

Article

Not peer-reviewed version

Viability of ABO Blood Typing with ATR-FTIR Spectroscopy

[Alfonso Fernández-González](#), [Álvaro Obaya](#), Christian Chimeno-Trinchet, [Tania Fontanil](#),
[Rosana Badía-Laiño](#) *

Posted Date: 31 July 2023

doi: 10.20944/preprints202307.2081.v1

Keywords: FTIR; ATR-FTIR; Linear Discriminant Analysis; LDA; Naïf Bayesian Classifier; NBC; chemometrics



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Viability of ABO Blood Typing with ATR-FTIR Spectroscopy

Alfonso Fernández-González ¹, Álvaro J. Obaya ², Christian Chimeno-Trinchet ¹, Tania Fontanil ³ and Rosana Badía-Laiño ^{1,*}

¹ Department of Physical and Analytical Chemistry, University of Oviedo

² Department of Functional Biology, University of Oviedo

³ Unidad de Terapias Avanzadas, Instituto Ordoñez

* Correspondence: recomol@uniovi.es

Featured Application: Fast ABO blood typing using Attenuated Total Reflectance – Fourier Transform Infrared Spectroscopy and artificial intelligence tools.

Abstract: Fourier Transform Infrared Spectroscopy (FTIR), provides valuable biochemical information for biomedical analysis. It aids in identifying cancerous tissues, diagnosing diseases like acute pancreatitis or Alzheimer's, and has applications in genomics, proteomics, and metabolomics. A combination of FTIR and chemometrics constitute an approach that shows promise in fields like biology, forensics, food quality control, and plant variety identification. The study aims to explore the feasibility of ATR-FTIR spectroscopy for identifying ABO-blood types using spectroscopic tools. We employ various classifying algorithms, including Linear Discriminant Analysis (LDA), Naïve Bayes Classifier (NBC), Principal Component Analysis (PCA), and combinations of these methods, to detect A and B antigens and determine the ABO blood type. The results show that these algorithms predict the blood type to a greater extent than random selection, although they do not match the precision of biochemical blood typing tools. Additionally, our findings suggest a higher sensitivity of the methodology in identifying B antigens compared to A antigens.

Keywords: FTIR; ATR-FTIR; Linear Discriminant Analysis; LDA; Naïf Bayesian Classifier; NBC; chemometrics

1. Introduction

Infrared spectroscopy has emerged as a powerful tool in biomedical analysis due to its ability to provide comprehensive biochemical information about cells, tissues or biological samples. This holistic data obtained from Fourier Transform Infrared Spectroscopy (FTIR), allows for the identification of cancerous tissues [1] or aids in the diagnosis of diseases such as acute pancreatitis of Alzheimer [2,3]. Moreover, vibrational spectroscopy, which encompasses both Raman and infrared spectroscopy, has proven to be highly valuable in genomics [4], proteomics [5] and metabolomics [6]. The latter is particularly interesting, as the combination of metabolomics information provided by FTIR and chemometric tools allows for the investigation of the viability and sexing of cattle embryos [7,8]. This FTIR-chemometrics coupling has also been fruitfully exploited in different fields like forensics, food quality control or plant variety identification [9–11], and seems to be a promising starting point to face a rapid spectroscopic approach to the blood typing following the ABO system.

Blood typing of the ABO system has been routinely used since its description in the early 1900s. It is considered the most important among the known 29 blood group systems [12]. ABO typing is based on the existence of three different alleles whose expressions results in the modifications of polysaccharide structures on red blood cells surface [13]. The combination of these alleles on the human genome gives as a result the four known blood types extensively used in blood transfusion for medical applications. The ABO blood group is thus determined by the presence of A and B antigens (or none) on erythrocytes (and other cells) and the presence of anti-A and anti-B antibodies in serum. Due to the presence of these antibodies, red cell agglutination happens when mixing with

plasma from a different blood type [14]. The ABO system depends on the expression of glycosyltransferases for specific sugar moieties, α 1, 3-N-acetylgalactosaminyltransferase for the A antigen and α 1, 3-galactosyltransferase for the B antigen. Expression of both transferases (one allele each) ends up with the presence of both antigens in red cells surface (AB type). Presence of the H allele in both copies of the genome results in the absence of any glycosyltransferase activity and consequently antigens A or B are not present (O type) [13].

Then, this enzymatic activity translates into the presence of different oligosaccharides in the membrane, whose biochemical difference lays in the nature of the most external monosaccharide. There is a basic sequence of glucose, galactose, N-acetylglucosamine and galactose joined in 1-3 positions, with a fucose in position 2 of the last galactose. This sequence is common for A, B and O antigens; however, A and B shows an additional saccharide: A-type has an N-acetylgalactosamine in position 3 of the galactose, while B-type shows a second galactose in that position.

In this work, we intend to explore the viability of ATR-FTIR spectroscopy to detect the biochemical differences of different blood types, which would eventually allow for the identification of the type only using spectroscopic tools.

2. Materials and Methods

Human blood samples were kindly provided by the *Centro Comunitario de Sangre y Tejidos de Asturias* using blood samples from anonymous donors. All samples were ABO-typed following manufacturer instructions of an immune-based assay of blood typing from *IBDCiencia* (Ref.: ME91253).

FTIR spectra were taken in a Varian 670-IR spectrometer equipped with a Golden Gate ATR device. Spectra were averaged from 16 scans and recorded from 600 cm^{-1} to 4000 cm^{-1} with resolution 4 cm^{-1} . Measurements were carried out by putting a droplet of blood or plasma onto the diamond crystal of ATR and evaporating it under a constant air current.

Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) calculations were performed using MatLab scripts developed by the authors. Naïve Bayesian Classifier (NBC), and k-nearest neighbours calculations were performed using MatLab functions.

3. Results and discussion

The actual blood distribution of Spanish population, according to Red Cross data [15] shows a very low presence of certain types (Table 1). Thus, keeping this proportion in the training and test datasets could lead to poor results because of the low amounts. Therefore, blood samples were selected in order to have a proper amount of each type, regardless of a possible overrepresentation compared to the natural distribution (Table 1).

Table 1. ABO Blood-type distribution and number of samples.

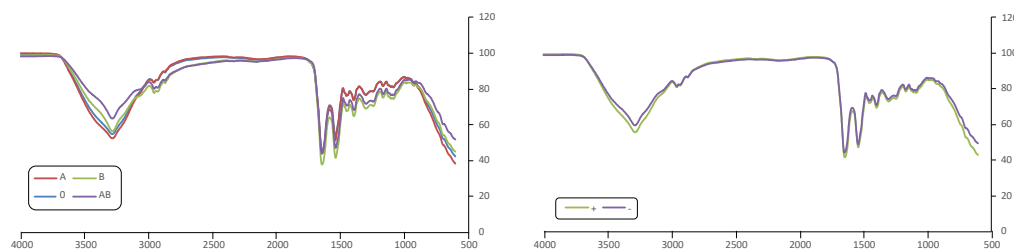
	0+	A+	B+	AB+	0-	A-	B-	AB-
Spanish population	35%	36%	8%	2,5%	9%	7%	2%	0,5%
Number of samples	24 (24%)	21 (21%)	36 (36%)	5 (5%)	5 (5%)	2 (2%)	3 (3%)	4 (4%)

The spectroscopic study was carried out from two different approaches, a first one considering both ABO blood type and Rh factor and a second one taking into account only ABO blood type. Average spectra for every class are depicted in Figure 1, showing no evident differences with the ABO-type or the Rh factor. Thus, the ABO-typing was evaluated from a mathematical approach, using IA tools. For this, the area under the most relevant peaks, total 18, were selected as input variables in the different algorithms (Table 2).

Table 2. FTIR selected peaks.

Region / cm ⁻¹	Region / cm ⁻¹	Region / cm ⁻¹
910 – 945	1213 – 1257	2829 – 2887
980 – 999	1257 – 1348	2887 – 2941
1059 – 1092	1348 – 1425	2941 – 2995
1092 – 1113	1425 – 1477	2995 – 3107
1113 – 1140	1477 – 1583	3107 – 3224
1140 – 1186	1583 – 1716	3224 – 3383

These peaks are mainly related to amide I (1639 cm⁻¹), amide II (1537 cm⁻¹) and amide III (1300 cm⁻¹) bands characteristic of proteins and also to carbohydrates, with strong very wide bands between 3520 cm⁻¹ and 3100 cm⁻¹ [16].

**Figure 1.** FTIR spectra of averaged ABO-type samples (left) and Rh samples (right).

3.1. Probabilities of random classification

In order to check whether the success rate of the proposed classification tools is better than a mere random classification, it is important to compare the true positive (T+), true negative (T-), false positive (F+) and false negative (F-) ratios of the proposed methodology with those obtained in a random classification. In principle, a random classification would assign a random class to a sample, following the frequency distribution of the whole population. That is, the probability of assigning the class *i* to a sample (*p*(*i*)) is the same that the frequency of class *i* in the population (*f*(*i*)). In such a case where *p*(*i*)=*f*(*i*), it is clear that the T+ ratio is $\sum_i \frac{p(i)^2}{N}$ where *p*(*i*) is the probability of assigning *i*-class to a sample and *N* is the total number of classes. Similarly, the rate of false positives and false negatives is identical and equal to $\frac{\sum_{i,j} p(i) \cdot p(j)_{i \neq j}}{N}$. These formulae can be extended for those cases where *p*(*i*)≠*f*(*i*), per example, if assignation into classes is performed equiprobably. Table 3 shows the T+, F+, T- and F- ratios expected for a random classification using the probabilities shown in Table 1 (*p*(A)=0.23, *p*(B)=0.39, *p*(AB)=0.09, *p*(O)=0.29).

Table 3 T+, T-, F+ and F- expected distribution in a random classification with frequency distribution *f* and probability of assignation *p*.

Ratio of	T+	T-	F+	F-
Mathematical expression	$\sum_i \frac{p(i) \cdot f(i)}{N}$	1-T+-F+-F-	$\frac{\sum_{i,j} p(i) \cdot p(j)_{i \neq j}}{N}$	$\frac{\sum_{i,j} p(i) \cdot p(j)_{i \neq j}}{N}$
Considering four classes A, B, AB & O				
With <i>p</i> (<i>i</i>) = <i>f</i> (<i>i</i>) = Table 1	9.7%	59.7%	15.3%	15.3%
Natural frequencies				
With <i>p</i> (<i>i</i>) = <i>f</i> (<i>i</i>) = Table 1	7.4%	57.4%	17.6%	17.6%
Sampling frequencies				
Considering three classes A, B & O				
With <i>p</i> (<i>i</i>) = <i>f</i> (<i>i</i>) = Table 1	13.8%	47.0%	19.6%	19.6%
Natural frequencies				
With <i>p</i> (<i>i</i>) = <i>f</i> (<i>i</i>) = Table 1	11.7%	44.9%	21.7%	21.7%

Sampling frequencies

It is obvious, then, that keeping the assignment probabilities as the true distribution frequencies provides the best T+ and T- ratios. All the classification algorithms will be checked against this random distribution.

3.2. Principal Component Analysis (PCA)

The graphical representation of the samples in the Principal Component (PC) space (Figure 2) does not seem to reveal a clear aggrupation, neither with respect to the ABO-type nor to the Rhesus factor. Nonetheless, A-type (red) and B-type (green) trend to segregate, which can be explained taking into account the different chemical nature of their respective antigens. AB-type appears mixed since it shares both antigens, and O samples, which do not present a specific antigen, seems to be randomly distributed.

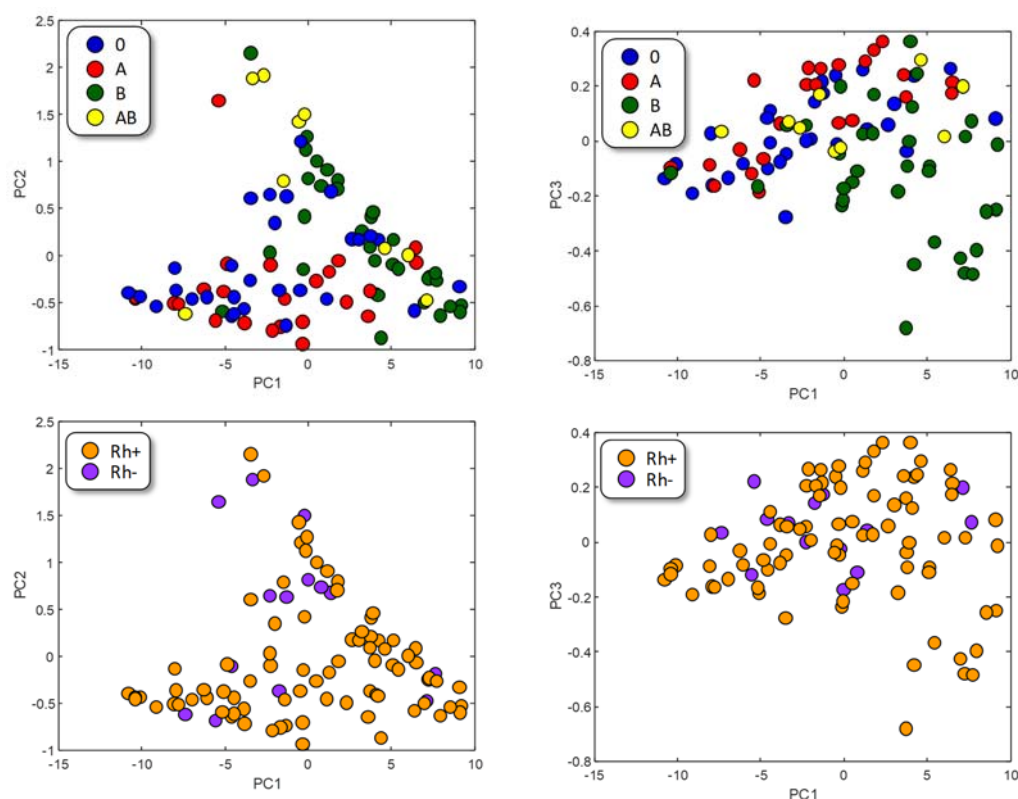


Figure 2 PC representation of samples. Color represents ABO-type (up) or Rh factor (down).

These results suggest that minimizing the intra-class variance while maximizing the inter-class variance could improve the classification success rate. This kind of classification is obtained when performing Linear Discriminant Analysis studies.

3.3. Linear Discriminant Analysis (LDA)

The whole pool of 111 samples was randomly divided in a training dataset (78 samples) and a test dataset (33 samples) and then the LDA training was performed. This procedure was repeated six times, each of them with different training and test datasets. Figure 3 (up) shows the visual representation of LDA algorithm. It is obvious that B-type samples are differently grouped than A-type samples. However, AB and O samples, although more or less self-grouped, overlaps with either B (in case of AB) or A (in case of O). These results suggest that infrared is more sensitive to B antigen, as B-antigen-containing samples (B & AB) appears grouped and segregated from non-containing-B-antigen samples (A & O). The obtained results were significantly better than a pure random

classification ($P < 0.02$, Table 4 up), obtaining a 48% of correctly classified samples versus a 29% of expected success in a mere random classification. Most difficulties are found in O samples classified as A (15%) and A classes classified as O (7%), which is in clear agreement with the visual representation of Figure 3.

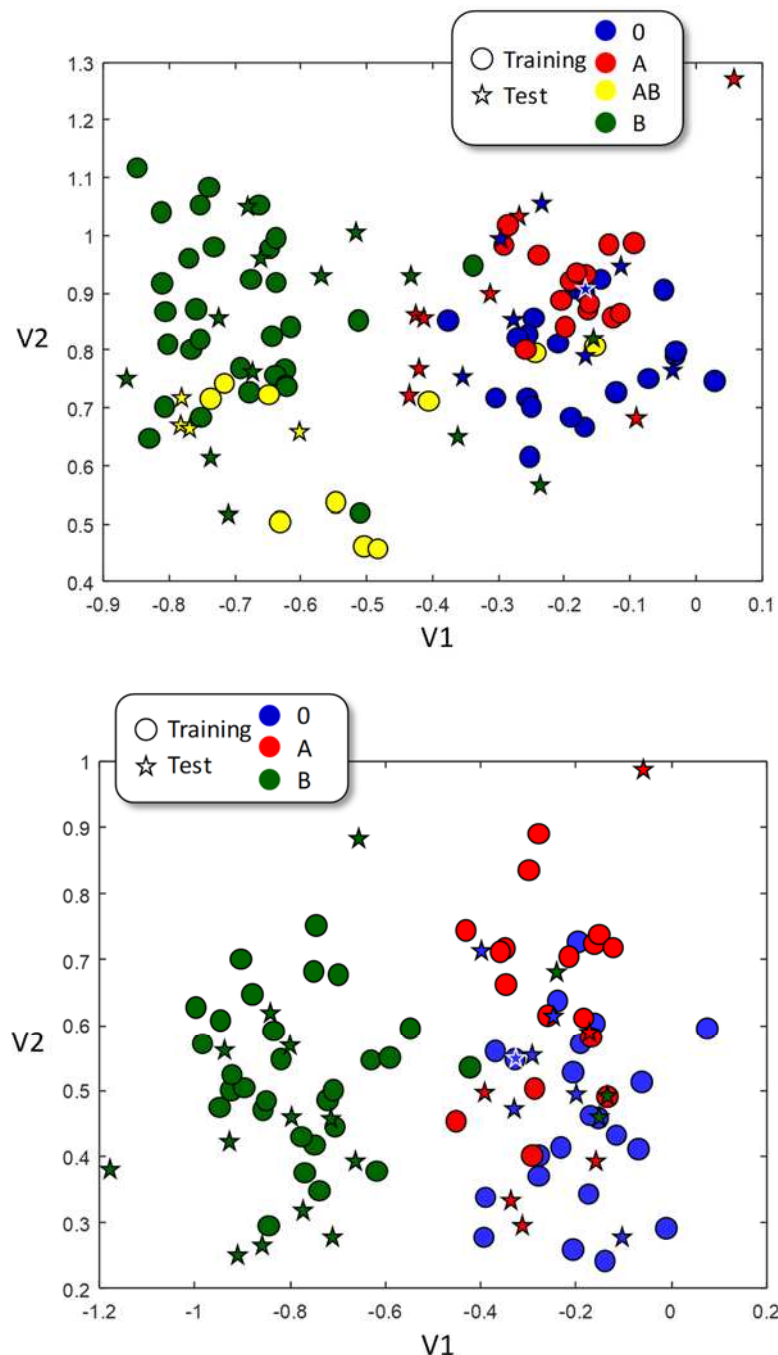


Figure 3 LDA representation for training (circles) and test (stars) datasets. Color of test dataset belongs to their real class, and not to the assigned one. Up, A, B, AB & O blood types used; down, A, B & O only.

Since AB samples contain both A and B antigens and, therefore, spectroscopic characteristics common to both A and B types, the inclusion of AB samples in the training dataset could probably lead to difficulties in the classification. Therefore, we trained a second model including A, B and O samples only. Samples were randomly split into training and test datasets, and the procedure repeated six times. The graphical representation of this LDA classification is shown in Figure 3b.

Training samples from A and B sets clearly segregate, establishing two well differentiated areas; training samples from O type, however, overlap with those from A type, but not with those of B type. Only B-type samples get an improvement in all categories, with an increment in the rates of true positives and negatives, and a drastic reduction in the percentage of false positives and negatives. However, and despite the improvement in the ratio of true positives of both A and O samples, also the ratio of false positives (for O type) and of false negatives (for A type) increases. In this case, a 63% of the samples are correctly classified with, again, a high wrong classification in O samples taken as A and A samples taken as O. This situation is not likely related to the chemical difference on the antigens, since there is a complex interrelationship between the blood type and the blood chemistry, which can participate in the ABO-type identification without being the antigens themselves.

Table 4 Training and test dataset confusion matrixes for LDA and success rates.

LDA classification						Random classification					
Real class	Assigned class					Real class	Assigned class				
		A	B	AB	0			A	B	AB	0
	A	6 %	<1 %	2 %	7 %		A	5 %	9 %	3 %	6 %
	B	6 %	26 %	9 %	4 %		B	9 %	15 %	5 %	10 %
	AB	2 %	6 %	6 %	0 %		AB	3 %	5 %	2 %	3 %
	0	15 %	0 %	3 %	10 %		0	6 %	10 %	3 %	7 %

	T+	T-	F+	F-
A-type success rate	↓ 6.1%	↑ 62.1%	↓ 22.2%	↑ 9.6%
B-type success rate	↘ 25.8%	↗ 50.0%	↑ 6.1%	↓ 18.2%
AB-type success rate	↓ 6.1%	↑ 73.2%	↗ 13.6%	↑ 7.1%
0-type success rate	↓ 9.6%	↑ 62.1%	↗ 10.6%	↘ 17.7%
Average success rate	↓ 11.9%	↑ 61.9%	↗ 13.1%	↗ 13.1%

LDA classification					Random classification				
Real class	Assigned class				Real class	Assigned class			
		A	B	0			A	B	0
	A	10 %	2 %	13 %		A	6 %	11 %	7 %
	B	3 %	37 %	4 %		B	11 %	23 %	14 %
	0	15 %	1 %	16 %		0	7 %	14 %	9 %

	T+	T-	F+	F-
A-type success rate	↓ 9.8%	↑ 58.0%	↓ 17.8%	↓ 14.4%
B-type success rate	↗ 36.8%	↑ 53.4%	↑ 2.9%	↗ 6.9%
0-type success rate	↓ 16.1%	↑ 51.1%	↓ 16.7%	↓ 16.1%
Average success rate	↓ 20.9%	↑ 54.2%	↘ 12.5%	↘ 12.5%

3.4. Naïf Bayessian Classifier

Naïf Bayessian Classifier (NBC) assigns a class to an unknown object based on the Bayes Probability Theorem. It is, therefore, necessary to know the probability of every variable (v_i) taking a certain value when the sample belongs to a class C_j (that is $P(v_i/C_j)$), and the probability of belonging to that class as well ($p(C_j)$). $P(v_i/C_j)$ is estimated from a training dataset randomly selected among the whole dataset. The probability of every class was obtained from the number of samples of every class in the training dataset, using as variables the peak area described in Table 2.

A NBC model was tested using six different training and test dataset randomly selected and in the same proportion than in LDA case. As can be seen in Table 5, the model has a higher success ratio than the random classification ($P < 0.01$), and provides a better true positive ratio than LDA for A, B and AB samples, but not for O.

As in the case of LDA, the presence of both A and B antigens in the AB type may cause difficulties in the classification. Thus, in a first instance, we perform the classification without AB type samples, using again six different training and test datasets. Table 5 shows the classification results both for NBC and a random classification, which are significantly different in every case with $P < 0.002$ according to a Student's-t test. A and B samples have a higher success rate with NBC classification than with a random one, although 0 shows slightly worse results, may be because the lack of antigens in the surface reduces the spectroscopic characteristics enabling the differentiation. Furthermore, LDA classification with only three classes works provides a better true positive ratio for every class.

It is interesting to state that, both in LDA and NBC, the incorporation of AB samples reduces the true positive ratio of the model.

Table 5 NBC results for A, B, AB & 0 (up) and A, B and 0 (down)

NBC classification						Random classification					
Real class	Assigned class					Real class	Assigned class				
		A	B	AB	0			A	B	AB	0
	A	9 %	5 %	1 %	2 %		A	5 %	8 %	3 %	6 %
	B	8 %	29 %	4 %	1 %		B	8 %	15 %	5 %	10 %
	AB	4 %	8 %	6 %	2 %		AB	3 %	5 %	2 %	3 %
	0	11 %	8 %	3 %	3 %		0	6 %	10 %	3 %	7 %

	T+	T-	F+	F-
A-type success rate	↓ 8.6 %	↑ 62.1 %	↓ 22.2 %	↑ 7.1 %
B-type success rate	↘ 29.3 %	↗ 38.4 %	↓ 19.7 %	↗ 12.6 %
AB-type success rate	↓ 6.1 %	↑ 73.7 %	↑ 7.6 %	↗ 12.6 %
0-type success rate	↓ 2.5 %	↑ 72.2 %	↑ 4.0 %	↓ 21.2 %
NBC Success rate	↓ 11.6 %	↑ 61.6 %	↘ 13.4 %	↘ 13.4 %

NBC classification					Random classification				
Real class	Assigned class				Real class	Assigned class			
		A	B	0			A	B	0
	A	10 %	7 %	6 %		A	6 %	11 %	7 %
	B	10 %	30 %	3 %		B	11 %	23 %	13 %
	0	16 %	11 %	7 %		0	7 %	13 %	8 %

	T+	T-	F+	F-
A-type success rate	↓ 9.8 %	↑ 50.6 %	↓ 25.9 %	↗ 13.8 %
B-type success rate	↘ 29.9 %	↗ 39.1 %	↘ 18.4 %	↑ 12.6 %
0-type success rate	↓ 6.9 %	↑ 56.9 %	↑ 9.2 %	↓ 27.0 %
Average Success rate	↓ 15.5 %	↑ 48.9 %	↗ 17.8 %	↗ 17.8 %

3.5. Principal Components combined with Naïf Bayesian Classifier

Taking into account the PCA and LDA results, it is very likely that the relationship of FTIR spectra with ABO-type exists with linear combination of several peak areas rather than the independent areas. For this reason, we tried a second NBC algorithm using as variables just the two first principal components. The whole dataset reduces its dimensionality from the original (18x95) (peaks x samples) to a (2x95) (PCs x samples). The differences between the PC-NBC algorithm and the random classification is statistically significant with $P < 0.01$ in every case but the misclassification of 0 samples as B (italics in the table). However, as shown in Table 6 the improvement is negligible, with a scarce enhancement in B and 0 correct type identification and a slight reduction in the false identification of A and 0 samples as B. This phenomenon, aligns with the notion that the spectroscopic identification of B antigen is easier than the identification of the A one or the lack of them.

Table 6 PC-NBC results for A, B, AB & 0

NBC classification						Random classification					
Real class	Assigned class					Real class	Assigned class				
		A	B	AB	0			A	B	AB	0
	A	4 %	7 %	1 %	9 %		A	5 %	8 %	3 %	6 %
	B	5 %	32 %	4 %	5 %		B	8 %	15 %	5 %	10 %
	AB	0 %	7 %	6 %	0 %		AB	3 %	5 %	2 %	3 %
	0	11 %	8 %	0 %	5 %		0	6 %	10 %	3 %	7 %

	T+	T-	F+	F-
A-type success rate	↓ 3.5 %	↑ 67.7 %	↗ 13.1 %	↘ 15.7 %
B-type success rate	↘ 31.8 %	↘ 31.3 %	↓ 23.7 %	↗ 13.1 %
AB-type success rate	↓ 5.6 %	↑ 83.8 %	↑ 4.0 %	↑ 6.6 %
0-type success rate	↓ 5.1 %	↗ 63.1 %	↗ 13.1 %	↘ 18.7 %
NBC Success rate	↓ 11.5 %	↗ 61.5 %	↗ 13.5 %	↗ 13.5 %

3.6. Single-antigen identification

Based on the results presented earlier, it becomes apparent that the algorithms exhibit a superior ability to differentiate samples containing the B antigen from the rest. Consequently, we proceeded to employ the aforementioned algorithms to ascertain whether they yield improved outcomes in identifying each antigen individually. On the flip side, it was crucial to assess the feasibility of employing a distinct chemometric approach for the identification of each antigen. Hence, we conducted tests utilizing LDA and NBC to discern between samples classified as A and Non-A, as well as B and non-B, whose results are recorded in table 7.

Table 7 Classification as A and B independently with LDA or NBC. * marks worse results than the random classification.

	T+	T-	F+	F-
Identifying A with LDA	↓ 9.6 %*	↑ 50.0 %*	↗ 19.2 %	↘ 21.2 %
Identifying B with LDA	↑ 46.5 %	↑ 42.4 %	↓ 35.4 %	↑ 7.6 %
Identifying A with NBC	↘ 21.2 %	↗ 35.3 %*	↓ 31.3 %*	↑ 12.1 %
Identifying B with NBC	↑ 44.4 %	↘ 26.8 %	↗ 16.2 %	↑ 12.6 %

All the results presented in Table 7 exhibit statistically significant differences compared to the outcomes derived from random classification, with a P-value of less than 0.01. However, it is crucial to note that the classification of A demonstrates less desirable scenarios: the ratios of true positive (T+) and true negative (T-) using LDA, as well as T- using NBC, indicate lower success rates than those achieved randomly. Similarly, the false positive ratio (F+) using NBC surpasses the random counterpart. This further substantiates our previous thesis that the B-antigen is more effectively detected by FTIR than its A counterpart.

Bearing this results in mind, we decided to evaluate whether the independent single identification of A and B antigens would lead to an improvement in the classification rate. Therefore, we used an independent multiple antigen classification (IMAC) consisting in training independently four algorithms (linear discriminant analysis for A (LDA-A) and for B (LDA-B) antigens; Naïf Bayesian Classifier for A (NBC-A) and B (NBC-B) antigens). Every tested sample was assigned an IMAC-number based on which of the four algorithms was + or – (Table 8), and the class was decided according to this number. The correspondence IMAC-number – blood type was decided by studying the histogram of samples vs IMAC-number for every blood type.

Table 8 – IMAC classification table

IMAC number	LDA – A	LDA – B	NBC – A	NBC – B	Blood type
0	-	-	-	-	O
1	-	-	-	+	O
2	-	-	+	-	O
3	-	-	+	+	A
4	-	+	-	-	B
5	-	+	-	+	B
6	-	+	+	-	B
7	-	+	+	+	B
8	+	-	-	-	O/A
9	+	-	-	+	A
10	+	-	+	-	A
11	+	-	+	+	A
12	+	+	-	-	B
13	+	+	-	+	AB
14	+	+	+	-	AB
15	+	+	+	+	AB

IMAC-numbers 4, 8 and 12 did not appear in any of the tested training datasets. In fact, only number 4 appears only once in the assayed test datasets. Then, the blood type for these numbers was decided according to a probability study. It is particularly difficult in the case of IMAC 8, as the probability of being A or O was very similar and there are neither experimental evidences nor probabilistic ones to prefer any type.

The results presented in Table 9 clearly indicate that IMAC yields superior outcomes when compared to random classification for each class, although achieving similar results to those obtained by the previous algorithms. Regarding the confusion matrix, it is interesting to state that samples A are mistaken with O (and vice-versa) in a higher extent than a mere random classification. A similar situation occurs with samples B and AB (and vice-versa). It becomes apparent, again, that the presence of B antigen is particularly influent in the classification.

Table 9 – IMAC Blood typing confusion matrix

NBC classification						Random classification					
Real class	Assigned class					Real class	Assigned class				
		A	B	AB	0			A	B	AB	0
	A	7%	3%	0%	10%		A	5%	8%	2%	4%
	B	3%	31%	7%	4%		B	8%	21%	5%	10%
	AB	1%	8%	4%	1%		AB	2%	5%	2%	3%
	0	12%	2%	0%	10%		0	4%	10%	3%	6%

	T+	T-	F+	F-
A-type success rate	↓ 7.1 %	↑ 66.2 %	↓ 14.7 %	↘ 12.1 %
B-type success rate	↘ 31.3 %	↗ 42.9 %	↘ 12.1 %	↓ 13.6 %
AB-type success rate	↓ 3.5 %	↑ 80.8 %	↑ 7.1 %	↑ 8.6 %
0-type success rate	↓ 10.1 %	↑ 62.1 %	↓ 14.1 %	↓ 13.6 %
IMAC Success rate	↓ 13.0 %	↑ 63.0 %	↘ 12.0 %	↘ 12.0 %

4. Conclusions

Incorporating infrared spectrometry and chemometrics for ABO blood typing yields improved results compared to random classification, although it still falls short of competing with

immunological methods. The presence of antigen 'B' appears to impart discernible characteristics to the blood sample, easily detected with FTIR. However, antigen 'A' remains largely elusive using this technique, posing challenges in its identification. It is possible that the difficulty in identifying blood type A may be attributed to the existence of subgroups, mainly A1 and A2, within this serological type. While there are also subgroups for blood type B, the difference between them is less pronounced, which appears to support the findings obtained in this study [13]. Although, the success rate is worse than in classic biochemical blood typing methods, this methodology shows up as a promising alternative.

Author Contributions: Conceptualization, A. Fernández-González and R. Badía; methodology, A. Fernández-González.; software, A. Fernández-González; formal analysis, T. Fontanil, A. Fernández-González, C. Chimeno-Trinchet; investigation, R. Badía; resources, A.J. Obaya and T. Fontanil; data curation, A. Fernández-González; writing—original draft preparation, A. Fernández-González, A.J. Obaya, C. Chimeno-Trinchet; project administration, R. Badía; funding acquisition, R. Badía; Informed consent and ethical aspects: A.J. Obaya All authors have read and agreed to the published version of the manuscript.

Funding: The authors gratefully acknowledge financial support from the Ministerio de Ciencia, Innovación y Universidades (MCIU), Agencia Estatal de Investigación (AEI), and European Regional Development Fund (FEDER), project # RTI2018-099756-B-I00 (MCIU/AEI/FEDER, UE).

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the utilization of leftover samples from analytical procedures prior to their use in blood donation.

Informed Consent Statement: Patient consent was waived due to the utilization of aliquots taken from analytical samples obtained from anonymous blood donors.

Data Availability Statement: Data are available upon request to the authors.

Acknowledgments: We would like to thank the *Centro Comunitario de Sangre y Tejidos de Asturias* who kindly provided us with blood samples from anonymous donors. We would also like to acknowledge the vibrational and electronic spectroscopy unit from the Scientific-Technical Services of the University of Oviedo.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Kumar S, Srinivasan A, Nikolajeff F, Role of Infrared Spectroscopy and Imaging in Cancer Diagnosis, *Curr. Med. Chem.* 2018 25(9): 1055-1072, doi: 10.2174/0929867324666170523121314
2. Petrov MS, Gordetzov AS, Amelyanov NV, Usefulness of infrared spectroscopy in diagnosis of acute pancreatitis, *Anz J. Surg.* 2007 77(5): 347-351, doi: 10.1111/j.1445-2197.2007.04057.x
3. Yang CM, Guang PW, Li L, Song H, Huang FR, Li YP, Wang LH, Hu JH, Early rapid diagnosis of Alzheimer's disease based on fusion of near- and mid-infrared spectral features combined with PLS-DA, *Optik* 2021 241: 166485, doi: 10.1016/j.ijleo.2021.166485
4. Dogan A, Lasch P, Neuschl C, Millrose MK, Alberts R, Schughart K, Naumann D, Brockmann GA, ATR-FTIR spectroscopy reveals genomic loci regulating the tissue response in high fat diet fed BXD recombinant inbred mouse strains, *BMC Genomics* 2013 14: 386, doi: 10.1186/1471-2164-14-386
5. Glassford SE, Byrne B, Kazarian SG, Recent applications of ATR FTIR spectroscopy and imaging to proteins, *Biochim. Biophys. Acta Proteins Proteom.* 2013 1834(12): 2849-2858, doi: 10.1016/j.bbapap.2013.07.015
6. Graca A, Magalhaes S, Nunes A, Biological predictors of aging and potential of FTIR to study age-related diseases and aging metabolomics fingerprint, *Curr. Metabolomics* 2017 5(2): 119-137, doi: 10.2174/2213235X04666161107150017
7. Muñoz M, Uyar A, Correia E, Díez C, Fernández-González A, Caamaño JN, Trigal B, Carrocera S, Seli E, Gómez E, Non-invasive assessment of embryonic sex in cattle by metabolomic fingerprint of in vitro culture medium, *Metabolomics* 2014 10: 443-451, doi: 10.1007/s11306-013-0587-9

8. Muñoz M, Uyar A, Correia E, Díez C, Fernández-González A, Caamaño JN, Martínez-Bello D, Trigo B, Humblot P, Ponsart C, Guyader-Joly C, Carrocera S, Martin D, Marquant Le Guienne B, Seli E, Gómez E, Prediction of pregnancy viability in bovine in vitro-produced embryos and recipient plasma with Fourier Transform infrared spectroscopy, *J. Dairy Sci.*, 2014 97: 5497-5507, doi: 10.3168/jds.2014-8067
9. Weber A, Hoplight B, Ogilvie R, Muro C, Khandasammy SR, Pérez-Almodóvar L, Sears S, Lednev IK, Innovative vibrational spectroscopy research for forensic application, *Anal. Chem.* 2023 95(1): 167-205, doi: 10.1021/acs.analchem.2c05094167
10. Cebi N, Bekiroglu H, Erarslan A, Rodríguez-Saona L, *Molecules* 2023 28(9): 3727, doi: 10.3390/molecules28093727
11. Murru C, Chimeno-Trinchet C., Díaz-García ME, Badía-Laiño R, Fernández-González A, Artificial Neural Network and Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy to identify the chemical variables related to ripeness and variety classification of grapes for protected designation of origin wine production, *Comput. Electron. Agric.* 2019 104922, doi: 10.1016/j.compag.2019.104922
12. Daniels GL, Fletcher A, Garratty G, Henry S, Jørgensen J, Judd WJ, Levene C, Lomas-Francis C, Moulds JJ, Moulds JM, Moulds M, Overbeeke M, Reid ME, Rouger P, Scott M, Sistonon P, Smart E, Tani Y, Wendel S, Zelinski T. International Society of Blood Transfusion. *Vox Sang.* 2004 Nov; 87 (4): 304-16. doi: 10.1111/j.1423-0410.2004.00564.x.
13. Hosoi E. Biological and clinical aspects of ABO blood group system. *J Med Invest.* 2008 Aug; 55 (3-4): 174-82. doi: 10.2152/jmi.55.174
14. Daniels G, Reid ME. Blood groups: the past 50 years. *Transfusion* 2010 Feb; 50 (2): 281-9. doi: 10.1111/j.1537-2995.2009.02456.x. Epub 2009 Nov 9. PMID: 19906040
15. Tipos de sangre | Todo sobre la donación de sangre | Cruz Roja – Donación de sangre. Available online: <https://www.donarsangre.org/todo-sobre-la-sangre/tipos-de-sangre/> (accessed on 27 september 2022)
16. Socrates, G. Biological Molecules - Macromolecules. In *Infrared and Raman Characteristic Group Frequencies, tables and charts*, 3rd ed.; John Wiley and Sons inc: Chichester, UK, 2001; pp. 328–340.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.