



Article

Plasma Interleukin-10 and Cholesterol Levels May Inform about Interdependences between Fitness and Fatness in Healthy Individuals

Francesco Sartor ^{1,2,*}, Jonathan P. Moore ² and Hans-Peter Kubis ²

¹ Department of Patient Care and Monitoring, Philips Research, 5656 AE Eindhoven, The Netherlands

² College of Human Sciences, Bangor University, Bangor LL57 2EF, UK; j.p.moore@bangor.ac.uk (J.P.M.); h.kubis@bangor.ac.uk (H.-P.K.)

* Correspondence: francesco.sartor@philips.com; Tel.: +31-(0)-615-509-627

Abstract: Relationships between demographic, anthropometric, inflammatory, lipid and glucose tolerance markers in connection with the fat but fit paradigm were investigated by supervised and unsupervised learning. Data from 81 apparently healthy participants (87% females) were used to generate four classes of fatness and fitness. Principal Component Analysis (PCA) revealed that the principal component was preponderantly composed of glucose tolerance parameters. IL-10 and high-density lipoprotein, low-density lipoprotein (LDL), and total cholesterol, along with body mass index (BMI), were the most important features according to Random Forest based recursive feature elimination. Decision Tree classification showed that these play a key role into assigning each individual in one of the four classes, with 70% accuracy, and acceptable classification agreement, $\kappa = 0.54$. However, the best classifier with 88% accuracy and $\kappa = 0.79$ was the Naïve Bayes. LDL and BMI partially mediated the relationship between fitness and fatness. Although unsupervised learning showed that the glucose tolerance cluster explains the highest quote of the variance, supervised learning revealed that the importance of IL-10, cholesterol levels and BMI was greater than the glucose tolerance PCA cluster. These results suggest that fitness and fatness may be interconnected by anti-inflammatory responses and cholesterol levels. Randomized controlled trials are needed to confirm these preliminary outcomes.

Keywords: VO₂max; anti-inflammatory; machine learning; PCA



Citation: Sartor, F.; Moore, J.P.; Kubis, H.-P. Plasma Interleukin-10 and Cholesterol Levels May Inform about Interdependences between Fitness and Fatness in Healthy Individuals. *Int. J. Environ. Res. Public Health* **2021**, *18*, 1800. <https://doi.org/10.3390/ijerph18041800>

Academic Editor: Paul B. Tchounwou

Received: 21 January 2021

Accepted: 7 February 2021

Published: 12 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the 1950s, first observational evidence emerged showing that physically active individuals had a lower risk of cardiovascular disease (CVD) [1]. This evidence was later corroborated by the protective effect found for cardiorespiratory fitness (CRF), as shown in the Aerobics Center Longitudinal Study in 1989 [2,3]. Since then, several reviews, systematic reviews, and meta-analysis have confirmed and highlighted the protective role of CRF regardless the level of fatness [4–7]. According to the “fat but fit paradox”, people who have a high level of CRF may be better protected from the risk of CVD than leaner people who have low CRF [8]. However, only a small proportion of US citizens can be considered “fat and fit”, and obesity is independently associated with low CRF, simply because obese people are generally less active [9].

Lahoz-Garcia et al. [10] showed an interesting partial mediation of CRF between diet and obesity in schoolchildren, meaning that higher CRF contributes, for the same diet, to a lower fat mass (FM). Consistently, others have found that moderate to vigorous physical activity levels, thus higher CRF, were independently associated with a lower atherogenic index of plasma, namely blood fat strongly related with CVD, regardless of diet; and that central adiposity mediated, in other words explains, the relationship between moderate to vigorous physical activity levels and atherogenic index of plasma [11]. This would rule

in favor of the protective role of higher CRF against CVD risk. Moreover, poor CRF has been associated with glucose intolerance [12] and a higher risk of insulin resistance in apparently healthy individuals [13]. Furthermore, it has been hypothesized that low CRF could provide an early sign of insulin resistance [14].

Obesity has been shown to be associated with low level systemic inflammation in connection with increased adipose tissue mass [15,16]. In turn there is evidence, in animal studies, of the possible role of inflammation on over-nutrition [17]. However, physical activity may counteract over-nutrition behavior at the hypothalamic level by means of anti-inflammatory signaling mediated interleukin-10 (IL-10) [17]. An anti-inflammatory role of IL-10 has been found also in rat skeletal muscle tissue [18]. In humans it was found consistently that intensive cycling is able to increase, 1 hour after the exercise, gene expression of several interleukins including IL-10, but not IL-6 [19]. High intensity exercise showed an acute, 30 minutes, IL-10 and IL-6 increase in overweight-obese inactive individuals, but this increase was not elicited by moderate intensity exercise [20]. Nevertheless, two weeks of high intensity exercise in overweight-obese unfit individuals did not show a chronic increase in IL-10 nor in IL-6 [21,22]. Rather, a chronic elevation of IL-10 found in obese women was reduced by 12 weeks of lifestyle intervention, including 30 minutes of exercise a day, only in those obese women who did not have metabolic syndrome [23]. Furthermore, higher serum concentration of IL-10 was found in older adults with a higher volume of physical activity [24]. Additionally, animal models show a possible protective role of anti-inflammatory signaling on cardiac function (i.e., left ventricular end-diastolic pressure) [25], a finding supported in human studies involving coronary heart disease patients, obese and diabetic individuals [26,27].

To further investigate the relationship between cardiovascular fitness and body composition characteristics i.e., fatness, we used a database, which combined demographic, blood lipids, insulin resistance, and inflammatory variables in association with CRF and FM% values. Our approach was to create a categorical variable composed of four classes, based on CRF and FM% levels. The four classes or categories are termed High Fatness with High Fitness (HFHF), High Fatness with Low Fitness (HFLF), Low Fatness with High Fitness (LFHF) and finally Low Fatness with Low Fitness (LFLF). The cutoff levels between categories were identified according to the literature [28,29]. We have applied a data driven approach consisting of four steps. First is an unsupervised learning phase, where the variables are clustered using Principal Component Analysis (PCA) [30]. PCA allows clustering of the variables into principal components. Second, a supervised learning phase was deployed to use those clusters in the feature importance selection. We opted for feeding the PCA components as well as the other variables into the feature importance selection algorithm because, although PCA combines uncorrelated variables with one another in such a way that each principal component will maximize variance, this does not mean that the components per se will be the most important classification features. Therefore, as a second step, we have used the same categorical four classes' dependent variable for a random forest based feature importance selection. In detail, we have used the Boruta algorithm, which is an improvement of the Random Forest feature selection model, also known as recursive feature elimination [31,32]. The Boruta algorithm adds randomness to the importance evaluation algorithm, so that the certainty about the importance of a given variable is increased. In short, a randomized copy of the variables is made at each iteration of the random forest importance computation. Thus, if a variable has a higher importance than the maximal importance of all randomized attributes it is retained. If there is some uncertainty, or if a variable has a lower importance it is rejected or discarded [32].

Third, a decision tree was used in order to define the discriminating path to the four classes of fitness and fatness. This classification model was used to visualize which independent variables would best split the data points into the four classes. However, classification was not limited to the decision tree. Another four classification models were used as well with the intent of testing which classification model would maximize the use of the selected independent variables, or features. The four alternative machine learning

classification models were Multiple Logistic Regression, Decision Tree, Naïve Bayes, and K-nearest neighbors. This step was necessary to test whether the features selected would effectively classify the data points. Finally, a fourth step, a mediation and moderation analysis [33] was conducted in order to investigate whether attenuation between CRF and FM% would occur when one of the variables extracted was used as covariate. We hypothesized that we would find attenuations, as previously shown in the literature [10,11], by means of variables linked to fat metabolism. The overall aim of this study was to use a data driven approach, employing machine-learning techniques, to generate new insights connecting fitness and fatness with demographic, blood lipids, insulin resistance, and inflammatory variables.

2. Materials and Methods

2.1. Study Design and Participants

The data analyzed in this study originated from two separate data collections conducted at Bangor University. Data from 81 apparently healthy participants (10 males and 71 females) were included in the analysis. All participants were informed about the study protocols and objectives, and provided written consent prior to the start of the studies. Study protocols were approved by the Ethics Committee of the School of Sports, Exercise and Health Sciences Department of Bangor University in conformity with the Declaration of Helsinki. The design of this study was purely observational.

2.2. Body Composition, Fat Mass Percentage, Blood Markers and Cardiorespiratory Fitness Assessment

Participants were pre-screened for cardiovascular diseases by means of the American Heart Association/American College of Sports Medicine Pre-Participation Questionnaire [34]. However, participants with elevated fasting levels of glucose, insulin and lipids were not per se excluded from this study. Body composition, fasting blood lipid profile and CRF ($VO_2\max$) were determined using standardized protocols described previously [21]. A cardiorespiratory fitness test was executed on a cycle ergometer (Corival 400, Lode, Groningen, The Netherlands), the protocol consisted of an incremental exercise test to exhaustion (1min at 50 + 20 W increments per minute). Oxygen uptake was measured breath by breath by means of a metabolic card (ZAN 600 CPET, Oberthulba, Germany). Fasting blood lipid profile (total Cholesterol, LDL and HDL), plasma insulin, plasma glucose, leptin and cytokines (IL-6, IL-10, and TNF- α) collection and analysis is also described in Sartor et al. [21]. Plasma glucose was analyzed by immobilized enzymatic assay (YSI 2300 STAT, Incorporated Life Sciences, Yellow Springs, OH, USA). Lipid profile was analyzed from plasma samples by optic enzymatic assay (Reflotron[®], Roche Diagnostics, Mannheim, Germany). Plasma insulin was analyzed by ELISA (ultrasensitive human insulin ELISA kit, Mercodia, Uppsala, Sweden). Cytokines (IL-10, IL-6 and TNF- α) and adipokines were also analyzed from fasting plasma samples by ELISA (Bender MedSystems GmbH, Austria and BioVendor, Laboratóní medicína, Czech Republic, respectively). Insulin sensitivity and β -cell function were estimated using fasting plasma insulin and glucose by means of the Homeostatic model assessment 2 (HOMA2) [35].

2.3. Classification Criteria

Four classes were extracted from the database described above; a Higher-Fatness with Higher-Fitness (HFHF) group, a Higher-Fatness with Lower-Fitness (HFLF) group, a Lower-Fatness with Higher-Fitness (LFHF) group, and finally a Lower-Fatness with Lower-Fitness (LFLF) group. The grouping criteria were taken from Gallagher et al. [28] for fatness, and the American College of Sports Medicine guidelines [29] for fitness. The criteria are represented in Table 1.

Table 1. Classification criteria for body fat percentage and relative VO₂max (mL/kg/min), age and sex.

Age	Males	Females
Young	if AGE < 40 years AND if Sex = 1 AND FatMass% ≥ 26 then Higher-Fatness	Elseif Sex = 0 AND FatMass% ≥ 39 then Higher-Fatness
Middle-Age	if 59 ≥ AGE ≥ 40 AND if Sex = 1 AND FatMass% ≥ 29 then Higher-Fatness	Elseif Sex = 0 AND FatMass% ≥ 41 then Higher-Fatness
Older	if AGE ≥ 60 AND if Sex = 1 AND FatMass% ≥ 31 then Higher-Fatness	Elseif Sex = 0 AND FatMass% ≥ 43 then Higher-Fatness
Young/Middle/Older	Else Lower-Fatness	Else Lower-Fatness
Young	If AGE < 29 AND if Sex = 1 AND if relVO ₂ max > 45.7 then Higher-Fitness	Elseif Sex = 0 AND if relVO ₂ max > 39.5 then Higher-Fitness
Middle-Age	If 39 ≥ AGE > = 30 AND if Sex = 1 AND if relVO ₂ max > 44.4 then Higher-Fitness	Elseif Sex = 0 AND if relVO ₂ max > 36.7 then Higher-Fitness
	If 49 ≥ AGE ≥ 40 AND if Sex = 1 AND if relVO ₂ max > 42.4 then Higher-Fitness	Elseif Sex = 0 AND if OreIVO ₂ max > 35.1 then Higher-Fitness
Older	If AGE > 50 AND if Sex = 1 AND if relVO ₂ max > 38.3 then Higher-Fitness	Elseif Sex = 0 AND if OreIVO ₂ max > 31.4 then Higher-Fitness
Young/Middle/Older	Else Lower-Fitness	Else Lower-Fitness

2.4. Data Analytics

2.4.1. Preprocessing

The full dataset collected at Bangor University premises was loaded into RStudio (Version 1.2.5033, 2009–2019 RStudio Inc., Boston, MA, USA). This initial dataset included 25 independent variables. A first missing data filter was applied and all variables with more than 70% missing data were discarded. After this step, 19 independent variables were retained. Two variables were converted into factorial variables, the classification variable as explained in Table 1 and the variable Sex. The retained variables were visualized to reveal imbalance. This visualization showed an imbalance towards females, as they represented 87% of our dataset. The imbalance was a consequence of the original research question of one data collection being confined to females. A zero- and near zero-variance predictors analysis was conducted, by means of nearZeroVar function (caret R package), to eliminate any independent variables that would not add anything in explaining variance (Table 2). However, no variables were rejected based on these criteria [36]. The preProcess function (caret R package) was used to center and scale the variables and missing data, were imputed using the bagImpute function which uses the bootstrap aggregating method [37]. Outliers were detected as values outside boxplot notches, using boxplot function (graphics R package). The notches were set as the median, plus or minus the standard error [38]. The detected outliers were excluded from the analysis.

Table 2. Zero- and near zero-variance predictors analysis.

	Frequency Ratio	Percent Unique	Zero Variance	Near Zero Variance
Sex	7.100000	2.469136	FALSE	FALSE
Age	1.555556	23.456790	FALSE	FALSE
Height	1.142857	28.395062	FALSE	FALSE
Weight	1.000000	77.777778	FALSE	FALSE
BMI	1.500000	75.308642	FALSE	FALSE
Chol	1.000000	72.839506	FALSE	FALSE
HDL	1.000000	62.962963	FALSE	FALSE
LDL	1.333333	69.135802	FALSE	FALSE
TG	5.250000	49.382716	FALSE	FALSE
Fgluc	1.500000	62.962963	FALSE	FALSE
Leptin	1.000000	76.543210	FALSE	FALSE
Insulin	1.000000	77.777778	FALSE	FALSE
BetacellF	1.000000	77.777778	FALSE	FALSE
InsSens	1.000000	80.246914	FALSE	FALSE
InsRes	1.000000	43.209877	FALSE	FALSE
TNFalpha	1.333333	66.666667	FALSE	FALSE
IL-6	1.333333	71.604938	FALSE	FALSE
IL-10	1.000000	50.617284	FALSE	FALSE
RER	1.200000	34.567901	FALSE	FALSE

BMI = Body Mass Index, Chol = Fasting Total Cholesterol, HDL = Fasting High Density Lipoprotein, LDL = Fasting Low Density Lipoprotein, TG = Fasting TriGlycerides, Fgluc = Fasting Glucose, BetacellF = β cell Function, InsSens = Insulin Sensitivity, InsRes = Insulin Resistance, TNFalpha = Tumor Necrosis Factor α , IL-6 = Interleukin-6, IL-10 = Interleukin-6, RER = Respiratory Exchange Ratio.

2.4.2. Principal Component Analysis and Feature Selection

Once the data were pre-processed a principal component analysis was conducted to find what combination of variables would explain the variability of the data. The function PCA (FactoMineR R package) as described in [39] was used. Eigenvalues, which represent the amount of the variation explained by each principal component, were extracted by `fviz_eig`. The number of retained components was set so that 70% of the total variance is explained. Correlation plots of all variables were produced using the `corrplot` function (`corrplot` R package). The importance of the twenty variables including five new Principal Components was evaluated by a recursive feature elimination technique based on the Boruta Random Forest method (Boruta R package) [32]. The Boruta function compares original importance attributes against importance achievable by shadow random variables, in iterations until convergence. The principal components were also included in the feature selection step, to test whether the most variation corresponded with the highest importance.

2.4.3. Decision Tree

A decision tree was built using the nine variables selected by the Boruta algorithm, with the exclusion of the PCA dimensions. As first step, the class imbalance was compensated by means of weights for simple random sample (i.e., $1/\text{probability}$). The decision tree was constructed using the `rpart` function (`rpart` R package) and vitalized by `rpart.plot` (`rpart.plot` R package). Tree depth was set as the smallest tree within one standard error of the minimum cross validation error [40].

2.4.4. Classification Models

Multiple logistic regression, decision tree, naïve Bayes, and κ -nearest neighbors classification model were trained on our dataset by means of the `train` function (`caret` R package) as described in Kuhn [36]. The classes were the four subgroups (HFHF, HFLF, LFHF, LFLF) described above. In order to perform the multinomial logistic regression, the multinom method was selected within the `train` function. In order to evaluate the performance of each single classifier, accuracy tables and confusion matrices were generated, using the `confusionMatrix` in `caret` and visualized thanks to `ggplot` (`ggplot2` R package) [36].

2.4.5. Mediation and Moderation Analysis

Mediation analysis was conducted by means of the mediation R package [41]. Before analyzing, the mediation and moderation raw data for each variable were assessed for normality and linearity by means of quantile-quantile plots (qqnorm function, from the basic stats R package), centered, and scaled when required, as described earlier. Linear regressions models, via the lm function (stats R package), were built between the mediator and the independent variable (relative VO_2max), and between the dependent variable (Fat Mass percentage) and the independent variable-mediator combined. The mediate function simulated the comparison between these two linear regressions, showing if the mediation would add a significant contribution in relating the independent and dependent variables. The mediation analysis resulted in the Average Causal Mediation Effects (ACME), the Average Direct Effects (ADE), and the combined effects (Total Effect), and the proportion mediated (Prop. Mediated). Moderation was executed by the gylma and stargazer R packages. A linear model was built between the dependent variable and independent variable plus the moderator, and between the dependent variable and the moderator plus the product.

2.5. Statistical Analysis

The descriptive statistics, means and standard deviations of all participants for the 15 included variables and for each of the four subgroups were analysed using the arsenal R package [42]. Data for the four subgroups were split using the filter function supported by the dplyr R package. One-way ANOVAs were performed to compare the four sub-groups and they were followed-up when appropriate both by the tableby function (arsenal R package). Significance level was set at 0.05.

3. Results

3.1. Subgrouping and Difference Analysis

As described in the method section, four subgroups were derived according to participants' CRF, body FM%, age, and sex. The subgroups sizes are not evenly distributed. Two subgroups HFHF and LFHF are rather small ($N = 9$, $N = 6$, respectively). In line with our intention to form four groups of different fatness and fitness levels, the ANOVA and follow-up showed significant differences between the two higher-fitness and lower-fitness levels. Moreover, the HFHF group and the LFHF groups also showed a significant difference in fitness, the lower in fatness being fitter (40.1 ± 2.9 mL/kg/min) than the higher in fatness (34.3 ± 4.3 mL/kg/min). As for the higher fatness/lower fatness split, this was fully achieved, as confirmed by the ANOVA and follow-ups (Table 3). As to be expected, BMI was significantly higher in the HFLF group compared with the LFHF and LFLF subgroups. There was a trend towards a higher BMI for the HFLF group when compared with the HFHF group, and a trend towards a higher BMI in the HFHF group compared with the LFHF group. It is to be noted that BMI does not fully reflect FM% (Table 3). Total fasting plasma Cholesterol levels showed significantly higher levels in the HFLF compared with the HFHF and LFHF groups. There was a strong trend towards a higher cholesterol level in the LFLF group compared with the HFHF group. The LFHF group showed higher HDL than the HFHF group. The LFLF group had a higher HDL level than the HFLF group. Moreover, there were two strong trends for a higher HDL in the LFLF group and the LFHF group versus the HFHF and the HFLF groups, respectively. LDL was higher in the HFLF group compared with the HFHF, LFHF, and LFLF groups. Finally, fasting plasma insulin was higher in the HFLF compared with the HFHF. Interestingly two, LFHF and LFLF, groups showed higher insulin values than the HFHF group (Table 3).

Table 3. Descriptive Statistics of Database, difference analysis and follow-up analyses.

	HFHF (N = 9)	HFLF (N = 47)	LFHF (N = 6)	LFLF (N = 19)	Total (N = 81)	ANOVA <i>p</i> Value	<i>t</i> -Test Follow-Up HFHF vs. HFLF <i>p</i> Value	<i>t</i> -Test Follow-Up HFHF vs. LFHF <i>p</i> Value	<i>t</i> -Test Follow-Up HFHF vs. LFLF <i>p</i> Value	<i>t</i> -Test Follow-Up HFLF vs. LFHF <i>p</i> Value	<i>t</i> -Test Follow-up HFLF vs. LFLF <i>p</i> Value	<i>t</i> -Test Follow-Up LFHF vs. LFLF <i>p</i> Value
Relative VO ₂ max (mL/kg/min)						<0.001	<0.001	0.013	<0.001	<0.001	0.346	<0.001
Mean(SD)	34.349 (4.237)	25.492 (6.432)	40.123 (2.992)	26.968 (3.272)	27.906 (6.932)							
Range	29.560–42.530	14.050–41.700	35.680–44.400	19.300–31.500	14.050–44.400							
Fat Mass %						<0.001	0.108	0.003	0.003	<0.001	<0.001	0.065
Mean(SD)	41.951 (5.686)	45.842 (6.686)	31.395 (4.828)	35.548 (4.510)	41.925 (7.871)							
Range	32.100–47.500	29.800–57.240	25.400–37.750	26.180–40.500	25.400–57.240							
Age, yrs						0.063						
Mean (SD)	42.444 (7.764)	34.787 (13.454)	24.500 (8.666)	33.526 (12.624)	34.580 (12.866)							
Range	33.000–50.000	19.000–57.000	19.000–42.000	20.000–49.000	19.000–57.000							
BMI						0.003	0.063	0.099	0.687	0.009	0.005	0.454
Mean (SD)	31.174 (1.572)	33.728 (3.949)	29.165 (2.828)	30.577 (4.217)	32.367 (4.061)							
Range	27.580–33.080	26.970–44.990	25.000–31.440	25.300–39.230	25.000–44.990							
Height, m						0.894						
Mean (SD)	1.671 (0.114)	1.662 (0.092)	1.657 (0.047)	1.681 (0.099)	1.667 (0.093)							
Range	1.570–1.950	1.500–1.950	1.580–1.710	1.540–1.950	1.500–1.950							
Weight, kg						0.138						
Mean (SD)	87.458 (13.453)	93.392 (14.219)	80.335 (10.806)	87.141 (18.855)	90.299 (15.427)							
Range	67.990–119.050	63.400–125.690	62.500–91.630	61.750–125.690	61.750–125.690							
Cholesterol, mmol/L						0.006	0.003	0.995	0.055	0.013	0.447	0.113
Mean (SD)	3.839 (0.623)	4.756 (0.830)	3.837 (0.753)	4.573 (1.003)	4.543 (0.904)							
Range	3.210–5.020	2.590–6.260	2.830–4.970	3.170–6.320	2.590–6.320							
HDL, mmol/L						0.008	0.900	0.025	0.056	0.057	0.004	0.873
Mean (SD)	1.023 (0.205)	1.040 (0.395)	1.368 (0.327)	1.407 (0.554)	1.149 (0.445)							
Range	0.610–1.420	0.370–2.490	1.110–1.790	0.700–2.590	0.370–2.590							
LDL, mmol/L						<0.001	0.002	0.367	0.388	<0.001	0.037	0.186
N-Miss	0	1	0	0	1							
Mean (SD)	2.456 (0.509)	3.221 (0.667)	2.165 (0.699)	2.769 (1.006)	2.949 (0.816)							
Range	1.860–3.240	1.910–4.460	1.230–2.890	0.360–4.560	0.360–4.560							
Fglucose, mmol/L						0.237						
N-Miss	0	0	0	1	1							
Mean (SD)	5.416 (0.652)	5.078 (0.802)	4.665 (0.565)	5.030 (0.351)	5.074 (0.700)							
Range	4.710–6.360	3.850–9.050	3.800–5.300	4.190–5.460	3.800–9.050							
Leptin, ng /mL						0.118						
N-Miss	0	0	1	0	1							
Mean (SD)	16.966 (10.307)	29.711 (15.740)	22.058 (18.130)	29.883 (16.539)	27.840 (15.895)							

Table 3. Cont.

	HFHF (N = 9)	HFLF (N = 47)	LFHF (N = 6)	LFLF (N = 19)	Total (N = 81)	ANOVA <i>p</i> Value	<i>t</i> -Test Follow-Up HFHF vs. HFLF <i>p</i> Value	<i>t</i> -Test Follow-Up HFHF vs. LFHF <i>p</i> Value	<i>t</i> -Test Follow-Up HFHF vs. LFLF <i>p</i> Value	<i>t</i> -Test Follow-Up HFLF vs. LFHF <i>p</i> Value	<i>t</i> -Test Follow-up HFLF vs. LFLF <i>p</i> Value	<i>t</i> -Test Follow-Up LFHF vs. LFLF <i>p</i> Value
Range	1.380–26.760	2.690–59.970	3.070–48.840	5.470–57.640	1.380–59.970							
Insulin, pmol/L						0.040	0.012	0.023	0.039	0.630	0.184	0.722
N-Miss	0	0	1	0	1							
Mean (SD)	5.667 (2.978)	11.053 (6.038)	9.722 (2.363)	9.021 (4.129)	9.881 (5.415)							
Range	1.730–9.640	1.210–28.090	7.350–12.400	2.410–17.470	1.210–28.090							
TNFalpha, pg/mL						0.992						
Mean (SD)	1.411 (1.669)	1.476 (2.253)	1.188 (1.952)	1.432 (2.046)	1.437 (2.093)							
Range	0.280–4.920	0.240–10.900	0.270–5.170	0.240–7.070	0.240–10.900							
IL-6, pg/mL						0.045	0.087	0.418	0.183	0.028	0.394	0.061
Mean (SD)	1.609 (0.525)	1.118 (0.811)	1.933 (0.983)	1.294 (0.589)	1.274 (0.777)							
Range	0.800–2.200	0.000–3.120	0.380–3.040	0.190–2.250	0.000–3.120							
IL-10, pg/mL						0.138						
N-Miss	0	0	1	0	1							
Mean (SD)	0.864 (0.224)	0.841 (0.332)	1.130 (0.848)	1.108 (0.662)	0.925 (0.470)							
Range	0.430–1.190	0.030–1.700	0.030–2.370	0.040–2.250	0.030–2.370							

HFHF = Higher-Fatness with Higher-Fitness group, HFLF = Higher-Fatness with Lower-Fitness group, LFHF = Lower-Fatness with Higher-Fitness group, LFLF = Lower-Fatness with Lower-Fitness group, VO₂max = maximal oxygen uptake, BMI = Body Mass Index, HDL = Fasting High Density Lipoprotein, LDL = Fasting Low Density Lipoprotein, TNFalpha = Tumor Necrosis Factor α , IL-6 = Interleukin-6, IL-10 = Interleukin-6. Significant p-levels are highlighted in bold.

3.2. Principal Component Analysis

The independent variables, once filtered for missing data, were clustered by means of principal component analysis. Five principal component dimensions were found that explained 70% of the variance (Figure 1). Dimension 1 was dominated by glucose tolerance features, dimension 2 by Leptin and Sex, dimension 3 was constituted by lipid profile, dimension 4 by triglycerides and glucose, and, finally, dimension 5 by BMI and weight. (Figure 1). In Figure 2 the classification and the weight of the single individuals is shown when the first two components are put in relation.

These five dimensions were further included in the feature selection process. Recursive feature elimination based on random forest showed that the stronger features in describing the four groups were IL-10, BMI, total cholesterol, HDL, LDL, dimension 1, beta cell function, dimension 4, IL-6, Age, dimension 3, and weight. In Figure 2 the interrelationship of the first two PCA components is shown and the four groups are clustered. Fitter groups tend to develop along dimension 1 while the less fit along dimension 2.

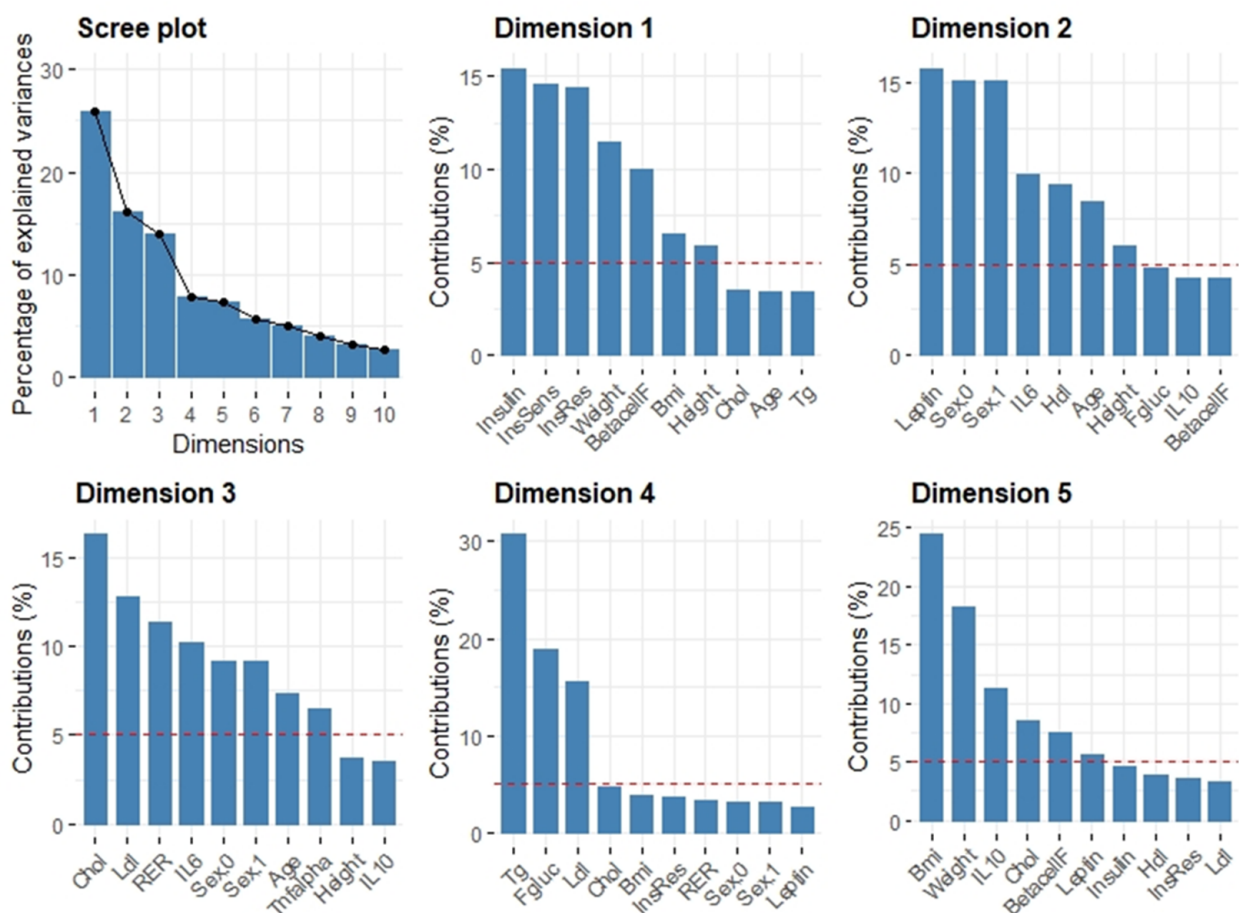


Figure 1. Output of the principal component analysis: BMI = Body Mass Index, Chol = Fasting Total Cholesterol, HDL = Fasting High Density Lipoprotein, LDL = Fasting Low Density Lipoprotein, TG = Fasting TriGlycerides, Fgluc = Fasting Glucose, Betacellf = β cell Function, InsSens = Insulin Sensitivity, InsRes = Insulin Resistance, TNFalpha = Tumor Necrosis Factor α , IL-6 = Interleukin-6, IL-10 = Interleukin-6, RER = Respiratory Exchange Ratio, Sex.0 = females, Sex.1 = males.

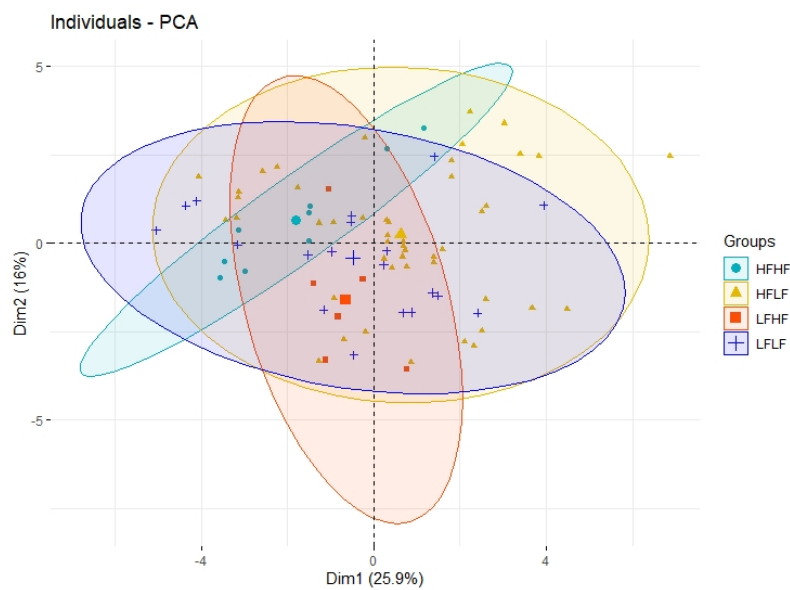


Figure 2. Clustering of the categorical variable, including the four fatness and fitness permutations. Relationship between the first principal component and the second principal component computed by PCA. The size of the icons for the single individuals shows their weight in classification. HFHF = Higher-Fatness with Higher-Fitness group, HFLF = Higher-Fatness with Lower-Fitness group, LFHF = Lower-Fatness with Higher-Fitness group, LFLF = Lower-Fatness with Lower-Fitness group.

3.3. Classification Models

The Random Forest based recursive feature elimination Boruta algorithm found twelve variables as certainly important in classifying the four fatness and fitness classes (Figure 3). Amongst these twelve are PCA dimensions 1,4 and 3, in order of importance. While the algorithm is uncertain about dimension 5 and discards dimension 2. IL-10, BMI, and cholesterol levels are clearly the most important variables. In Figure S1 the first 10 selected variables are shown as boxplot. Additionally, in Figure S1 linear correlations between variables are displayed, showing how the retained variables still carry most of the correlations.

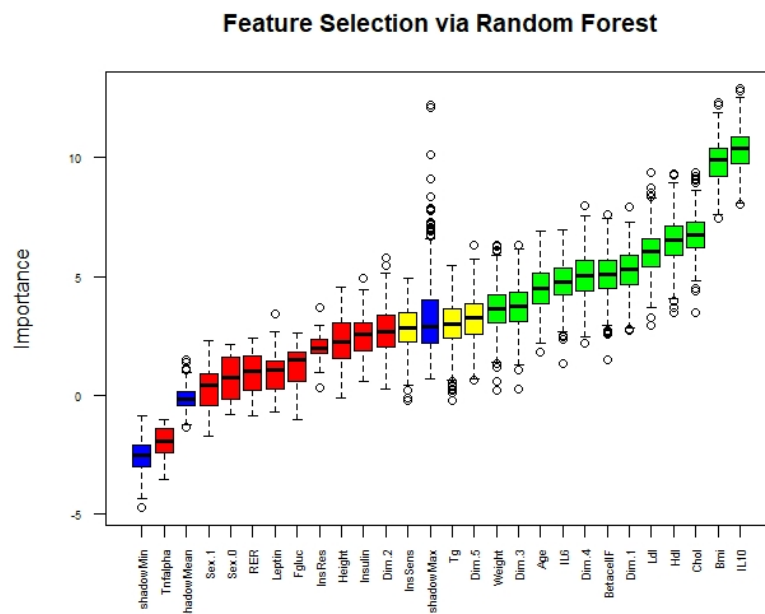


Figure 3. Recursive feature elimination; in green are depicted the variables that are certain. IL6. = Interleukin-6, IL-10 = Interleukin-6, RER = Respiratory Exchange Ratio, Sex.0 = females, Sex.1 = males, Dim.1 = Dimension 1 of the PCA, Dim.2 = Dimension 2 of the PCA Dim.3 = Dimension 3 of the PCA, Dim.4 = Dimension 4 of the PCA, Dim.5 = Dimension 5 of the PCA.

When the twelve variables, including the PCA dimensions, selected by the Boruta importance algorithm were used to generate the classification model, we found acceptable classification performances. In fact, the Multiple Logistic Regression model showed a classification accuracy of 0.77 (95% CI: 0.6717, 0.8627), significantly higher than the No Information Rate (0.4691), and a κ -coefficient of 0.65, Figure 4. The Decision Tree model, displayed in Figure 5, although having the lowest accuracy (0.70, 95% CI: 0.5919, 0.8001) amongst the models generated here, still had an accuracy significantly higher than its No Information Rate (0.432), and an acceptable κ -coefficient (0.54) (Figure 4). The Naïve Bayes classifier showed the highest accuracy (0.88, 95% CI: 0.7847, 0.9392), significantly higher than the No Information Rate (0.58), and a moderate κ -coefficient equal to 0.79 (Figure 4). Finally, the K-Nearest Neighbors classifier had an accuracy of 0.73 (95% CI: 0.6181, 0.8213), which was, however, not higher than the No Information Rate (0.76), with a rather weak agreement, a κ -coefficient of 0.47 (Figure 4). Overall, the latter performed worse than the other classification models.

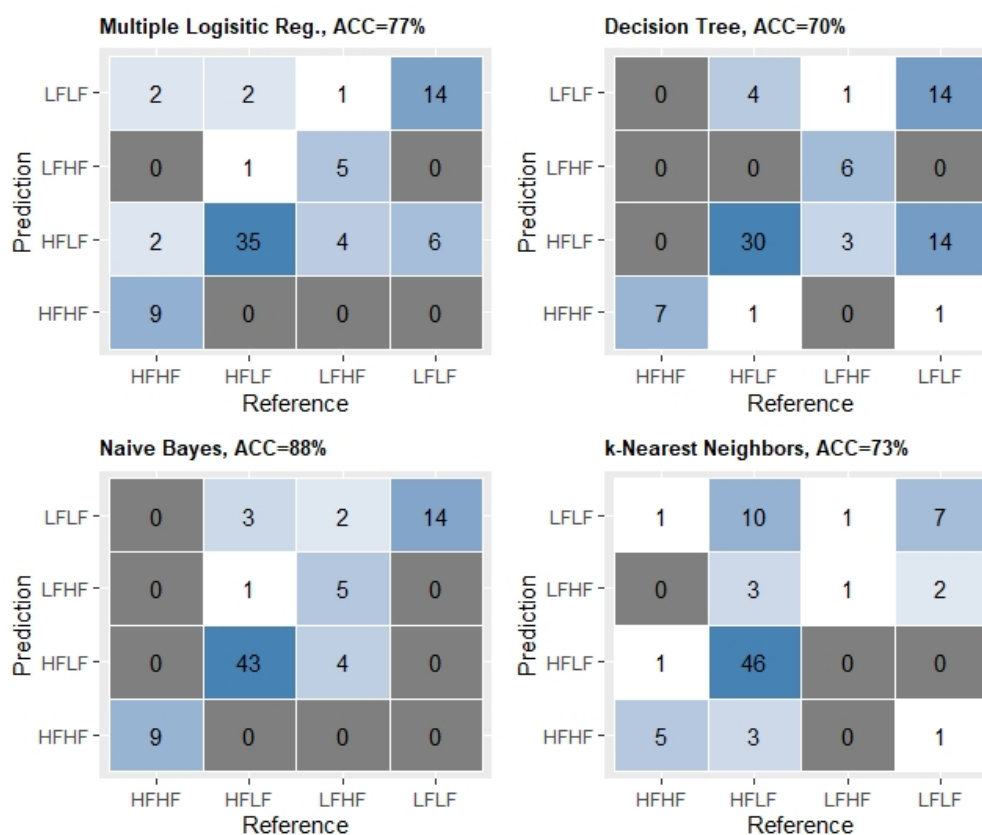


Figure 4. Confusion Matrices, and accuracy of the four classification models. ACC = accuracy, HFHF = Higher-Fatness with Higher-Fitness group, HFLF = Higher-Fatness with Lower-Fitness group, LFHF = Lower-Fatness with Higher-Fitness group, LFLF = Lower-Fatness with Lower -Fitness group.

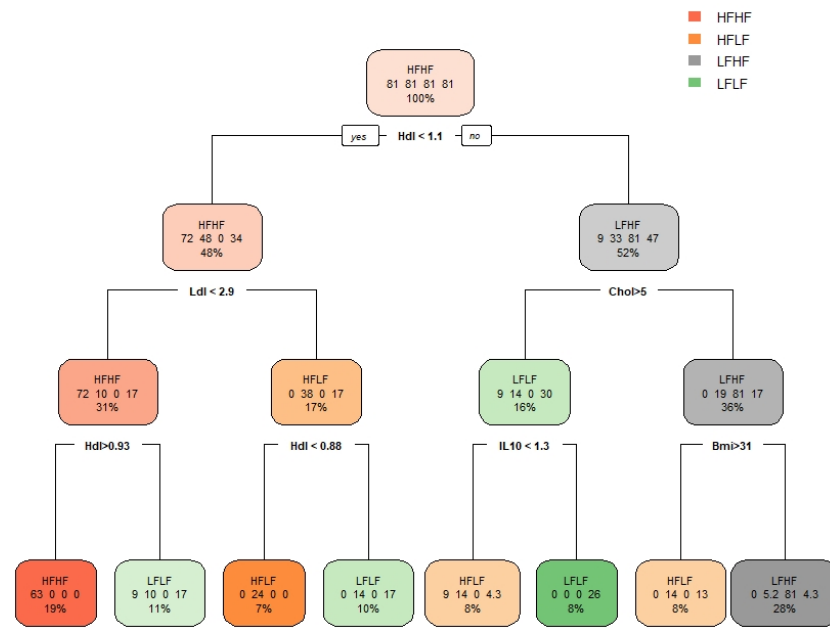


Figure 5. Decision tree, where HDL = High Density Lipoprotein, LDL = Low Density Lipoprotein, IL-10 = Interleukin-10 and BMI = Body Mass Index are expressed in their original dimensions (mmol/L, mmol/L, pg/mL, respectively). HFHF = Higher-Fatness with Higher-Fitness group, HFLF = Higher-Fatness with Lower-Fitness group, LFHF = Lower-Fatness with Higher-Fitness group, LFLF = Lower-Fatness with Lower-Fitness group.

3.4. Mediation and Moderation Analysis

All selected variables were analyzed for mediation and moderation. As shown by the quantile-quantile plots in Figure 6, LDL and BMI did not require further scaling and/or centering and were the only two variables to show a significant partial mediation effect between CRF and FM% (Figure 7). Details of the causal mediation analysis are captured in Table 4.

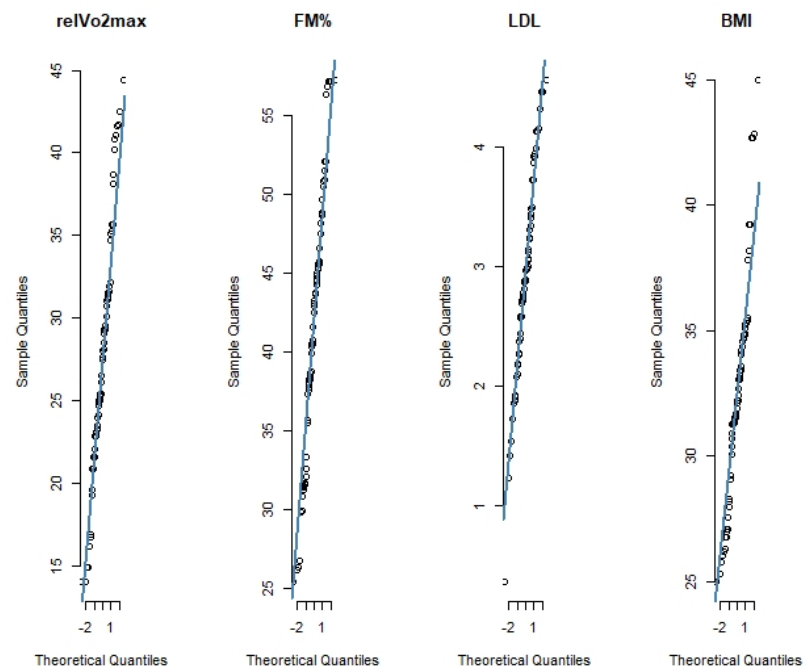


Figure 6. Quantile-quantile plots of the variables that showed partial mediation.

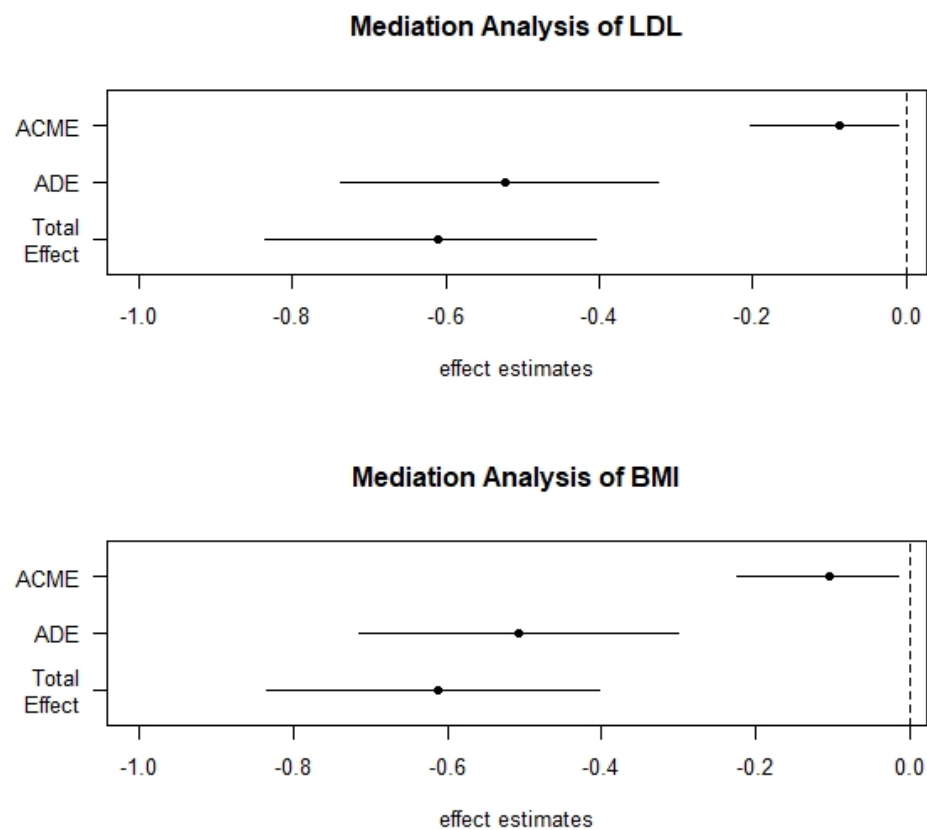


Figure 7. Decomposed Mediation Analysis plot: ACME = Average Causal Mediation Effect, ADE = Average Direct Effect, LDL = Low Density Lipoprotein, BMI = Body Mass Index.

Table 4. Causal Mediation Analysis, Quasi-Bayesian Confidence Intervals.

	Estimate	95% CI Lower	95% CI Upper	p-Value
ACME (LDL)	-0.0843	-0.1813	-0.01	0.024 *
ADE (LDL)	-0.5221	-0.7414	-0.30	<0.001 ***
Total Effect LDL)	-0.6063	-0.8271	-0.40	<0.001 ***
Prop. Mediated (LDL)	0.1308	0.0164	0.31	0.024 *
ACME (BMI)	-0.1078	-0.2205	-0.02	0.012 *
ADE (BMI)	-0.4996	-0.7034	-0.30	<0.001 ***
Total Effect (BMI)	-0.6075	-0.8211	-0.40	<0.001 ***
Prop. Mediated (BMI)	0.1728	0.0397	0.36	0.012 *

LDL = Low Density Lipoprotein, BMI = Body Mass Index, ACME = Average Causal Mediation Effect, ADE = Average Direct Effect, Prop. Mediated = Proportion of the effect Mediated. Significant values: *** <0.001, * <0.05. N = 81, Simulations: 1000.

4. Discussion

This present study embraces artificial intelligence as a tool to provide new insight into the fat but fit paradox [8]. Using unsupervised and supervised machine learning approaches to interrogate existing physiological data, this work indicates connection between markers of dyslipidemia, inflammation and cardiorespiratory fitness that reveal possible functional interaction of physiological systems underpinning the “fat but fit paradox”.

4.1. Descriptive Statistics in Relation to Fatness and Fitness

We have created four classes, or groups, in line with population normative cut-off values [28,29]. Consistently, these groups differed significantly from one another in terms of fitness and fatness (Table 3). Fasting total cholesterol levels and LDL were significantly higher in the HFLF group, while HDL was higher in the groups with lower fatness. The decision tree depicted in Figure 5 shows how well HDL and LDL alone could differen-

tiate the HFHF group from the other groups. Although IL-10 did not show significant differences between the four groups, whereas IL-6 did, IL-10 seemed to be involved in the differentiation of individuals with lower fitness level in function of their fatness (Figure 5). Moreover, the Analysis of Variance amongst the four groups also showed differences in fasting insulin levels. Fasting insulin was the highest in the HFLF group and the lowest in the HFHF group (Table 3). This is of particular interest because it seemed to be associated with fitness rather than with fatness levels. Fitness has been shown to play an important role in protecting against glucose intolerance [43]. This may be related to the well-known effect of muscle contractile activity, hence exercise training, on insulin sensitivity [44].

4.2. Machine Learning

Principal component analysis clustered the various markers available in this study so that they could better explain the variance of the fatness and fitness categorical variable. This resulted in PCA Dimension 1, mainly composed of glucose tolerance indicators, such as fasting insulin, insulin sensitivity, and insulin resistance, as well as beta cell function derived from the HOMA2 model (Figure 1). However, supervised learning, namely the random Forest based feature selection algorithm, revealed that the importance of IL-10, cholesterol levels (i.e., HDL, LDL and total Cholesterol) along with BMI in classifying the four classes was greater than that of the above mentioned glucose tolerance PCA cluster. The interesting aspect of our approach is that our analysis clearly points towards dominant features, namely IL-10, LDL, HDL, BMI, for categorizing our four groups, in competition with other features, which are just as well known to be influenced by fatness and fitness. Besides the potential exercise dependent link between IL-10 and insulin/leptin sensitivity in the hypothalamus in animal studies [17], exercise was found to increase IL-10 levels in overweight-obese human subjects [20]. An interlink between fatness and IL-10, however, was found in obese subject after weight loss, revealing higher IL-10 levels [45]. Therefore, distinct features of our data could point towards an important discriminating function of IL-10 and LDL/BMI for fatness and fitness classification and could be linked to these findings. Moreover, exercise has been found to effect LDL as well as HDL levels [46].

4.3. Partial Mediation

Fatness and fitness are significantly inversely related [47]. This was confirmed by our data. In addition to this, however, we found that CRF is indirectly related to FM% through the mediation of LDL and BMI. Previous literature found that BMI could mediate CRF and cardio-metabolic risk in schoolchildren [48]. Another investigation in schoolchildren using a large dataset showed that CRF may have a beneficial effect on lipid profile, insulin metabolism and inflammation independent of fatness [49]. Our results seem to lead in the same direction. Specific effects of exercise training, of a high enough intensity, to promote aerobic capacity improvements have been linked to a decrease in concentration of atherogenic ox-LDL [46,50]. In addition, upregulation of fatty acid metabolism and transport through exercise dependent signaling pathways (particularly peroxisome proliferator-activated receptor) [51,52] and concurrent alterations in lipid profiles [53,54] are well described. Interestingly, IL-10 was found to be linked with LDL level as IL-10 was shown to induce uptake of LDL by fluid-phase endocytoses in macrophages leading to lowered LDL plasma levels [55].

4.4. Implications

We are aware that our study is retrospective. Thus, it provides a limited level of evidence. It is beyond the purpose of this study to accept or not the hypothesis that fitness plays a protective role in people with higher level of fatness. Yet the discriminating role that anti-inflammatory and cholesterol levels seem to make sense when addressing fatness and fitness may point in the direction of “healthy obesity” when the CRF level is high [4].

4.5. Limitations

The current study is based on data from 81 individuals. Several variables, such as systolic and diastolic Blood Pressure and C-Reactive Protein, had to be excluded from the analysis because of missing data. The observations and conclusions drawn from this study would need to be verified in a larger dataset. This study does not provide direct experimental evidence, but is merely observational and retrospective. These considerations need to be taken into account when evaluating our results and conclusions. Our dataset has more females than males, and although sex did not appear to play a key role in determining the classification, we cannot exclude that, with a higher number of males this factorial variable would or could have had a greater weight. Finally, by dividing our dataset into four classes we observed that these were not evenly distributed. This issue was partially mitigated by balancing, using class weights.

5. Conclusions

Our data analytics approach has shown a potential key role of IL-10 as well as HDL, LDL, total Cholesterol and BMI in the classification of people according to their fatness and fitness levels. Unsupervised learning showed that a cluster of glucose tolerance related variables explains the highest quote of the variance of the categorical variable. However, supervised learning did not select this PCA cluster. Mediation analysis showed that LDL and BMI partially explain the association between fitness and fatness. These results suggest that CRF and FM% may be interconnected by anti-inflammatory responses and cholesterol blood levels. This may be in line with the protective role of cardiorespiratory fitness suggested in recent years. However, large randomized controlled trials are needed to validate this hypothesis experimentally and conclusively.

Supplementary Materials: The following are available online at <https://www.mdpi.com/1660-4601/18/4/1800/s1>, Figure S1: Boxplot of selected variables, and correlation matrices before and after the selection.

Author Contributions: Conceptualization, H.-P.K., J.P.M. and F.S.; methodology, F.S.; software, F.S.; formal analysis, F.S.; investigation, H.-P.K., J.P.M. and F.S.; resources, H.-P.K.; data curation, H.-P.K.; writing—original draft preparation, F.S.; writing—review and editing, H.-P.K., J.P.M. and F.S.; visualization, F.S.; supervision, H.-P.K.; project administration, H.-P.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Schools of Sport Health and Exercise Sciences Research Ethics Committee of Bangor University (protocol code M15-15/16 and approval date 16/05/2016).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to restrictions related to the data protection regulations.

Acknowledgments: The authors would like to thank Themistoklis Marinos for helping with collection of data.

Conflicts of Interest: F.S. works for Royal Philips. All other authors do not have conflict to disclose.

References

1. Morris, J.N.; Heady, J.A.; Raffle, P.A.B.; Roberts, C.G.; Parks, J.W. Coronary heart-disease and physical activity of work. *Lancet* **1953**, *262*, 1111–1120. [[CrossRef](#)]
2. Blair, S.N.; Kohl, H.W.; Paffenbarger, R.S.; Clark, D.G.; Cooper, K.H.; Gibbons, L.W. Physical Fitness and All-Cause Mortality: A Prospective Study of Healthy Men and Women. *JAMA J. Am. Med. Assoc.* **1989**, *262*, 2395–2401. [[CrossRef](#)]
3. Ortega, F.B.; Lavie, C.J.; Blair, S.N. Obesity and cardiovascular disease. *Circ. Res.* **2016**, *118*, 1752–1770. [[CrossRef](#)] [[PubMed](#)]
4. Yerrakalva, D.; Mullis, R.; Mant, J. The associations of “fatness,” “fitness,” and physical activity with all-cause mortality in older adults: A systematic review. *Obesity* **2015**, *23*, 1944–1956. [[CrossRef](#)] [[PubMed](#)]

5. Frith, E.; Loprinzi, P.D. The protective effects of a novel fitness-fatness index on all-cause mortality among adults with cardiovascular disease. *Clin. Cardiol.* **2017**. [[CrossRef](#)]
6. Lee, J. Influences of Cardiovascular Fitness and Body Fatness on the Risk of Metabolic Syndrome: A Systematic Review and Meta-Analysis. *Am. J. Health Promot.* **2020**. [[CrossRef](#)] [[PubMed](#)]
7. Oppewal, A.; Hilgenkamp, T.I.M. Is fatness or fitness key for survival in older adults with intellectual disabilities? *J. Appl. Res. Intellect. Disabil.* **2020**, *33*, 1016–1025. [[CrossRef](#)]
8. Ortega, F.B.; Ruiz, J.R.; Labayen, I.; Lavie, C.J.; Blair, S.N. The Fat but Fit paradox: What we know and don't know about it. *Br. J. Sport. Med.* **2018**, *52*. [[CrossRef](#)]
9. Duncan, G.E. The “fit but fat” concept revisited: Population-based estimates using NHANES. *Int. J. Behav. Nutr. Phys. Act.* **2010**, *7*, 47. [[CrossRef](#)]
10. Lahoz-García, N.; García-Hermoso, A.; Milla-Tobarra, M.; Díez-Fernández, A.; Soriano-Cano, A.; Martínez-Vizcaíno, V. Cardiorespiratory Fitness as a Mediator of the Influence of Diet on Obesity in Children. *Nutrients* **2018**, *10*, 358. [[CrossRef](#)]
11. Edwards, M.K.; Loprinzi, P.D. Physical activity and diet on atherogenic index of plasma among adults in the United States: Mediation considerations by central adiposity. *Eur. J. Clin. Nutr.* **2018**, *72*, 826–831. [[CrossRef](#)]
12. Lee, D.C.; Sui, X.; Church, T.S.; Lee, I.M.; Blair, S.N. Associations of cardiorespiratory fitness and obesity with risks of impaired fasting glucose and type 2 diabetes in men. *Diabetes Care* **2009**, *32*, 257–262. [[CrossRef](#)] [[PubMed](#)]
13. Clarke, S.L.; Reaven, G.M.; Leonard, D.; Barlow, C.E.; Haskell, W.L.; Willis, B.L.; DeFina, L.; Knowles, J.W.; Maron, D.J. Cardiorespiratory Fitness, Body Mass Index, and Markers of Insulin Resistance in Apparently Healthy Women and Men. *Am. J. Med.* **2020**, *133*, 825–830.e2. [[CrossRef](#)]
14. Leite, S.A.; Monk, A.M.; Upham, P.A.; Chacra, A.R.; Bergenstal, R.M. Low cardiorespiratory fitness in people at risk for type 2 diabetes: Early marker for insulin resistance. *Diabetol. Metab. Syndr.* **2009**, *1*, 8. [[CrossRef](#)]
15. Ferrante, A.W. Obesity-induced inflammation: A metabolic dialogue in the language of inflammation. *J. Intern. Med.* **2007**, *262*, 408–414. [[CrossRef](#)] [[PubMed](#)]
16. Coelho, A.J. Systemic Inflammation in the Morbidly Obese Patient. In *Gastric Bypass*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 125–132.
17. Ropelle, E.R.; Flores, M.B.; Cintra, D.E.; Rocha, G.Z.; Pauli, J.R.; Morari, J.; De Souza, C.T.; Moraes, J.C.; Prada, P.O.; Guadagnini, D.; et al. Leptin Sensitivity through IKK β and ER Stress Inhibition. *PLoS Biol.* **2010**, *8*, 1000465. [[CrossRef](#)]
18. Batista, M.L.; Rosa, J.C.; Lopes, R.D.; Lira, F.S.; Martins, E.; Yamashita, A.S.; Brum, P.C.; Lancha, A.H.; Lopes, A.C.; Seelaender, M. Exercise training changes IL-10/TNF- α ratio in the skeletal muscle of post-MI rats. *Cytokine* **2010**, *49*, 102–108. [[CrossRef](#)] [[PubMed](#)]
19. Nieman, D.C.; Henson, D.A.; Davis, J.M.; Dumke, C.L.; Utter, A.C.; Murphy, E.A.; Pearce, S.; Gojanovich, G.; McAnulty, S.R.; McAnulty, L.S. Blood Leukocyte mRNA Expression for IL-10, IL-1Ra, and IL-8, but Not IL-6, Increases after Exercise. *J. Interf. Cytokine Res.* **2006**, *26*, 668–674. [[CrossRef](#)]
20. Dorneles, G.P.; Haddad, D.O.; Fagundes, V.O.; Vargas, B.K.; Kloecker, A.; Romão, P.R.T.; Peres, A. High intensity interval exercise decreases IL-8 and enhances the immunomodulatory cytokine interleukin-10 in lean and overweight-obese individuals. *Cytokine* **2016**, *77*, 1–9. [[CrossRef](#)]
21. Sartor, F.; de Morree, H.M.; Matschke, V.; Marcora, S.M.; Milousis, A.; Thom, J.M.; Kubis, H.P. High-intensity exercise and carbohydrate-reduced energy-restricted diet in obese individuals. *Eur. J. Appl. Physiol.* **2010**. [[CrossRef](#)] [[PubMed](#)]
22. Barry, J.C.; Simtchouk, S.; Durrer, C.; Jung, M.E.; Mui, A.L.; Little, J.P. Short-term exercise training reduces anti-inflammatory action of interleukin-10 in adults with obesity. *Cytokine* **2018**, *111*, 460–469. [[CrossRef](#)]
23. Eposito, K.; Pontillo, A.; Giugliano, F.; Giugliano, G.; Nicoletti, G.; Giugliano, D. Association of Low Interleukin-10 Levels with the Metabolic Syndrome in Obese Women. *J. Clin. Endocrinol. Metab.* **2003**, *88*, 1055–1058. [[CrossRef](#)] [[PubMed](#)]
24. Jankord, R.; Jemiolo, B. Influence of physical activity on serum IL-6 and IL-10 levels in healthy older men. *Med. Sci. Sports Exerc.* **2004**, *36*, 960–964. [[CrossRef](#)] [[PubMed](#)]
25. Nunes, R.B.; Tonetto, M.; Machado, N.; Chazan, M.; Heck, T.G.; Veiga, A.B.G.; Dall'Ago, P. Physical exercise improves plasmatic levels of IL-10, left ventricular end-diastolic pressure, and muscle lipid peroxidation in chronic heart failure rats. *J. Appl. Physiol.* **2008**, *104*, 1641–1647. [[CrossRef](#)]
26. Goldhammer, E.; Tanchilevitch, A.; Maor, I.; Beniamini, Y.; Rosenschein, U.; Sagiv, M. Exercise training modulates cytokines activity in coronary heart disease patients. *Int. J. Cardiol.* **2005**, *100*, 93–99. [[CrossRef](#)]
27. Kolliputi, N.; Goyal, R.; Panguluri, S.K.; Mishra, P.K.; Kesharwani, V.; Chavali, V.; Hackfort, B.T.; Tyagi, S.C. Exercise ameliorates high fat diet induced cardiac dysfunction by increasing interleukin 10. *Front. Physiol.* **2015**, *6*, 124. [[CrossRef](#)]
28. Gallagher, D.; Heymsfield, S.B.; Heo, M.; Jebb, S.A.; Murgatroyd, P.R.; Sakamoto, Y. Healthy percentage body fat ranges: An approach for developing guidelines based on body mass index. *Am. J. Clin. Nutr.* **2000**, *72*, 694–701. [[CrossRef](#)] [[PubMed](#)]
29. Pescatello, L.S. *ACSM's Guidelines for Exercise Testing and Prescription*, 9th ed.; Wolters Kluwer/Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2014.
30. Zhang, Z.; Castelló, A. Principal components analysis in clinical studies. *Ann. Transl. Med.* **2017**, *5*, 12. [[CrossRef](#)]
31. Granitto, P.M.; Furlanello, C.; Biasioli, F.; Gasperi, F. Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. *Chemom. Intell. Lab. Syst.* **2006**, *83*, 83–90. [[CrossRef](#)]
32. Kursu, M.B.; Rudnicki, W.R. Feature selection with the boruta package. *J. Stat. Softw.* **2010**, *36*, 1–13. [[CrossRef](#)]

33. Baron, R.M.; Kenny, D.A. *The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations*; Psychological Association, Inc.: Washington, DC, USA, 1986; Volume 51.
34. Thompson, P.D.; Arena, R.; Riebe, D.; Pescatello, L.S. ACSM's new preparticipation health screening recommendations from ACSM's guidelines for exercise testing and prescription, ninth edition. *Curr. Sports Med. Rep.* **2013**, *12*, 215–217. [[CrossRef](#)]
35. Wallace, T.M.; Levy, J.C.; Matthews, D.R. Use and abuse of HOMA modeling. *Diabetes Care* **2004**, *27*, 1487–1495. [[CrossRef](#)]
36. Kuhn, M. Building predictive models in R using the caret package. *J. Stat. Softw.* **2008**, *28*, 1–26. [[CrossRef](#)]
37. Kuhn, M. The Caret Package. Available online: <http://topepo.github.io/caret/index.html> (accessed on 20 March 2020).
38. Chambers, J.M.; Cleveland, W.S.; Kleiner, B.; Tukey, P.A. *Graphical Methods for Data Analysis*; CRC Press: Boca Raton, FL, USA, 2018; ISBN 9781351080750.
39. Husson, F.; Le, S.; Pagès, J. *Exploratory Multivariate Analysis by Example Using R*; Chapman and Hall: New York, NY, USA; CRC: Boca Raton, FL, USA, 2017.
40. Boehmke, B.; Greenwell, B. *Hands-On Machine Learning with R*; Chapman and Hall: New York, NY, USA; CRC: Boca Raton, FL, USA, 2019.
41. Tingley, D.; Yamamoto, T.; Hirose, K.; Keele, L.; Imai, K. Mediation: R Package for Causal Mediation Analysis. *J. Stat. Softw.* **2014**, *59*, 1–38. [[CrossRef](#)]
42. Heinzen, E.; Sinnwell, J.; Atkinson, E.; Gunderson, T.; Dougherty, G.; Votruba, P.; Lennon, R.; Hanson, A.; Goergen, K.; Lundt, E.; et al. An Arsenal of “R” Functions for Large-Scale Statistical Summaries. Available online: <https://cran.r-project.org/web/packages/arsenal/arsenal.pdf> (accessed on 20 March 2020).
43. Solomon, T.P.J.; Malin, S.K.; Karstoft, K.; Knudsen, S.H.; Haus, J.M.; Laye, M.J.; Kirwan, J.P. Association Between Cardiorespiratory Fitness and the Determinants of Glycemic Control Across the Entire Glucose Tolerance Continuum. *Diabetes Care* **2015**. [[CrossRef](#)]
44. Gregory, J.M.; Muldowney, J.A.; Engelhardt, B.G.; Tyree, R.; Marks-Shulman, P.; Silver, H.J.; Donahue, E.P.; Edgerton, D.S.; Winnick, J.J. Aerobic exercise training improves hepatic and muscle insulin sensitivity, but reduces splanchnic glucose uptake in obese humans with type 2 diabetes. *Nutr. Diabetes* **2019**. [[CrossRef](#)] [[PubMed](#)]
45. Jung, S.H.; Park, H.S.; Kim, K.S.; Choi, W.H.; Ahn, C.W.; Kim, B.T.; Kim, S.M.; Lee, S.Y.; Ahn, S.M.; Kim, Y.K.; et al. Effect of weight loss on some serum cytokines in human obesity: Increase in IL-10 after weight loss. *J. Nutr. Biochem.* **2008**, *19*, 371–375. [[CrossRef](#)]
46. Vasankari, T.J.; Kujala, U.M.; Vasankari, T.M.; Ahotupa, M. Reduced oxidized LDL levels after a 10-month exercise program. *Med. Sci. Sports Exerc.* **1998**, *30*, 1496–1501. [[CrossRef](#)] [[PubMed](#)]
47. Serrano-Sá Nchez, J.A.; Delgado-Guerra, S.; Olmedillas, H.; Guadalupe-Grau, A.; Arteaga-Ortiz, R.; Sanchis-Moysi, J.; Dorado, C.; Calbet, J.A.L. Adiposity and Age Explain Most of the Association between Physical Activity and Fitness in Physically Active Men. *PLoS ONE* **2010**. [[CrossRef](#)]
48. Díez-Fernández, A.; Sánchez-López, M.; Gulías-González, R.; Notario-Pacheco, B.; García-Prieto, J.C.; Arias-Palencia, N.; Martínez-Vizcaíno, V. BMI as a mediator of the relationship between muscular fitness and cardiometabolic risk in children: A mediation analysis. *PLoS ONE* **2015**, *10*. [[CrossRef](#)] [[PubMed](#)]
49. Kwon, S.; Trudy, L.B.; Janz, K. Associations of cardiorespiratory fitness and fatness with cardiovascular risk factors among adolescents: The NHANES 1999–2002. *J. Phys. Act. Health* **2010**, *7*, 746–753. [[CrossRef](#)] [[PubMed](#)]
50. Tiainen, S.; Kiviniemi, A.; Hautala, A.; Huikuri, H.; Ukkola, O.; Tokola, K.; Tulppo, M.; Vasankari, T. Effects of a Two-Year Home-Based Exercise Training Program on Oxidized LDL and HDL Lipids in Coronary Artery Disease Patients with and without Type-2 Diabetes. *Antioxidants* **2018**, *7*, 144. [[CrossRef](#)] [[PubMed](#)]
51. Luquet, S.; Lopez-Soriano, J.; Holst, D.; Fredenrich, A.; Melki, J.; Rassoulzadegan, M.; Grimaldi, P.A. Peroxisome proliferator-activated receptor δ controls muscle development and oxydative capability. *FASEB J.* **2003**, *17*, 2299–2301. [[CrossRef](#)] [[PubMed](#)]
52. Spangenburg, E.E.; Brown, D.A.; Johnson, M.S.; Moore, R.L. Alterations in peroxisome proliferator-activated receptor mRNA expression in skeletal muscle after acute and repeated bouts of exercise. *Mol. Cell. Biochem.* **2009**, *332*, 225–231. [[CrossRef](#)]
53. Wei, C.; Penumetcha, M.; Santanam, N.; Liu, Y.G.; Garelnabi, M.; Parthasarathy, S. Exercise might favor reverse cholesterol transport and lipoprotein clearance: Potential mechanism for its anti-atherosclerotic effects. *Biochim. Biophys. Acta Gen. Subj.* **2005**, *1723*, 124–127. [[CrossRef](#)]
54. Butcher, L.R.; Thomas, A.; Backx, K.; Roberts, A.; Webb, R.; Morris, K. Low-intensity exercise exerts beneficial effects on plasma lipids via ppar γ . *Med. Sci. Sports Exerc.* **2008**, *40*, 1263–1270. [[CrossRef](#)] [[PubMed](#)]
55. Lucero, D.; Islam, P.; Freeman, L.A.; Jin, X.; Pryor, M.; Tang, J.; Kruth, H.S.; Remaley, A.T. Interleukin 10 promotes macrophage uptake of HDL and LDL by stimulating fluid-phase endocytosis. *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **2020**, *1865*, 158537. [[CrossRef](#)]