American Journal of Chemistry and Application
2018; 5(3): 45-50
http://www.aascit.org/journal/ajca
ISSN: 2375-3765

AASCIT American Association for Science and Technology

Chemometry of Clinically Importent Fatty Acids in the Blood Serum Using Near Infrared Spectrometer

Andrei Kalinin[1, *], Victor Krasheninnikov[1], Alexandr Sviridov[2], Vladimir Titov[3]

[1]Institute of Spectroscopy, Russian Academy of Sciences, Moscow, Troitsk, Russia
[2]Institute on Photon Technologies of Federal Scientific Research Centre «Crystallography and Photonics», Russian Academy of Sciences Moscow, Troitsk, Russia
[3]Russian Cardiology Research and Production Complex, Ministry of Health, Moscow, Russia

Email address

kalinin@isan.troitsk.ru (A. Kalinin)
[*]Corresponding author

Citation

Andrei Kalinin, Victor Krasheninnikov, Alexandr Sviridov, Vladimir Titov. Chemometry of Clinically Importent Fatty Acids in the Blood Serum Using Near Infrared Spectrometer. *American Journal of Chemistry and Application*. Vol. 5, No. 3, 2018, pp. 45-50.

Received: March 14, 2018; **Accepted**: April 2, 2018; **Published**: May 30, 2018

Abstract: Triglycerides of fatty acids (FAs) are essential components of the human diet, while possessing an unique biological activity. The correlation of some diseases with the ratios of triglycerides of the saturated, monounsaturated and polynoic FAs in blood of patients is widely known. The content of FAs in biological media is traditionally determined using gas or liquid chromatographic methods. These methods are inaccessible to clinical laboratories because of their inefficiency, complexity and high cost. Thus, our goal was to develop a method and technical conditions for rapid assay in the serum of these FAs, namely, saturated palmitic, monounsaturated oleic and double-unsaturated linoleic, and polynoic arachidonic FAs with the use of a near infrared spectrometer and a regression (calibration) method on the projections of latent structures. As a result, the correlation between the absorption spectra of the serum samples with the mass fractions of named FAs and total triglycerides and cholesterol was established. Parameters of the calibration models and the conditions of spectrometry were assessed. Limit sensitivity 0.1 mg/l of arachidonic FA was achieved. So, the applicability of the FT spectrometer in the wavelength range 1.0-1.8 μ for the analysis of the named FAs in serum without sample pretreatment was confirmed.

Keywords: Fatty Acids, Blood Serum, Near Infrared FT Spectrometer

1. Introduction

It is well known [1-3] that cardiovascular diseases and cancer - the main causes of human death, are statistically associated with an increased content of saturated fatty acid (FAs) triglycerides compared to unsaturated FAs in the blood. The most significant FAs should include saturated palmitic (C16:0) and unsaturated oleic (C18:1) and linoleic (C18:2) FAs. In addition, in some studies [4-7], correlations were established in the stages of certain diseases (gallstones, type 2 diabetes mellitus, fatty liver disease (NAFLD) and insulin resistance). Three trans isomers (11C-18:1, 12C-18:1 and 5C-14:1) and three cis isomers (11C-18:1, 13C-18:1 and 7C-16:1) were correlated either positively or negatively with age, body mass index, plasma and lipoprotein cholesterol and/or blood pressure [8]. The positive relationship between the content of total trans fatty acids and the health outcomes in another recent work [9] was confirmed. There are also known the results of tracking the connection between the state of body overload of athletes and the change in the FAs profile of the blood plasma, and, as far as we know, IR spectrometry was used for the first time in this study [10].

While treating cardiovascular diseases there is an urgent need for rapid methods of determining the content of these FAs in blood. Traditional methods of FA analyses are the gas and liquid chromatography, they are of little use for express applications in clinical labs because of inefficacy, the complexity of sample preparation and high cost of analysis. The usage of near infrared (NIR) spectroscopy, combined with the method of multi-component calibration - regression on projections of latent structures (PLS) [11], seems promising for mass express-control of the above-mentioned FAs in serum of patients. The purpose of this work was to study the possibility of applying NIR spectrometers of

various types for the rapid registration of the content of the above FAs and in the development of measurement conditions with the minimum necessary pre-treatment of plasma samples.

2. Materials and Methods

2.1. Sampling

Spectrometric analysis of FA triglycerides includes laborious calibration of the spectrometer using a huge number of standard samples with varying concentrations of several components. For this study, four sets of blood samples in amount of 27, 30, 42 and 28 specimens were taken from patients in the clinical department of cardiovascular diseases of Russian Cardiology Research and Production Complex and serum from them was prepared for research in accordance with the medical regulations in February, March, April 2016 and April 2017. (please, see Table 1).

Table 1. *Parameters of the serum calibration samples according to the reference data.*

Set number, (number of samples)	Analyte, units of measurement	Min.	Max.	Mean	RMS
1 (27)	ТГ, mmol/l	0.33	8.83	2.3	1.3
	Х, mmol/l	0.07	7.12	3.8	1.4
2 (30)	ТГ, mmol/l	0.48	5.88	1.85	1.2
	Х, mmol/l	3.0	7.16	4.51	1.17
	C16:0, µg /l	480	5880	1123	712
	C18:1, µg /l	320	5664	839	201
	C18:2, µg /l	442	6421	1021	266
	C20:4, µg /l	106	406	281	109
3 (42)	C16:0, µg /l	305	5277	961	797
	C18:1, µg /l	253	5544	878	821
	C18:2, µg /l	419	6421	1119	949
	C20:4, µg /l	96	467	228	98
4 (28)	ТГ, mmol/l	0.47	40.6	9.6	3.8
	Х, mmol/l	2.4	16.8	7.8	4.5
	C16:0, µg /l	364	5611	1031	612
	C18:1, µg /l	203	4540	808	720
	C20:4, µg /l	126	546	283	102

The content of total triglycerides (TG) and total cholesterol (Ch) were measured in serum by enzymatic colorimetric techniques (GPO-PAP) Triglycerides -7743 Novo, with the liquid form of Miura, 250, and (CHOD-PAP) Cholesterol-7751 Novo and Miura, 250, respectively. The individual FA triglyceride content were evaluated by using extraction, separation and methylation of FAs with the use of HPLC DIONEX Ultimate 3000 with Reprosil-Pur column C18-AQ 3 micron filler 150 × 4.6 mm. The detector used was mass spectrometer ABSCIEX QTRAP 5500 model working in registration mode of positively charged ions. After the data measurement with the above arbitration methods, the serum samples were frozen to -18°C, stored for no more than three weeks and then defrosted to record their absorption spectra.

2.2. Acquiring the Absorption Spectra

NIR spectrometry of FA triglyceride content in serum is difficult because of dominating water absorption. Earlier, we have found it possible to determine the individual FA in fat-oil mixtures with a small (to 6% mass) level of water using the overtones and combinations of fundamental absorption bands near wavelengths of 1.15, 1.4 and 1.7microns, for which the peak of the last band is about 4 times greater than of the first [12]. To isolate the FAs absorption against the background of a significant water uptake, we used the coincidence of fat absorption peak, and the relative failure of the water absorption near 1.7 microns (Figure 1, a), b). The absorption spectra of serum samples in glass tube of 6 mm diameter were acquired using MPA - FTNIR spectrometer, Bruker, Germany, at wavelength range 0.9 - 1.8 microns, during 4 minutes of accumulation time (64 scans), and the sample temperature of 36°C. The minimum required sample volume was 0.3 ml.

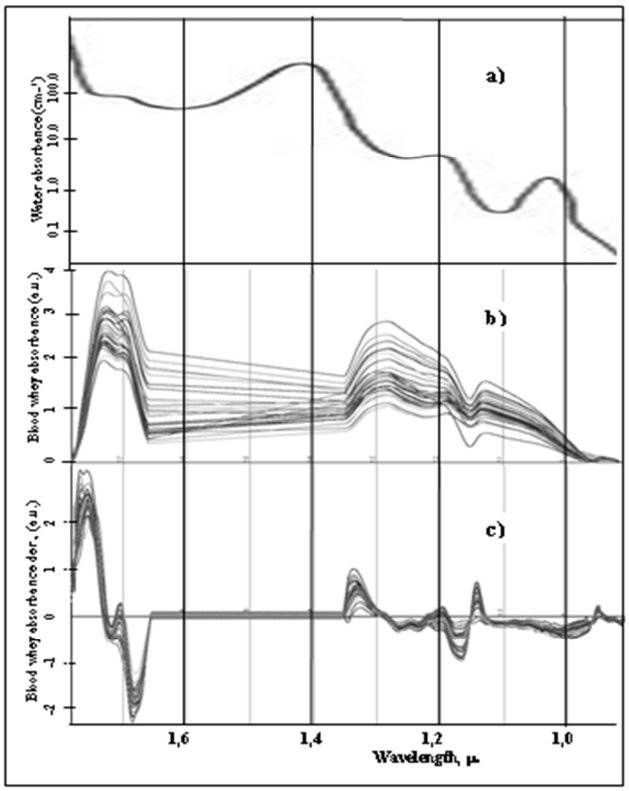

Figure 1. *Spectral values of: a) the water absorption coefficient (in the logarithmic scale), b) the optical density of 27 serum samples, c) the differentiated optical density of the same samples.*

The spectra of optical density $D = \log I_0/I_t$, where I_0 - the flux of radiation with wavelength of 0.9 - 1.8 microns through the cell with distilled water in it, and I_t – the flux passed through the cell with the sample. The reference values of mass fraction of TG, Ch, and FAs C16:0, C18:1, C18:2 and C20:4 in each of the samples were accumulated into self-made PLS regression program ISCAP [13] for building and testing the PLS models. In Figure 1b) digitized at 2600 points of the wavelength scale, the spectra of optical density D of serum samples of the fourth set are presented

The spectra were then smoothed according to Golay-

Savitsky (2 polynomials, with an interval of 30 points) and differentiated - see Figure 1c.

2.3. Constructing and Testing of Multi-component PLS Calibration Models

The solution of the problem of multidimensional linear calibration (regression) from the matrix equation Y = XB + E for multicomponent analysis based on spectra consist of:

- the matrix Y for content of FAs or fats and oils in a set of calibration samples;
- the matrix X for the spectra of the same set of calibration samples;
- the matrix E for error residuals for the each sample of the set;
- the solution - matrix B of calibration coefficients.

Building of a calibration models for the spectra matrix X and FA fractions, Y, and its optimization including:

- construction of various PLS models,
- testing (estimation the value of RMS error using matrix E, choice of model) and

Then it is possible to determine the fractions of ŷ in an unknown sample from its measured spectrum x and the chosen model.

The matrix B from the equation Y = XB is found by the method of least squares: by minimizing the sum of the squared deviations of the values obtained on the calibration samples $(Y - XB)^t (Y - XB)$, we find estimates of the coefficients of the matrix B: $B = (X^t X)^{-1} X^t Y$. The main mathematical problem is the inversion of the matrix X^tX. Obviously, if the number of standard samples is less than the number of variables in X, then the inverse matrix does not exist. Moreover, even for a sufficiently large number of samples, the inverse matrix may not be due to a significant (up to linear) similarity (collinearity) of the spectra.

The tool against collinearity is the data compression using the Principle Component Analysis (PCA).

The PCA forms new variables ta = pa1x1 + pa2x2 +... (a = 1,..., A), called the main components (LVs), so that: $X = TP^T + E$

LVs, or scores, make up the matrix T, together with the loadings matrix P, they reproduce the matrix X quite accurately, and only the error remains in the residuals of E. Wherein:

1) the variables T are orthogonal
2) their number is much less than in the matrix X, i.e. the similarity of the spectra is overcome and T can be used instead of X to construct the regression.

The method of PLS regression:

- compresses both matrices, X and Y;
- LV factors are calculated in turn by the NIPALS algorithm, [14],
- get 2 kits of scores T, U and loadings P, Q plus the load-weights matrix W,
- iteratively improve the model to maximize cov (T, U)
- fulfill the prediction:

$$\hat{Y} = Xnew\ B,$$

$$B = W\ (PTW)^{-1} QT$$

The constructed models (matrices B) were tested with a set of standards not used for calibrations and models with the best reliability factors-the multiple correlation coefficient - r^2_p, the complexity n - the number of LVs - and the error index-the random component of the standard deviation SEP for predicting each analyte in the test samples were selected.

Regression by means of projections to latent structures (PLS, also known as partial least squares) is a useful alternative to the linear multiple regression model fitted by "least squares" if:

1) the number of x-variables (in spectrum) is relatively high compared with the number of observations;
2) the x-variables are correlated;
3) there is more than one y-variable (response, content of component) and these variables are correlated.

In case of a cross validation the root mean square error of cross validation (RMSECV) can be taken as a criterion to judge the quality of the method (for characteristics whose name ends in C (for example, RMSEC), the E_c (training) matrix is used, and for those that end in P (for example, RMSEP), the matrix E_t (check) is taken):

$$\text{RMSEC (k)} = \sqrt{\frac{1}{I_c}\sum_{i=1}^{I_c} e_{ki}^2}.$$

In case of a test set validation this value is called the root mean square error of prediction.

$$\text{RMSEP (k)} = \sqrt{\frac{1}{I_t}\sum_{i=1}^{I_c} e_{ki}^2}.$$

The RMSE values depend on the k-response number.

Bias in calibration (BIASC) and in verification (BIASP) –

$$\text{BIASC (k)} = \frac{1}{I_c}\sum_{i=1}^{I_c} e_{ki}. \quad \text{BIASP (k)} = \frac{1}{I_t}\sum_{i=1}^{I_t} e_{ki}.$$

The BIAS values depend on the k-response number too.

Standard errors in calibration (SEC) and in verification (SEP) –

$$\text{SEC (k)} = \sqrt{\frac{1}{I_c}\sum_{i=1}^{I_c}(e_{ki} - BIASC)^2}, \quad \text{SEP (k)} = \sqrt{\frac{1}{I_t}\sum_{i=1}^{I_t}(e_{ki} - BIASP)^2}.$$

The values of SE depend on k - the response number.

Correlation coefficients r^2 *(k)* between standard y_{ki} and estimated responses $\hat{y}_{ki}$ is:

$$r^2(k) = \frac{I\sum y_{ki}\hat{y}_{ki} - \sum y_{ki}\sum \hat{y}_{ki}}{\sqrt{I\sum y^2_{ki} - (\sum y_{ki})^2}\sqrt{I\sum \hat{y}^2_{ki} - (\sum \hat{y}_{ki})^2}},$$

They are also calculated separately for the calibration r^2_C

(k) and verification $r^2_V(k)$ kits.

In all these formulas, the quantities e_{ki} are the elements of the matrices E_c or E_t.

Using the PLS regression program ISCAP [12], calibration models were constructed for the optical density spectra in the wavelength range 0.9 - 1.8 μ, and the mass fractions of total triglycerides, total cholesterol in the 1st, 2nd and 4th sets of samples and chromatographic reference data on the content of FAs C16:0, C18:1, C18:2 and C20:4 in the 2nd, 3rd and 4th sets of samples. Efficiency of them was evaluated by "leave-one-out" cross validation technique in accordance with [10] by the values of multiple correlation coefficient *r2cv* (reliability), standard deviation s from reference data and n - the number of latent variables (LVs, complexity). Next, we have tested each model on serum samples from the other sets, which were not used for calibration. Parameters of prediction, *r2p*,

Particular attention was paid to the comparison of vectors of the regression coefficients $W_{FA,i}$ indicating spectral regions, which are essential for the determination of each of the FAs - see for example $W_{C16:0,6}$ and $W_{C20:4,5}$ in Figure 2.

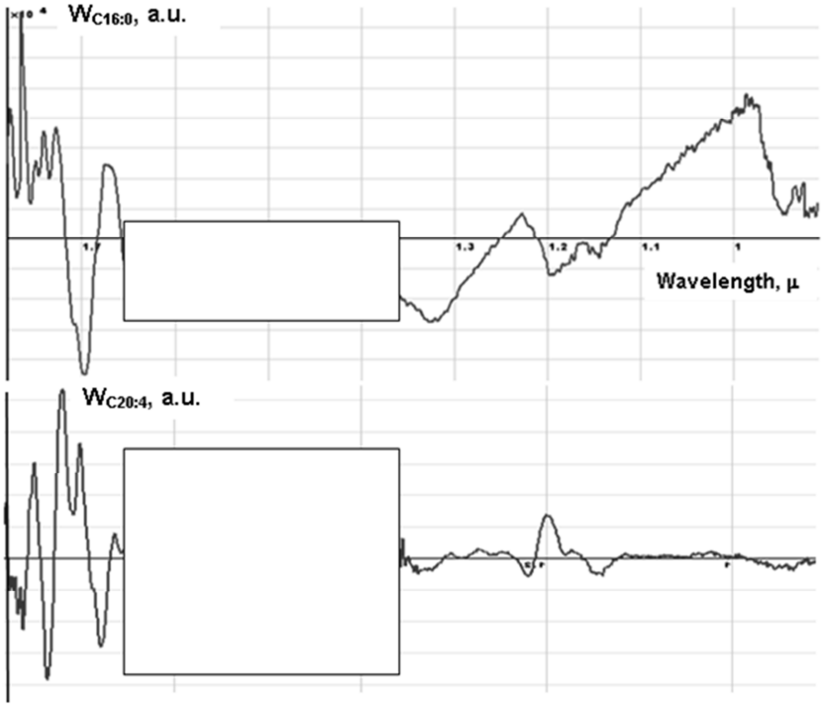

Figure 2. *Vectors of regression coefficients $W_{C16:0,6}$ and $W_{C20:4,5}$ for the PLS models using differentiated spectra with 6 and 5 LVs, respectively: rectangles are closed ranges, where the signals I_0 and I_t were near to 0 because of water absorption.*

3. Results and Discussion

Registration of the optical density variation in the spectra of serum samples for the range of 0.9-1.8 μ allowed to evaluate the correlation with the variation of the fractions of total triglycerides and cholesterol, individual C16:0, C18:1, C18:2 and C20:4 FAs. The best prediction performance of total triglycerides and cholesterol was occurred with the help of the GSDer model (see Table 2), built on smoothed and once-differentiated spectra of 55 samples. Models for the prediction of C16:0, C18:1 and C18:2 FAs, based on the cross-validation results with the test samples, showed the values of the r^2_p in the range of 0.91-0.93.

Table 2. *Parameters of tested models for prediction of total triglycerides (TG, mmol/l), cholesterol (Ch, mmol/l) and FAs using independent samples: K - number of calibration samples, L - number of test samples, H - spectrum processing method: GSDer and 2GSDer - once or twice smoothed and differentiated spectra, Rb - baseline removed spectra, n - number of LVs, r^2_p - multiple correlation coefficient,* s *- standard deviation, for prediction.*

№(K, L)	H	TG			Ch			C16:0			C18:1			C18:2			C20:4		
		n	r^2_p	s	n	r^2_p	s	n	r^2_p	s	n	r^2_p	s	n	r^2_p	s	n	r^2_p	s
1+2+4 (55, 22)	GSDer	6	0.94	1.37	5	0.86	1.49	8	0.85	119	7	0.86	194	8	0.81	183	7	0.81	102
2+1 (30, 25)	GSDer	5	0.83	1.36	5	1.79	0.61	9	0.86	117	7	0.84	126	8	0.83	112	10	0.76	113
	2GSDer	9	0.96	1.81	4	1.69	0.65	8	0.87	118	7	0.86	226	8	0.85	141	12	0.82	122
	Rb	10	0.87	1.4	9	0.87	10.42	6	0.91	215	7	0.91	222	7	0.93	121	10	0.61	128
2+3 (26, 42)	GSDer	7	0.96	1.23	5	0.88	1.4	2	0.93	112	2	0.91	122	4	0.93	114	9	0.63	118
	2GSDer	10	0.81	1.55	12	0.77	1.56	2	0.92	113	2	0.91	124	2	0.91	116	8	0.63	118
2+4 (39, 25)	GSDer	7	0.96	1.23	5	0.88	1.4	2	0.93	112	2	0.91	122	4	0.93	112	9	0.63	121
	2GSDer	10	0.81	1.55	12	0.77	1.56	2	0.92	113	2	0.91	124	3	0.91	118	10	0.64	118

Based on the samples of the third batch, PLS models for prediction of arachidonic FA C20: 4 were also constructed. Verification of their parameters by the "leave-one-out" method for the optimal model (for 39 samples) showed the values r^2_{cv} = 0.69, s = 0.27%, n = 4, which indicates the presence of correlation, but requires testing on samples not included in the calibration set. Testing them with reference samples of 2nd set gave rough readings, See Table 2. This result, however, is valuable in that the non-major arachidonic

FA in serum was determined at a minimum value of its fraction 96 μg / l, which is a record high sensitivity for NIR spectrometry in serum.

On the other hand, statistical analysis of chromatographic data on the content of palmitic, oleic and linoleic acid in the serum showed high values of the pair (according to Pearson) correlation between them. This circumstance allows us to restrict ourselves to determining one of them or their sum, which reduces the requirements for the sensitivity of the method and facilitates the achievement of the necessary metrological parameters of the analysis.

As it is known [9, 11], the usage of the PLS model parameter - the vector of regression coefficients (see Figure 2), allows to select the spectral regions with regression coefficients of most significance for determining each FA, and to use these regions for building models that don't overlap with spectral regions significant for prediction of other components. For example, in Figure 2, the peaks of the regression vectors in the spectral regions 0.9-1.1, near 1.2 and 1.33 μ had significantly different values for C16:0 and C20:4, which made it possible to test the selectivity of the FA prediction. To estimate the possible effect of saturated palmitic C16:0 FA variations on the determination of arachidonic C20:4 FA the generalized results of NIR overtone and combination assignment to C = C bond vibrations of samples according to the results of [13] were used.

We have constructed the models named GSDerSel and 2GSDerSel for the spectral regions characterizing the C = C bond vibrations in C20:4 FA, namely, for wavelengths 1.15 - 1.25, 1.65-1.75 μ and have obtained the best reliability and error of determination: r^2_{cv} = 0.86, s = 0.27% and r^2_{cv} = 0.81, s = 0.23%, respectively, which were better, than the parameters for full spectra range model.

In addition to the PLS models for predicting the content of clinically important C16:0, C18:1, C18:2 and C20:4 FAs, the models were constructed to evaluate the parameters of the regression for serum spectra with Ch content, which possibly "interferes" with the determination of total TG and their regression vectors were investigated.

It turned out that the construction of models in regions of the spectrum where the regression coefficients are high for the determination of total TG and small for cholesterol X does not improve the indices of total TG prediction, so, the assumption of an "interfering" effect of cholesterol was not confirmed.

4. Conclusions

It is shown in the present work that individual triglycerides of diagnostically significant FAs in blood serum can be quickly evaluated without sample preparation using a NIR spectrometer in the wavelength range 0.9-1.8 μ. Since the latest work on the spectrometry of therapeutically significant FAs in foodstuffs, including our works [11, 12], showed the possibility of controlling their intake with food, and our present work opens the possibility of an operative control of the named FAs in the blood, the combination of both approaches makes it possible to intensify the therapy of cardiovascular diseases.

From the standpoint of general biology, it is clearly positive if adults, especially the elderly, refuse to consume butter and reduce the content of foods rich in palmitic acid: beef, sour cream, fat cheeses [14], and so normalize the ratio of saturated FA to unsaturated in the body. The prospect is to test the applicability of a portable and inexpensive produced domestically NIR spectrometer of this range in the present task, and the prospect of blood sampling from the finger, rather than from the vein, as is customary in traditional clinical biochemistry, is especially attractive. The creation of such an analyzer is a real step in the prevention of metabolic pandemics (atherosclerosis and atheromatosis, metabolic syndrome, insulin resistance, obesity).

Acknowledgements

The authers are glad to thanks a lot Dr. Nik Boldyerev for his kind assistanse of aquiring FTNIR spectra and discussions

The work was supported by the Russian Foundation for Basic Research, grant No. 16-02-00277.

The authers declared the absence of any conflicts of interest

References

[1] VN Titov, AV Aripovsky, SI Kaba, PO Kolesnik, MI Vezdel, Yu. K. Shiryaeva, Clinical laboratory diagnostics, No. 7 (2012) 3- 8 (in Russian)

[2] University of Maryland Medical Center, available at: http://umm.edu/health/medical/altmed/supplement/omega3-fatty-acids accessed 26.10.2016.

[3] D. J. Holub and B. J. Holub, Omega-3 fatty acids from fish oils and cardiovascular disease, Molecular and Cellular Biochemistry, 2004, 263, 217–225.

[4] A. Kh. Kadyrov, F. Kh. Mansurova, MB Toshev, M. M. Suriev, Determination of the content of higher fatty acids in the sera of healthy individuals and patients with cholelithiasis at various stages of lithogenesis, Reports of the Academy of Sciences of the Republic Tajikistan, 2008, 51, No. 11, 858-863,. (in Russian).

[5] Yang Y. J., Choi M. H., Paik M. J., Yoon H. R., Chunga B. C., Gas chromatographic–mass spectrometric determination of plasma saturated fatty acids using pentafluorophenyldimethylsilyl derivatization. J Chromatogr. B, 2000, 742, pp. 37-46.

[6] Yi L. Z., He J., Liang Y. Z., Yuan D. L., Chau F.-T. Plasma fatty acid metabolic profiling and biomarkers of type 2 diabetes mellitus based on GC/MS and PLS-LDA. FEBS Letters, 2006, 580, pp. 6837-6845.

[7] E. Á. Hernández, et all., Acute dietary fat intake initiates alterations in energy metabolism and insulin resistance, The Journal of Clinical Investigation, 2017, 127, №2, 695-708.

[8] L. C. Hudgins. J. Hirsch and E. A. Emken, Correlation of isomeric fatty acids in human adipose tissue with clinical risk factors for cardiovascular diseas, Am J Clin Nutr, 1991: 53: 474-82.

[9] R. J. de Souza, et all, Intake of saturated and trans unsaturated fatty acids and risk of all cause mortality, cardiovascular disease, and type 2 diabetes: systematic review and meta-analysis of observational studies, BMJ 2015; 351: h3978 | doi: 10.113 6/bmj.h3978

[10] C. Petibois, G. Cazorla, A. Cassaigne, and G. Déléris, Application of FT-IR Spectrometry to Determine the Global Metabolic Adaptations to Physical Conditioning in Sportsmen, Applied Spectroscopy (2002), 56, № 10, pp. 1259-1267.

[11] Wold S., Sjostrom M., Eriksson L., PLS regression, Chem. Intelligent. Lab. Syst. 58 (2001) 7: 109–30.

[12] Kalinin A., Krasheninnikov V., Sadovskiy S., Denisovich E., Yurova E., J. Near Infrared Spectr., 16 (2008) 3: 343–8.

[13] A. V. Kalinin, V. N. Krasheninnikov, A. P. Sviridov, V. N. Titov, Journal of Applied Spectroscopy, 2016, 83, N5, pp. 811-819,

[14] A. V. Kalinin, V. N. Krasheninnikov, A. P. Sviridov, V. N. Titov, Clinical laboratory diagnostics, 60 (2015) №11 13-20 (in Russian)

[15] F. Westad, A. Schmidt and M. Kermit, J. Near Infrared Spectrosc. 16 (2008) 3: 265-73.

[16] Titov V. N, Pathogenesis. 2013; 11 (1): 18–26.