

Intersubject Variability of Blood Analysis Reference Values: Assessment of Age and Locality Influence by Means of a Linear Discriminant Analysis Model

João C. M. Barreira,¹ Roberto Costa,^{1,2} M. Graça Pombo,² and Isabel C. F. R. Ferreira^{1*}

¹CIMO-ESA, Instituto Politécnico de Bragança, Campus de Santa Apolónia, Apartado 1172, Bragança, Portugal

²Laboratório Dra Matilde Sampaio – BragançaLab, Rua Emídio Navarro, n° 2, R/Chão, Bragança, Portugal

Glycemic and lipidic profiles might be influenced by several factors. The influence of the age group in the most extensively evaluated lipidic and glycemic parameters is more or less well-accepted. To verify this empirical notion, 996 human subjects aged between 21 and 90 years from different localities were characterized according to age. To assess lipid profile, total cholesterol and cholesterol associated with lipoprotein fractions (low-density lipoprotein cholesterol and high-density lipoprotein cholesterol) and triglycerides were determined. Regarding glycemic profile, glucose, and glycated hemoglobin were measured. The majority of the population had values of lipidic parameters fit into the reference values, presenting low or moderate risk for developing cardiovascular disease. Blood glucose was often far above

the desirable, but this can be devalued due to the HbA1c values, which were overwhelmingly located in the normal range. The categorization of data in different age groups did not allow defining statistically significant differences. Despite the discriminant linear model was presented, the results indicate that age group did not act as a strong discriminant factor. Somehow unexpectedly, the most significant differences were found among the different localities, which tended to show a similarity according with their latitude. Furthermore, there were no significant correlations in the parameters associated with lipidic profile, but there was a direct correlation between glucose levels and HbA1c (glycemic parameters). *J. Clin. Lab. Anal.* 27:237–244, 2013. © 2013 Wiley Periodicals, Inc.

Key words: LDL-C; HDL-C; triglycerides; glycated hemoglobin; biostatistics

INTRODUCTION

The definition of reference values for the main blood parameters should be as well adapted as possible to the specific characteristics of a determined population. The lipidic components are among the compounds with higher relevance. The European Systematic Coronary Risk Evaluation (SCORE) is the recommended risk estimation system and estimates a 10-year risk of cardiovascular (CV) disease mortality based on age, gender, country of origin, systolic blood pressure, smoking status, and either total cholesterol or the total/HDL-c ratio (1).

Lipid guidelines typically focus on total cholesterol \pm low-density lipoprotein cholesterol (LDL-C) levels with less emphasis on high-density lipoprotein cholesterol (HDL-C) or triglyceride (TG) assessment, thus potentially underestimating CV risk and the need for lifestyle

or treatment optimization. In current practice, isolated total cholesterol measurements are frequently used and, if low, subjects are reassured. However, low total cholesterol levels may hide HDL-C and TG abnormalities (2). For instance, multiple statin intervention trials, while confirming the benefit of reduction of LDL-C, have demonstrated an increase in risk associated with low levels of HDL-C (3). However, early detection of elevated serum

*Correspondence to: Isabel C. F. R. Ferreira, CIMO-ESA, Instituto Politécnico de Bragança, Campus de Santa Apolónia, apartado 1172, Bragança 5301-855, Portugal. E-mail: iferreira@ipb.pt

Received 26 April 2012; Accepted 15 January 2013

DOI 10.1002/jcla.21590

Published online in Wiley Online Library (wileyonlinelibrary.com).

cholesterol makes possible to identify an important modifiable risk factor for coronary artery disease (4).

These lipid metabolic changes, commonly known as dyslipidemias, are caused by disorders in any stage of lipid metabolism that may cause impact on serum lipoprotein levels. Occasionally, the genetic defects require the presence of secondary factors for their clinical expression (dyslipidemias of mixed etiology). The majority of primary dyslipidemias is polygenic, with influence from multiple genetic and environmental factors (5).

There has also been much debate regarding the role of elevated TGs in atherosclerosis because of their clear association with vascular risk. High TG might not be directly associated with CV disease, although they may aggravate disease if subjects have concomitant elevated LDL-C and/or low HDL-C. Raised TG are also an integral component of the insulin resistance syndrome, itself associated with glucose intolerance, hypertension, and diabetes, which increases the vascular risk in the presence of LDL-C and HDL-C abnormalities (6). On the other hand, maintaining plasma glucose standard concentrations is crucial to prevent and/or delay the onset of chronic complications. Glycated hemoglobin (HbA1c) test is considered as a benchmark for assessing glycemic control and represents a summary of glycemic excursions during the 6–8 weeks prior to its determination (7). This kind of hemoglobin is extremely useful in diabetes monitoring, providing a retrospective index of integrated glucose values in plasma for an extended period (8–10 weeks) (8). In view of these facts, the characterization of lipidic and glycemic profiles assumes particular importance, since it might act as a fundamental risk assessment evaluator, preventing, or lowering the occurrence of CV diseases.

Linear discriminant analysis (LDA) is a multivariate statistical technique whose objectives are: (i) identify the variables that best distinguish (or discriminate) two groups of individuals structurally different and mutually exclusive; (ii) apply these variables to create a discriminant function that represents parsimoniously the differences between the groups; and (iii) using this discriminant function to classify a priori new individuals in the groups. In this work, the main objective was to evaluate the influence of age group on lipidic and glycemic profiles, aiming to understand if the reference values should include potential correction factors defined according with subject's age. The influence of locality was also evaluated to check eventual interaction effects.

MATERIALS AND METHODS

Subjects

Data from 996 subjects (526 females and 470 males) were obtained in the Clinical Analysis Laboratory Dr

Matilde Sampaio (BragançaLab, Portugal). The subjects were resident in the district of Bragança, Northeast of Portugal, and aged between 21 and 90 years. They were divided according to age (21–30 years, 17 subjects; 31–40 years, 38 subjects; 41–50 years, 66 subjects; 51–60 years, 211 subjects; 61–70 years, 309 subjects; 71–80 years, 270 subjects; 81–90 years, 85 subjects) and locality (501 females and 446 males with the following distribution: Bragança (41°48' N; –6°45' O), 371 subjects; Alfândega da Fé, (41°20' N; –6°59' O) 261 subjects; Vinhais (41°50' N; –7°00' O), 192 subjects; Torre de Moncorvo (41°11' N; –7°03' O), 87 subjects; Miranda do Douro (41°30' N; –6°16' O), 36 subjects; the localities/counties with low number of subjects were removed in this evaluation). In both cases the ratio males:females was 1.12.

Biological Samples

The analyses were performed in serum (total cholesterol, LDL-C, HDL-C, TG, and glucose) or whole blood (HbA1c). Peripheral blood samples were obtained by venipuncture and placed into polypropylene tubes containing gel for seric determinations or the anticoagulant EDTA (ethylenediaminetetraacetic acid) for whole blood parameters. Serum was removed after centrifugation at 1500 g for 10 min at 4°C.

To assess lipid profile, total cholesterol, LDL-C, HDL-C, and triglycerides were determined. Regarding the glycemic profile, glucose, and HbA1c were quantified.

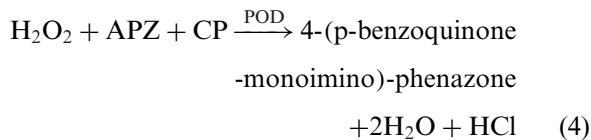
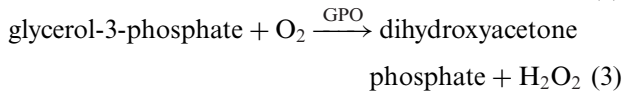
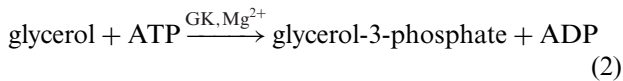
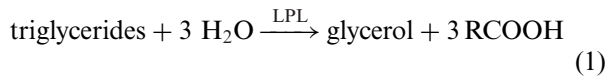
Seric Determinations

To complete this study, several analytical determinations were performed in the serum obtained from blood samples of users of the Clinical Analysis Laboratory Dr Matilde Sampaio, BragançaLab, using a COBAS INTEGRA 400 plus analyzer (Roche Diagnostics).

Triglycerides quantification

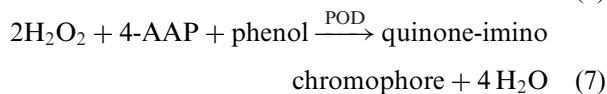
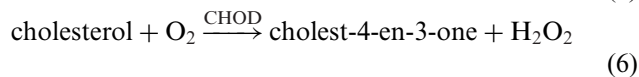
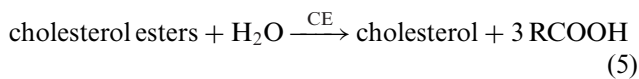
An *in vitro* test for quantitative determination of the concentration of triglycerides was carried out in human serum, using an enzymatic colorimetric assay. This method is based on Wahlefeld (9) work using a microbial lipoprotein lipase (LPL) for the rapid and complete hydrolysis of triglycerides [1], followed by phosphorylation to glycerol-3-phosphate catalyzed by glucose kinase (GK) [2] and oxidation to dihydroxyacetone phosphate and hydrogen peroxide catalyzed by phosphate oxidase (GPO) [3]. The hydrogen peroxide then reacts with 4-aminophenazone (APZ) and 4-chlorophenol (CP), under the catalytic action of peroxidase (POD) to form a red chromophore (Trinder endpoint reaction) [4]. The color intensity of the red chromophore was determined at 659 nm, being directly proportional to the concentration

of triglycerides. As a benchmark, it is recommended a value below 150 mg/dl (1.7 mmol/l).



Total cholesterol quantification

An in vitro test for quantitative determination of the concentration of cholesterol was carried out in human serum through an enzyme-based colorimetric assay (9). Cholesterol esters are cleaved by cholesterol esterase (CE), producing free cholesterol and fatty acids [5]. Cholesterol oxidase (CHOD) catalyzes the oxidation of cholesterol to cholest-4-en-3-one and hydrogen peroxide [6]. In the presence of POD, hydrogen peroxide affects the oxidative coupling of phenol and 4-aminoantipyrine (AAP), forming a red quinone-imine chromophore [7]. The color intensity was determined measuring the absorbance at 512 nm, being directly proportional to the concentration of cholesterol. The levels for adults in terms of risk or coronary heart disease are: recommended: <200 mg/dl (5.2 mmol/l); moderate risk: 200–239 mg/dl (5.2–6.2 mmol/l); high risk: ≥240 mg/dl (≥6.2 mmol/l).

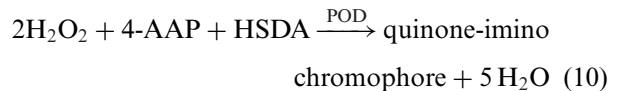
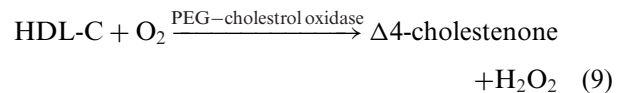
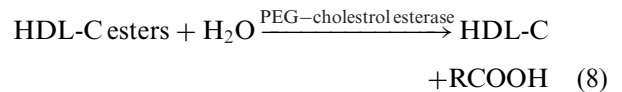


HDL-C quantification

An in vitro test for quantitative determination of the concentration of HDL-c was carried out in human serum using a homogeneous enzyme-based colorimetric assay (9). In the presence of magnesium ions and dextran sulfate, water-soluble complexes with LDL, VLDL, and chylomicrons are formed, being resistant to enzymes modified by polyethylene glycol (PEG). The concentration of HDL-C is determined enzymatically by CE and CHOD

coupled with amino groups (40%). CE quantitatively decomposes cholesterol esters into free cholesterol and fatty acids [8]. In the presence of oxygen, cholesterol is oxidized by CHOD to Δ4-cholestenone and hydrogen peroxide [9]. The intensity of the violet quinone-imine chromophore [10], directly proportional to the concentration of HDL-C, was determined by measuring the absorbance at 583 nm.

Considering the risk of coronary heart disease, the recommended levels for adults are: recommended: >60 mg/dl (>1.55 mmol/l); moderate risk: 40–60 mg/dl (1.04–1.55 mmol/l); high risk: <40 mg/dl (<1.04 mmol/l).



^a N-(2-hydroxy-3-sulfopropyl)-3,5-dimethoxyaniline sodium

LDL-C quantification

LDL-C contents were obtained from measurements of total cholesterol, triglycerides, and HDL-C by the Friedewald approximation:

$$\text{LDL-C} = \text{total cholesterol} - \text{HDL-C} - \frac{\text{triglycerides}}{5}$$

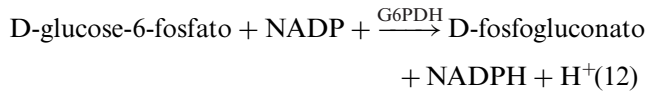
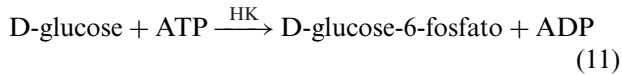
This calculation is valid for concentrations of triglycerides under 400 mg/dl (4.52 mmol/l). The recommended levels for adults, regarding coronary heart disease risk are:

Desirable: <130 mg/dl; moderate risk: 130–159 mg/dl; high risk: >160 mg/dl.

Glucose quantification

An in vitro test for quantitative determination of the concentration of glucose concentration was carried out in human serum by a colorimetric assay with hexokinase (HK) enzymatic reference (9). HK catalyzes the phosphorylation of glucose by ATP, yielding glucose-6-phosphate

and ADP [11]. To continue the reaction, a second enzyme, glucose-6-phosphate dehydrogenase (G6PDH) is used to catalyze the oxidation of glucose-6-phosphate by NADP⁺, leading to NADPH [12]. The concentration of NADPH formed is directly proportional to the concentration of glucose and was determined by measuring the absorbance at 340 nm. The reference values lay among 70 and 110 mg/dl (3.89 and 6.11 mmol/l).



Whole Blood Determinations

HbA1c quantification

HbA1c was determined with an analyzer HA-1840 (A. Menarini Diagnostics), which uses reversed-phase ionic change high performance liquid chromatography (10). The fractions are separated by electrostatic interactions with the gel column, whose surface contains hydrophobic groups and ion exchange groups. The signal of the separated hemoglobin fractions was red at two distinct wavelengths: HgA1 fraction at 415 nm and HbF fraction at 500 nm. The reference values for this analyte are 4.5–7% for healthy subjects or patients with controlled diabetes; 7–8.5% for patients with inadequate treatment or subclinical diabetes with normal glucose value but with altered tolerance test; > 8.5% for diabetics without control.

Statistical Analysis

An analysis of variance (ANOVA) with Type III sums of squares was performed using the General Linear Model (GLM) procedure of the SPSS software, version 18.0 (SPSS, Inc.). All dependent variables were analyzed using a 2-way ANOVA, being the main factors the “age group” (21–30, 31–40, 41–50, 51–60, 61–70, 71–80, and 81–90 years) and the “locality” (Alfândega da Fé, Bragança, Macedo de Cavaleiros, Torre de Moncorvo, Vinhais) of the tested subjects. The assumption of equal variances among groups (homoscedastic distribution) was tested through Levene’s test, with further application of Tukey honestly significant difference (HSD), $\alpha = 0.05$. Tukey HSD uses the Studentized range statistic to make all pairwise comparisons between groups and sets the experiment wise error rate to the error rate for the collection for all pairwise comparisons. In addition, a LDA was used as a technique to classify the seven age groups as well as the five localities according to their lipidic and glyceemic seric profiles. A stepwise technique, using the Wilks’ Λ method

with the usual probabilities of F , was applied for variable selection. It is fundamental to test the ability of discriminant functions to significantly discriminate the groups. Hence, it is intended to test if the group centroids (this is, the mean of all discriminant functions for each group) are equal or if there is at least a different one. This procedure uses a combination of forward selection and backward elimination procedures, where the selection of each new variable is preceded by the verification of the significance of all the previously selected variables (11, 13). With this approach, it is possible to identify the significant variables among the lipidic and glyceemic parameters obtained for each sample. To verify which canonical discriminant functions were significant, the Wilks’ Λ test was applied. The F value for each Wilks’ Λ value resulting from the inclusion or removal of a variable is calculated as follows:

$$F = \left(\frac{n-g-p}{g-1} \right) \left(\frac{1 - \frac{\lambda_{p+1}}{\lambda_p}}{\frac{\lambda_{p+1}}{\lambda_p}} \right)$$

where n is sample global dimension, g is the number of groups, p is the number of independent variables, λ_p is the Wilks’ Λ value prior to the inclusion/removal of the new variable and λ_{p+1} is the Wilks’ Λ value after the inclusion/rejection of the new variable. This statistic follows F-Snedecor distribution with $(g - 1)$ and $(n - g - p)$ degrees of freedom, and the associated significance probability measures the significance of the inclusion/removal of the new variable. To avoid overoptimistic data modulation, a leaving-one-out cross-validation procedure was carried out to assess the model performance. The LDA statistical analysis and the other statistical tests were performed at a 5% significance level using the SPSS software mentioned above.

RESULTS

Prior to the statistical interpretation of results, the main outcomes will be described in order to characterize the studied population. All the results are presented considering different age groups (Table 1) and localities (Table 2) as the mean value of each age group (Table 1) or locality (Table 2) independently of gender (G), as well as the mean for each gender for all age groups (A) or localities (L). In this way, the results allow a more comprehensive analysis of the influence exerted by each one of the main factors under study, avoiding interaction effects. Nevertheless, results showed that the interaction $A \times G$ acted as a significant source of variation ($P < 0.05$) only for glucose ($P = 0.043$) and HbA1c ($P = 0.043$), while the interaction $L \times G$ did not exerted this effect in any case. Regarding the results found in each main factor, the differences were only statistically significant for locality influence, in the cases

TABLE 1. Biochemical Parameters (mg/dl) in Function of the Assayed Age Groups (A). The Results are Expressed as Mean \pm SD ($n = 996$; 470 Males and 526 Females)

		Total cholesterol	HDL-c	LDL-c	Triglycerides	Glucose	HbA1c (%)
Age (A)	21–30	202 \pm 34	50 \pm 8	129 \pm 31	117 \pm 60	100 \pm 36	5 \pm 1
	31–40	188 \pm 44	47 \pm 7	112 \pm 37	142 \pm 61	134 \pm 68	6 \pm 2
	41–50	207 \pm 45	49 \pm 11	130 \pm 37	141 \pm 63	139 \pm 70	7 \pm 2
	51–60	204 \pm 46	47 \pm 8	126 \pm 43	156 \pm 76	133 \pm 60	6 \pm 2
	61–70	197 \pm 42	47 \pm 9	122 \pm 39	137 \pm 64	134 \pm 52	7 \pm 2
	71–80	200 \pm 46	47 \pm 8	125 \pm 41	144 \pm 66	131 \pm 57	7 \pm 2
	81–90	200 \pm 39	48 \pm 9	124 \pm 34	137 \pm 60	121 \pm 47	6 \pm 2
	<i>P</i> -value	0.226	0.371	0.379	0.069	0.076	0.094
Gender (G)	Male	197 \pm 44	46 \pm 9	122 \pm 40	145 \pm 73	132 \pm 58	6 \pm 2
	Female	202 \pm 44	48 \pm 8	126 \pm 39	141 \pm 61	132 \pm 55	7 \pm 2
	<i>P</i> -value	0.357	0.274	0.375	0.785	0.594	0.636
A \times G	<i>P</i> -value	0.454	0.449	0.412	0.558	0.043	0.036

of total cholesterol, HDL-C, glucose, and and HbA1c; therefore, multiple comparisons were only performed in these situations (Table 2). However, some general observations could be pointed out for the remaining cases by analyzing the estimated marginal means (EMM) plots.

The subjects belonging to 41–50 years age group (A) registered the highest total cholesterol (207 \pm 45 mg/dl) and LDL-C (130 \pm 37 mg/dl), while the lowest values of these parameters were obtained in the 31–40 years A (Table 1). HDL-C seemed to be less related with total cholesterol, since its highest value (50 \pm 10 mg/dl) was detected for the 21–30 years A. Regarding triglycerides, the highest value (156 \pm 76 mg/dl) was found for subjects with ages varying from 51 to 60 years, while the lowest (117 \pm 30 mg/dl) was registered for the 21–30 years A. Glucose (100 \pm 36 mg/dl) and HbA1c (5 \pm 1%) reached minimal values in the 21–30 years A.

Despite no statistical differences could be found for the assayed A, it is noteworthy that the parameters with potential negative effect for the coronary heart diseases reached maximal values in the 41–50 A, since younger sub-

jects remain, in general, physiologically protected, while older ones are probably under some therapeutic follow up.

Concerning locality (L) influence (Table 2), the highest total cholesterol value was obtained in Alfândega da Fé (206 \pm 45 mg/dl) and Torre de Moncorvo (206 \pm 54 mg/dl), while the lowest was registered in Miranda do Douro (187 \pm 28 mg/dl).

The HDL-C fraction did not vary greatly, while LDL-C exhibited a noticeable variation, with maximum values in Alfândega da Fé (129 \pm 40 mg/dl) and Torre de Moncorvo (130 \pm 46 mg/dl) and lowest values in Miranda do Douro (113 \pm 28 mg/dl), showing again a close association with total cholesterol content.

The triglycerides profiles denoted high similarity; however, a tendency to higher values could be observed in Vinhais (148 \pm 62 mg/dl), while the lowest values tended to be obtained in Miranda do Douro (131 \pm 65 mg/dl). The glyceic parameters showed significant differences, with maximal values of glucose in Vinhais (143 \pm 59 mg/dl) and minimal values in Torre de Moncorvo (110 \pm 33 mg/dl); HbA1c was also higher in Vinhais, as well as in Miranda do Douro (7 \pm 2% in both localities), while

TABLE 2. Biochemical Parameters (mg/dl) in Function of the Selected Localities (L). The Results are Expressed as Mean \pm SD ($n = 947$; 446 Males and 501 Females). In Each Column and for Each Parameter, Different Letters Mean Significant ($P < 0.05$) Differences

		Total cholesterol	HDL-c	LDL-c	Triglycerides	Glucose	HbA1c (%)
Locality (L)	Alfândega da Fé	206 \pm 45 a	48 \pm 10	129 \pm 40 a	144 \pm 73	121 \pm 53 bc	6 \pm 2 b
	Bragança	198 \pm 42 ab	47 \pm 8	122 \pm 39 ab	142 \pm 68	133 \pm 58 ab	6 \pm 2 ab
	Miranda do Douro	187 \pm 28 b	48 \pm 8	113 \pm 28 b	131 \pm 65	140 \pm 50 ab	7 \pm 2 a
	Torre de Moncorvo	206 \pm 54 a	48 \pm 10	130 \pm 46 a	143 \pm 65	110 \pm 33 c	6 \pm 1 b
	Vinhais	195 \pm 45 ab	46 \pm 7	120 \pm 40 ab	148 \pm 62	143 \pm 59 a	7 \pm 2 a
	<i>P</i> -value	0.012	0.068	0.016	0.718	<0.001	<0.001
Gender (G)	Male	197 \pm 44	46 \pm 9	122 \pm 40	146 \pm 73	130 \pm 57	6 \pm 2
	Female	203 \pm 45	48 \pm 8	126 \pm 40	142 \pm 62	130 \pm 55	6 \pm 2
	<i>P</i> -value	0.100	0.060	0.155	0.956	0.886	0.326
L \times G	<i>P</i> -value	0.913	0.764	0.924	0.789	0.986	0.525

the lowest values were registered in Torre de Moncorvo and Alfândega da Fé (Table 2).

DISCUSSION

In general, the different age groups revealed higher homogeneity for all the assayed parameters, when compared with the locality influence. It should be pointed out that when the Tukey's classification could be applied, Bragança and Vinhais, two localities with geographical proximity, were equally classified in almost all cases, being observed a very similar outcome for Torre de Moncorvo and Alfândega da Fé, that are also very close to each other.

Regarding the effect induced by gender, the results did not reveal statistically significant differences among men and women. Nevertheless, the assayed female subjects tended to have higher total cholesterol, HDL-C, and LDL-C values, while male subjects leaned to present more elevated glucose levels.

For the cases in which the multiple comparisons could not be performed, the EMM plots (data not shown) confirmed the previously described tendencies, either concerning A or L influence.

Furthermore, most of the subjects had total cholesterol values ranging from 150 to 240 mg/dl, with results varying from 100 to 388 mg/dl. Considering reference values, the majority of population is located under the recommended value (<200 mg/dl) or moderate risk range (200–239 mg/dl), with only a small proportion of subjects with high risk (>240 mg/dl) of coronary artery disease. Regarding HDL-C, most subjects had results varying from 35 to 60 mg/dl, with a high number of subjects located on the moderate risk range (35–55 mg/dl), despite results varied from 26 to 105 mg/dl.

Concerning LDL-C, the obtained results varied from 17 to 290 mg/dl, with a major proportion concentrated between 70 and 150 mg/dl. It is noteworthy that only a small part of the population-presented values classified as high risk (>160 mg/dl) for coronary heart disease, while most subjects presented desirable values (<130 mg/dl). Triglycerides ranged from 42 to 395 mg/dl, with an elevated number of subjects presenting values in the 50–200 mg/dl range; the majority of population presented, in fact, a good result with concentrations under 150 mg/dl. However, there is also a considerable number of subjects in the high-risk range.

Regarding glycemie parameters, glucose varied from 65 to 405 mg/dl, with most subjects located in the 80–180 mg/dl range, a concerning result since the desirable values should not be higher than 110 mg/dl. In the case of HbA1c, the obtained results varied from 3.7 to 14.3%, with most subjects concentrated in the 4–8.5% range, re-

specting the adequate reference standard (7–8.5%), an important finding since HbA1c is the most adequate test for quantification of cronic complications hazzard in diabetes patients (14).

In order to understand if the studied parameters (continuous variables) were correlated, their frequency distribution was analyzed using a 3D density plot. These correlations were tested with at least one of the parameters presenting values out of the reference values, since this could act as a more reliable disease indicator. In none of the cases (one or both parameters out of reference values) it was found a good correlation among HDL-C and triglycerides, LDL-C and triglycerides, or LDL-C and total cholesterol. The only parameters that proved to be strongly correlated were glucose and HbA1c, either with one (Fig. 1A) or both (Fig. 1B) biochemical parameters outside of the reference limits.

Regarding the main objective of this work, the differences induced by A were evaluated by stepwise LDA, resulting in a discriminant model with three significant functions ($P < 0.05$ for Wilks' Λ test) (Fig. 2). Concerning A (Fig. 2A), the three functions explained 100.0% of the observed variance, with the first function responsible for 52.09%, the second 28.75%, and the third 19.16%. The discriminant power proportion of the n th function may be estimated by the ratio among its own value and the sum of all discriminant functions values. The canonical discriminant function coefficients allowed obtaining the following model:

$$\text{Function 1: } D_1 = 0.005 \times \text{triglycerides} + 0.028$$

$$\times \text{glucose} - 1.028 \times \text{HbA1c} + 2.225;$$

$$\text{Function 2: } D_2 = 0.012 \times \text{triglycerides} - 0.001$$

$$\times \text{glucose} + 0.281 \times \text{HbA1c} - 3.361;$$

$$\text{Function 3: } D_3 = -0.009 \times \text{triglycerides} + 0.020$$

$$\times \text{glucose} - 0.094 \times \text{HbA1c} - 0.722.$$

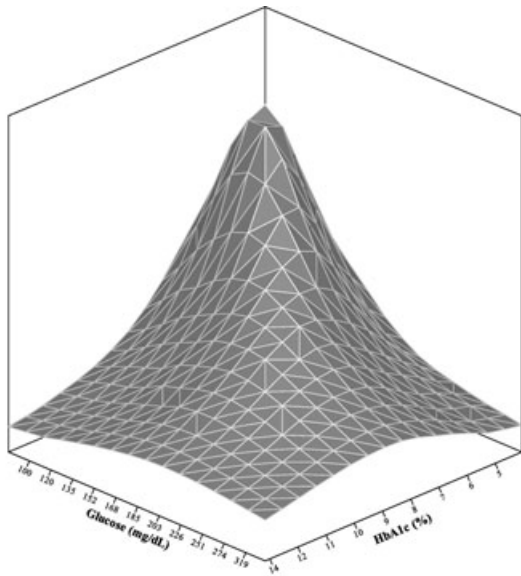
Regarding classification function coefficients, the following functions were obtained:

$$[21-30] c_1 = 0.019 \times \text{triglycerides} - 0.042 \times \text{glucose} \\ + 2.473 \times \text{HbA1c} - 7.108;$$

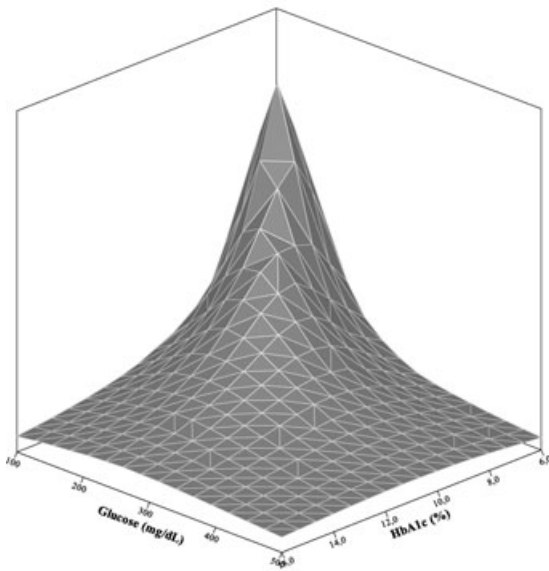
$$[31-40] c_2 = 0.022 \times \text{triglycerides} - 0.046 \times \text{glucose} \\ + 2.977 \times \text{HbA1c} - 9.907;$$

$$[41-50] c_3 = 0.021 \times \text{triglycerides} - 0.049 \times \text{glucose} \\ + 3.154 \times \text{HbA1c} - 10.495;$$

$$[51-60] c_4 = 0.025 \times \text{triglycerides} - 0.049 \times \text{glucose}$$



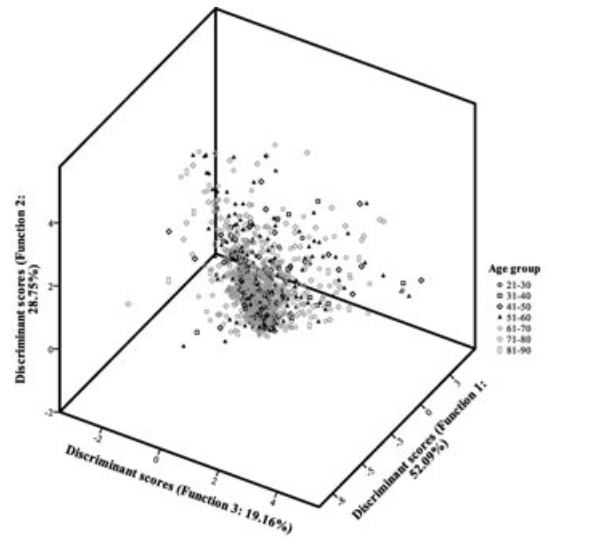
A



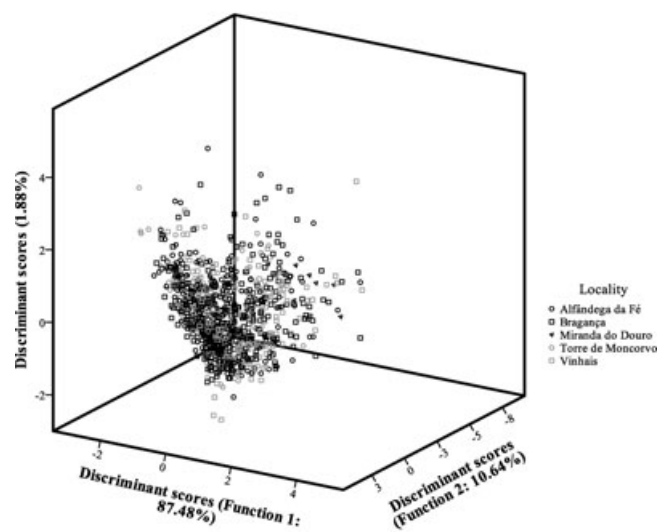
B

Fig. 1. 3D density graph plotting of glucose and HbA1c frequency distributions. (A) One of the parameters outside the reference limits; (B) both parameters outside the reference limits.

$$\begin{aligned}
 &+ 3.015 \times \text{HbA1c} - 10.307; \\
 [61-70] \ c_5 &= 0.021 \times \text{triglycerides} - 0.051 \times \text{glucose} \\
 &+ 3.149 \times \text{HbA1c} - 10.231; \\
 [71-80] \ c_6 &= 0.023 \times \text{triglycerides} - 0.055 \times \text{glucose} \\
 &+ 3.268 \times \text{HbA1c} - 10.628; \\
 [81-90] \ c_7 &= 0.021 \times \text{triglycerides} - 0.064 \times \text{glucose} \\
 &+ 3.491 \times \text{HbA1c} - 10.767.
 \end{aligned}$$



A



B

Fig. 2. Canonical analysis of age group (A) and locality (B) influence based on the biochemical assayed parameters.

However, the defined functions were not able to separate the results in individualized groups, confirming the already detected similarity among the obtained results. In fact, the selected variables did not show discriminant ability, as the classification performance indicated, with only 14.4% of the original groups and 13.3% of the cross-validated groups being correctly classified.

Regarding L influence (Fig. 2B), the three defined functions included 100.0% of the observed variance, distributed as 87.48% for the first, 10.64 for the second and

1.88 for the third function. The classification function coefficients resulted in the following model:

$$\text{Function 1: } D_1 = -0.013 \times \text{cholesterol} + 0.010 \\ \times \text{glucose} + 0.212 \times \text{HbA1c} - 0.119;$$

$$\text{Function 2: } D_2 = 0.001 \times \text{cholesterol} + 0.033 \\ \times \text{glucose} - 1.033 \times \text{HbA1c} + 1.980;$$

$$\text{Function 3: } D_3 = 0.019 \times \text{triglycerides} + 1E^{-5} \\ \times \text{glucose} + 0.280 \times \text{HbA1c} - 5.548.$$

However, and despite the found significant differences, the model was not able to separate the results in individual clusters corresponding to the five naturally occurring groups. In terms of classification ability, the results were again unsatisfactory, since only 8.9% of the original groups and 17.8% of the cross validates groups were correctly classified. Nevertheless, since the classification performance was slightly higher than the obtained for the age group influence, the classification function coefficients are also presented:

$$\text{Alfândega da Fé } c_1 = 0.103 \times \text{cholesterol} - 0.067 \\ \times \text{glucose} + 3.584 \times \text{HbA1c} \\ - 19.129;$$

$$\text{Bragança } c_2 = 0.099 \times \text{triglycerides} - 0.062 \\ \times \text{glucose} + 3.547 \times \text{HbA1c} \\ - 18.684;$$

$$\text{Miranda do Douro } c_3 = 0.093 \times \text{triglycerides} - 0.070 \\ \times \text{glucose} + 3.942 \times \text{HbA1c} \\ - 19.142;$$

$$\text{Torre de Moncorvo } c_4 = 0.104 \times \text{triglycerides} - 0.073 \\ \times \text{glucose} + 3.646 \times \text{HbA1c} \\ - 19.077;$$

$$\text{Vinhais } c_5 = 0.097 \times \text{triglycerides} - 0.064 \\ \times \text{glucose} + 3.746 \times \text{HbA1c} \\ - 19.373.$$

CONCLUSION

In general, the results proved to have higher similarity according with age group than with locality. This seems to indicate that the blood parameters references values exempt the inclusion of any age correction factors, since

the obtained values exhibit close amplitude and magnitude. Actually, this is also reflected in the similarity of the obtained coefficients, either for the canonical discriminant function, or for the classification function. This study revealed that, although the studied localities are relatively close, the results obtained for the assayed biochemical parameters revealed statistically significant differences. Surprisingly, this influence surpassed the effect of age group, but deeper conclusions could only be obtained with additional analysis, like specific diet or use of cholesterol control therapeutic agents. Nevertheless, this work contributed to a more comprehensive knowledge of lipidic and glycemic indexes among the assayed population. Future works should be oriented to establish correlation among these blood profiles and the emergence or development of some common metabolic diseases, allowing defining reliable biomarkers and helpful prevention strategies.

REFERENCES

1. Conroy RM, Pyörälä K, Fitzgerald AP, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J* 2003;24:987–1003.
2. Menown IBA, Murtagh G, Maher V, Cooney MT, Graham IM, Tomkin G. Dyslipidemia therapy update: The importance of full lipid profile assessment. *Adv Ther* 2009;26:711–718.
3. Menown IBA. HDL cholesterol—A new target for cardiovascular intervention. *Cardiol News* 2006;9:11–13.
4. Catapano AL, Reiner Ž, De Backer G, et al. ESC/EAS guidelines for the management of dyslipidaemias. The task force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS). *Atherosclerosis* 2011;217:3–46.
5. Borel JP, Maquart FX, Gillery P, Exposito M. *Bioquímica para o Clínico*. Instituto Piaget Ed., Lisboa, 2001.
6. Balkau B, Deanfield JD, Despres J-P, et al. International day for the evaluation of abdominal obesity (IDEA). *Circulation* 2007;116:1941–1951.
7. Tanaka M, Ishii H. Improvement in bedtime plasma glucose level serves as a predictor of long-term blood glucose control in patients with type 2 diabetes: a study with monotherapy of 50/50 premixed insulin analogue three times daily injection. *J Diabetes Complicat* 2011;25:83–89.
8. Burtis CA, Ashwood ER. *Fundamentals of Clinical Chemistry*, fourth edition, Philadelphia, W. B: Saunders Company, 1996.
9. Roche. *Biochemistry, Kits for analysis*, 2011.
10. Menarini. *Biochemistry, Kits for analysis*, 2011.
11. López A, García P, Garrido A. Multivariate characterization of table olives according to their mineral nutrient composition. *Food Chem* 2008;106:369–378.
12. Maroco J. *Análise Estatística, com utilização do SPSS*. Lisboa: Edições Sílabo; 2003.
13. Rencher AC. *Methods of Multivariate Analysis*. New York, John Wiley; 1995.
14. Bem AF, Kunde J. A importância da Determinação da Hemoglobina Glicada no Monitoramento das Complicações Crônicas do Diabetes *mellitus*. *J Bras Patol Med Lab* 2006;42:185–191.