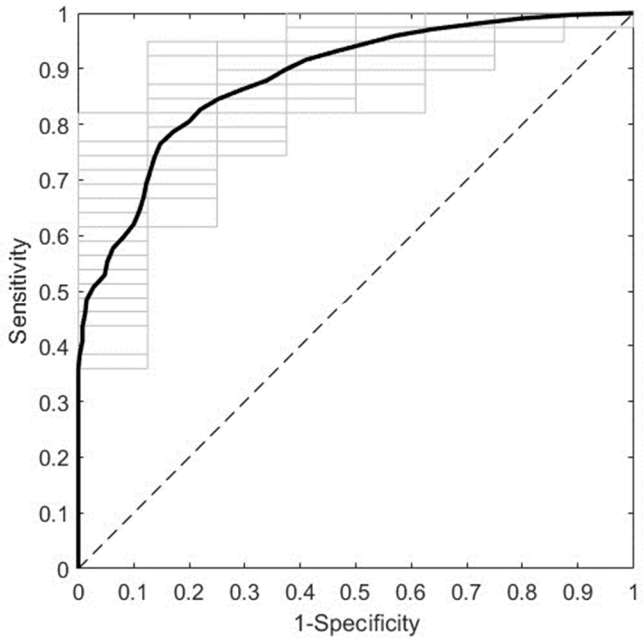


Figure 4. ROC curve of the best predictive model for the second trimester cohort. The model was trained with NIR spectra (R3, 5100-4000 cm^{-1}) from second trimester serum samples, after pretreatment by first derivative (width=15) and mean centering. The average and the individual curves of 50 DCV repetitions are colored in black and gray, respectively. *ROC: receiver operating characteristic; NIR: near-infrared; R3: range 3; DCV: double cross-validation.*

4. 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.

4.1. 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 1** and **Table 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 1** and **Table 3**). Higher frequencies or levels of these variables were observed in pregnant women with GDM. This observation makes sense, since they are known

 doi:10.20944/preprints202403.1737.v1

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 S5** and **Table 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].

4.2. 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 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**) 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.

4.3. 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 that obtained with the Full range, of 0.5726 ± 0.0410 (**Table 2**). In the case of second trimester serum samples, R3 exhibited a NER of 0.7894 ± 0.0431, much higher than the one obtained with the Full range, of 0.4642 ± 0.0321 (**Table 4**). This tendency was maintained when medical data was added (**Table S5** and **Table 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

 doi:10.20944/preprints202403.1737.v1

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 (**Fig. 1** and **Fig. 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 2** and **Table 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 2**) and 5100-4000 cm^{-1} (**Table 4**), respectively. They were associated with AUROCs of 0.5768 ± 0.0635 (**Fig. 2**) and 0.8836 ± 0.0259 (**Fig. 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 S11** and **Table 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 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.

4.5. 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 AUROCs 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.

 doi:10.20944/preprints202403.1737.v1

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 AUROCs 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 AUROC 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 AUROCs of 60.0-66.9% [49] before GDM diagnosis, and AUROCs 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 (AUROC of 0.5768 ± 0.0635, **Fig. 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 (AUROC of 0.8836 ± 0.0259, **Fig. 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.

4.6. 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.

4.7. 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

 doi:10.20944/preprints202403.1737.v1

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.

4.8. 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.

5. Conclusions

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.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org

Author Contributions: Conceptualization, Enrique Guzmán-Gutiérrez and Juan Araya; Data curation, Daniela Mennickent; Formal analysis, Daniela Mennickent; Funding acquisition, Daniela Mennickent, Enrique Guzmán-Gutiérrez and Juan Araya; Investigation, Daniela Mennickent and Sebastián Gutiérrez-Vega ; Methodology, Daniela Mennickent; Project administration, Enrique Guzmán-Gutiérrez and Juan Araya; Resources, Claudio Aguayo, Federico Marini, Enrique Guzmán-Gutiérrez and Juan Araya; Software, Federico Marini; Supervision, Enrique Guzmán-Gutiérrez and Juan Araya; Validation, Daniela Mennickent and Federico Marini; Visualization, Daniela Mennickent and Lucas Romero-Albornoz; Writing – original draft, Daniela Mennickent and Lucas Romero-Albornoz; Writing – review & editing, Sebastián Gutiérrez-Vega , Claudio Aguayo, Federico Marini, Enrique Guzmán-Gutiérrez and Juan Araya.

Funding: This work was financially supported by Agencia Nacional de Investigación y Desarrollo through PhD scholarship 21,190,736 (DM), FONDECYT 11,181,153 (JA), and FOVI 210,057 (EG-G); Universidad de Concepción/Ministerio de Educación via Proyecto UCO 1866 (DM); and Gobierno Regional del Bío-Bío by means of FIC-R “Capital Humano Avanzado en Inteligencia Artificial para el Bío-Bío” (DM).

Acknowledgments: The authors would like to thank the volunteers who participated in this study; as well as the health staff of CESFAM Victor Manuel Fernández, Santa Sabina, and Tucapel; and Joanna Tereszczuk from CMA Bio-Bío, for their technical support.

Conflicts of Interest: The authors have no relevant financial or non-financial interests to disclose.

References

1. McIntyre, H.D.; Catalano, P.; Zhang, C.; Desoye, G.; Mathiesen, E.R.; 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, H.R.; Ross, G.P. 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.; Dagklis, T. 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, M.M.; Punnose, J.; Dhatt, G.S. 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, E.H.; Fox, R.A.; Dennison, R.A.; Usher-Smith, J.A.; Meek, C.L.; Aiken, C.E. 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, F.W.F.; Peters, J.R. 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, E.A.; Rossi, S.W.; Hoesli, I.; Göbl, C.S. 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, S.W.; Coetzee, E.J. 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, Y.N.; Noubiap, J.J.; Danwang, C.; Sama, J.D.; Azabji-Kenfack, M.; Mbanya, J.C.; Sobngwi, E. 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, H.R.; Smith, G.C.S. 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.; Georgiopoulos, G.; Nicolaides, K.H.; Charakida, M. 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, L.M.; Gray, J.W. 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; ISBN 978-981-15-8647-7.
16. Bunaciu, A.A.; Fleschin, Ş.; Hoang, V.D.; Aboul-Enein, H.Y. Vibrational Spectroscopy in Body Fluids Analysis. *Crit Rev Anal Chem* **2017**, *47*, 67–75, doi:10.1080/10408347.2016.1209104.
17. Balan, V.; Mihai, C.T.; Cojocaru, F.D.; Uritu, C.M.; Dodi, G.; Botezat, D.; Gardikiotis, I. 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. MINSAL *Guía Perinatal*; 2015; ISBN 9789563480764.
21. Morais, C.L.M.; Lima, K.M.G.; Singh, M.; Martin, F.L. 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. In *Comprehensive Analytical Chemistry*; Elsevier, 2018; Vol. 82, pp. 265–299 ISBN 9780444640444.
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.; 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* **2023**, *18*, e0280513, doi:10.1371/journal.pone.0280513.

27. Garmendia, M.L.; Mondschein, S.; Montiel, B.; Kusanovic, J.P. Trends and Predictors of Gestational Diabetes Mellitus in Chile. *Int J Gynecol Obstet* **2020**, *148*, 210–218, doi:10.1002/ijgo.13023.
28. IDF *IDF Diabetes Atlas*; 2019; ISBN 9782930229874.
29. Azcarate, S.M.; Ríos-Reina, R.; Amigo, J.M.; Goicoechea, H.C. Data Handling in Data Fusion: Methodologies and Applications. *TrAC - Trends Anal Chem* **2021**, *143*, doi:10.1016/j.trac.2021.116355.
30. Post, M.J.; van der Putten, P.; van Rijn, J.N. Does Feature Selection Improve Classification? A Large Scale Experiment in OpenML. In *Lecture Notes in Computer Science*; Springer, 2016; pp. 158–170 ISBN 978-3-319-46348-3.
31. Cocchi, M. *Data Fusion Methodology and Applications*; Elsevier, 2019; ISBN 9780444639844.
32. Workman Jr., J.; Weyer, L. *Practical Guide and Spectral Atlas for Interpretive Near-Infrared Spectroscopy*; CRC Press, 2012; ISBN 9780429110511.
33. Beć, K.B.; Grabska, J.; Huck, C.W. 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, N.U.; Liu, X.; Zheng, Q.; Liang, Y.; Zhu, Y.; Iqbal, J.; 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, B.A.; Mojsak, P.; Rojo, D.; Telejko, B.; Paczkowska–Abdulsalam, M.; Hryniewicka, J.; Zielinska–Maciulewska, A.; Szelachowska, M.; Gorska, M.; Barbas, C.; 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.; Zhang, J.; Bai, L.; Wu, H.; Chen, B. 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.; Lorenzo, M.P.; Barbas, C.; Ramos, M.P. 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, S.L.; Poston, L.; Farrar, D.; Sattar, N.; Nelson, S.M.; Wright, J.; Mason, D.; Suderman, M.; 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, L.M.; Martins, A.S.; Duarte, D.; Barros, A.S.; Galhano, E.; Pita, C.; Almeida, M.D.C.; Carreira, I.M.; Gil, A.M. 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, S.O.; Pinto, J.; Graça, G.; Duarte, I.F.; Barros, A.S.; Galhano, E.; Pita, C.; Almeida, M.D.C.; Goodfellow, B.J.; Carreira, I.M.; 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.; Dessì, A.; Fanos, V.; Atzori, L.; Facchinetti, F. 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, H.C.; Ebbels, T.M.D.; Bundy, J.; Holmes, E.; Lindon, J.C.; Nicholson, J.K. 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, X.Y.; Song, L.P.; Zheng, H.T.; Wei, S.D.; Wen, X.L.; Huang, B.; Liu, D. Bin 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, L.A.; Lomelin-Gascon, J.; Ortega-Montiel, J.; Montoya, A.; Mendizabal, L.; Arregi, M.; Martinez-Martinez, M. de L.A.; Camarillo Romero, E.D.S.; 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**.
49. Zhao, C.; Dong, J.; Jiang, T.; Shi, Z.; Yu, B.; Zhu, Y.; Chen, D.; Xu, J.; Huo, R.; Dai, J.; 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, Y.L.; Jia, Y.J.; Xing, B.H.; Shi, D.D.; Dong, X.J. 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, E.M.; Parkin, R.K.; Mitchell, P.S.; 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, D.L.D.; de Morais, C. de L.M.; Cornetta, M. da C. de M.; Camargo, J.D. de A.S.; de Lima, K.M.G.; Crispim, J.C. de O. 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.