TITLE: THE FH PHENOTYPE: MONOGENIC FAMILIAL HYPERCHOLESTEROLAEMIA, POLYGENIC HYPERCHOLESTEROLAEMIA AND OTHER CAUSES

SHORT TITLE: FH PHENOTYPE: MONOGENIC, POLYGENIC AND OTHER CAUSES

AUTHORS

Mariano C, PhD[1,2]; Alves AC; PhD[1,2]; Medeiros A, MSc[1,2]; Chora JR, MSc[1,2]; Antunes M, PhD[3,4]; Futema M, PhD[5]; Humphries SE, PhD[5]; Bourbon M, PhD[1,2].

AFFILIATIONS

[1]Cardiovascular Research Group, Research and Development Unit, Department of Health Promotion and Chronic Diseases, National Institute of Health Doutor Ricardo Jorge, Lisbon, Portugal;

[2]University of Lisbon, Faculty of Sciences, BioISI – Biosystems & Integrative Sciences Institute, Lisbon, Portugal;

[3]Department of Statistics and Operations Research, Faculty of Sciences, University of Lisbon, Lisbon, Portugal;

[4]Centre of Statistics and its Applications – CEAUL, Faculty of Sciences, University of Lisbon, Lisbon, Portugal;

[5]Centre for Cardiovascular Genetics, Institute of Cardiovascular Sciences, University College London, London, United Kingdom.

CORRESPONDING AUTHOR

Mafalda Bourbon

Instituto Nacional de Saúde Doutor Ricardo Jorge

Av. Padre Cruz, 1649-016 Lisboa, Portugal

Tel.: (+351) 217 519 200 | (+351) 217 508 126

E-Mail: mafalda.bourbon@insa.min-saude.pt

ABSTRACT (250)

Background : Familial Hypercholesterolaemia (FH) is a monogenic disorder characterised by high LDL-C values and increased cardiovascular risk. The correct identification of the origin of the dyslipidaemia is important for patient management, including the implementation of the best therapeutic measures; patients with *LIPA* and *ABCG5/8* mutations need a completely different management approach than patients with defects in *LDLR/APOB/PCSK9*. The aim of this work was to characterise the genetic cause of the FH phenotype in a cohort of Portuguese patients with a clinical diagnosis of FH.

Methods: Between 1999 and 2007, 731 index patients (311 children and 420 adults) have been referred to our laboratory who met the Simon Broome diagnostic criteria . The *LDLR*, *APOB*, *PCSK9*, *APOE*, *LIPA*, *LDLRAP1*, *ABCG5*/8 genes were analysed using standard Next Generation sequencing methods. Polygenic hypercholesterolaemia was determined when an individual had a 6-SNP LDL-C genetic risk score (GRS) over the score 25th percentile.

Results : An FH-causing mutation in *LDLR*, *APOB* or *PCSK9* was found in 39% of patients (92% in *LDLR*, 5% *APOB* and 1% *PCSK9*), 14% have polygenic hypercholesterolaemia and 1% have other lipid disorders. In 55% of the patients in the paediatric group and in 51% of patients in the adult group, a cause for their hypercholesterolaemia was identified. Compared to a group of 1563 healthy Portuguese controls, in 7% of patients, a low LDL-C GRS was observed, suggesting they are likely to have another cause of monogenic dyslipidaemia.

Conclusions :-In the genetically heterogeneous Portuguese population a genetic cause for the FH phenotype can be found in 51%-55% of patients All known causes of the FH phenotype should be investigated in FH cohorts to ensure accurate diagnosis and appropriate management.

Keywords: Familial hypercholesterolaemia; monogenic dyslipidaemia; polygenic hypercholesterolaemia; genetic risk score; phenocopies.

1. INTRODUCTION

Familial Hypercholesterolaemia (FH) is an autosomal dominant condition characterised by substantially increased plasma concentrations of low-density lipoprotein cholesterol (LDL-C) from birth, leading to premature atherosclerosis. FH is also one of the most common inherited disorders associated to premature coronary heart disease (pCHD), with a frequency around 1:250 in most populations[1].

The genetic causes of FH are loss-of-function mutations, mainly in the LDL receptor gene (*LDLR*)[2] or apolipoprotein B gene (*APOB*)[3–5], and gain-of-function mutations in the proprotein convertase subtilisin/kesin type 9 gene (*PCSK9*)[6]. However, an increasing number of FH phenocopies are being identified and a few individuals with a clinical diagnosis of FH have been found to have rare variants in other genes, such as apolipoprotein E (*APOE*)[7,8], ATP-binding cassette sub-family G member 5 or 8 (*ABCG5/8)*[9]or lysosomal acid lipase (*LIPA*)[10,11]. Recently, studies have reported that a significant proportion of clinically-diagnosed FH patients where a mutation causing disease was not found could have a polygenic cause for their hypercholesterolaemia due to the inheritance of common LDL-C raising single nucleotide polymorphisms (SNPs) with a cumulative effect, leading to an increase in LDL-C to the level of FH diagnostic criteria[12,13]. Even in patients with a disease-causing mutation, this polygenic contribution is also present, contributing to a variable FH phenotype in patients with the same FH-causing mutation[13]. Several combinations of SNPs have been suggested to estimate the polygenic contribution[10,12,13].

In terms of cardiovascular risk assessment, it is of particular clinical value to distinguish between monogenic and other types of dyslipidaemia, since for all LDL intervals the cardiovascular disease (CVD) risk has been demonstrated to be higher in FH patients with a causative mutation compared to patients with the same LDL values[14]. This highlights the concept of time of LDL exposure, or LDL-C "Burden". Because of this, monogenic FH patients warrant treatment with high intensity and effective lipid lowering therapy ~~aggressively~~ to decrease their CVD risk, while those with a polygenic or environmental dyslipidaemia have a lower CVD risk and can be managed, for example by encouraging lifestyle and dietary changes, and with the use of more moderate dosage of lipid-lowering therapies.

In clinically-defined FH cohorts worldwide, an FH-causing variant is found in about 40-50% of the cases, although the prevalence of genetically identified FH patients will vary due to differences in molecular diagnostic methodologies, and also to differences in the clinical criteria applied[1,4,15–17]. In the remaining 50% of the cases, the cause for the hypercholesterolaemia must be sought for better patient management and prognosis.

Taking all these aspects in consideration, it is clearly important to ensure the correct identification of the aetiology of the dyslipidaemia in a subject, in order to implement appropriate interventions for CVD prevention. Here, we report the characterisation of the FH phenotype in the

Portuguese FH Study cohort. We also validated the 6-SNP LDL-C genetic risk score (GRS)[12,13] in the Portuguese population.

2. MATERIAL AND METHODS

The Portuguese FH Study is a research project coordinated by the National Institute of Health (INSA) supported mainly by external funds and free of charge for all patients and health institutions. INSA Ethical Committee and the National Data Protection Commission previously approved the study protocol and database. Written informed consent was obtained from all participants before their inclusion in the study.

2.1. Study population

A total of 887 index patients were enrolled in the Portuguese FH Study from 1999 to 2017, referred from different clinical specialties with a clinical suspicion of FH. In the present report, we only included the 731 index patients (311 children and 420 adults) with a clinical diagnosis of FH according to the Simon Broome (SB) criteria, as previously described[18], with a single adaptation - individuals aged 16-18 were included with the SB criteria for children due to the mild phenotype seen in FH patients within this age range. Additionally, 1777 relatives (393 children and 1384 adults) were referred to the Portuguese FH Study cascade-screening program (with or without a clinical diagnosis of FH).

2.2. Monogenic dyslipidaemia analysis

Genetic diagnosis was performed by the molecular analysis of *LDLR* (including the study of splice regions and large rearrangements), *APOB* (two fragments of exons 26 and 29), and *PCSK9* genes, as previously reported[19]. Selected patients, where a variant was not found in the previously studied genes, were further investigated for other monogenic causes of dyslipidaemia; this was performed by polymerase chain reaction and Sanger sequencing of the following genes: *APOE*, *LIPA*, LDLR adapter protein 1 (*LDLRAP1*), *ABCG5* and *ABCG8*. Sequences were analysed with Staden software[11,20] and the references used for analysis were NM_000527 for *LDLR*, NM_000384 for *APOB*, NM_174936 for *PCSK9*, NM_000041 for *APOE*, NM_015627 for *LDLRAP1*, NM_022436 for *ABCG5*, NM_022437 for *ABCG8* and NM_000235 for *LIPA*. Complementary DNA numbering was considered according to the Human Genome Variation Society (HGVS) nomenclature[21] with nucleotide c.1 being A of the ATG initiation codon p.1. The molecular study of the albumin gene (*ALB*) was performed elsewhere[22].

All variants were checked with Mutalyzer 2.0, as recommended by Human Genome Variation Society (HGVS). Variants were classified as pathogenic, likely pathogenic, variant of unknown significance (VUS), likely benign or benign, according to the American College of Medical Genetics and Genomics (ACMG) guidelines[23] following specific adaptations described in Chora *et*

al.[23] The variants reported in the present study were considered novel if they were not described before, and novel for Portugal if they were found for the first time in Portugal, but have been previously reported in another country. *In silico* analysis was performed as described before[24].

2.3. Polygenic hypercholesterolaemia analysis

A total of 1,563 genomic DNA samples from the e_COR Study and 455 index cases from the Portuguese FH Study from which quality DNA was available (186 children and 269 adults) were sent to aScidea Computational Biology Solutions Company (Barcelona, Spain) to be genotyped for a set of 6 SNPs, using the OpenArray™ technology[25] (Life Technologies, Carlsbad, California, US).

The LDL-C GRS was calculated using the 6 SNPs previously reported in the characterisation of polygenic hypercholesterolaemia, namely, cadherin EGF LAG seven-pass G-type receptor 2 (*CELSR2)/* sortilin 1 (*SORT1*) (rs629301), *APOB* (rs1367117), *ABCG5/8* (rs4299376), *LDLR* (rs6511720) and *APOE* (rs7412 and rs429358) and respective effect sizes (weighted sum of beta-coefficients of the risk allele)[13] (Supplementary Table 1).

The e_COR Study population[26] was used as reference group for the validation of the LDL-C GRS in the Portuguese population. LDL-C scores were distributed into quarters; individuals below the 25th percentile (P25th) were considered as having low polygenic score, between the P25th and the 75th percentile (P75th), intermediate polygenic score, and above the P75th, high polygenic score. This score was applied to the Portuguese FH Study population accordingly.

2.4. Biochemical characterisation of lipids and lipoproteins

For both cohorts, The Portuguese FH Study and e_COR study, the biochemical tests for Total Cholesterol (TC), direct LDL-C, high-density lipoprotein cholesterol (HDL-C), triglycerides (TG), Apolipoprotein A1 (apoA1), and Apolipoprotein B (apoB) were performed by enzymatic colorimetric and immunoturbidimetric methods. Serum levels of lipoprotein (a) [Lp(a)] were determined by an immunoturbidimetric method, as previously described[27].

2.5. Correction factors regarding lipid-lowering therapy

Whenever untreated lipid values (TC, LDL-C and apoB) for individuals under statins medication were not available, these were estimated using correction factors: measured TC and LDL-C[14,28,29], as well as apoB[30], was divided by 0.8 (20% TC reduction), 0.7 (30% LDL-C reduction), and 0.763 (23.7% apoB reduction), respectively. Untreated TG, HDL-C and apoA1 values were not estimated, since the effects of lipid-lowering therapy with statins are not significant in these biomarkers[31–34].

2.6. General statistical analysis

Statistical analyses were performed using R (version 3.1.2) software[35]. For comparison analysis of lipids, lipoproteins and LDL-C GRS values between independent groups, the non-parametric Two-sample Wilcoxon or Kruskal-Wallis tests were applied for two or more independent samples, respectively. When the trait distributions met the assumptions of normality (Shapiro-Wilk or Kolmogorov-Smirnov tests) and homogeneity of variance (Bartlett test), the parametric ANOVA or Student t tests were applied for two or more independent samples. Data are shown as Mean (±SD) For comparison of proportions, the 95% confidence intervals (CI) were used. Whenever the two CI did not overlap, it was considered that there was evidence to conclude that the proportions are statistically different. In the remaining cases (overlap of the two proportions confidence intervals), the two proportions were compared using chi-square or Fisher's tests. The multiple of median (MoM) was calculated for the LDL-C, TG and apoB measured values to analyse how far those values deviate from the median of a reference population. For phenotype *versus* genotype analysis the biochemical values at referral were used (corrected as described when necessary).

3. RESULTS

3.1. Demographic and clinical data

Demographic and clinical data on cardiovascular disease risk factors of all index cases are shown in Supplementary Table 2, including the non-treated lipid profile (when available), and the complete fasting lipid profile performed at our Institute for all individuals at referral to the Portuguese FH Study. Mean age (years) at referral was 9.94 (SD 3.69) for children and 45.67 (SD 13.32) for adults. Approximately 20% of the children were under pharmacological treatment at referral compared with 75% of the adults. The majority (>95%) of the patients are of Portuguese nationality distributed within all Portuguese regions.

3.2. Monogenic dyslipidaemia

A total of 731 index cases was analysed as described in methods for *LDLR*, *APOB* and *PCSK9* genes. In 282 patients (128 children and 154 adults) a pathogenic or likely pathogenic variant was found, including 9 homozygous (2 children and 7 adults); these will be referred as FH mutation positive (FH/M+). In 398 patients (159 children and 239 adults) no potential pathogenic variants were found; these will be referred as FH mutation negative (FH/M-). Additionally, 18 children and 26 adults were found to have a VUS.

Based on phenotype, selected FH/M- patients were analysed for other possible causes of monogenic dyslipidaemia.

3.2.1. Familial Hypercholesterolaemia

In about 39% of all index patients (n=282/731), at least one pathogenic or likely pathogenic variant was identified in the *LDLR*, *APOB* or *PCSK9* genes. In the paediatric index cohort (n=311),

41% individuals had genetically heterozygous FH (HeFH) (n=126) or homozygous FH (HoFH) (n=2). In the adult index cohort (n=420), 37% individuals had genetically HeFH (n=147) or HoFH (n=7). A VUS in the *LDLR* and *APOB* genes were found in 6% of children (n=18) and 6% of adults (n=26) corresponding to 35 individuals with a VUS in the *LDLR* and 9 individuals in the *APOB* gene. The cascade-screening programme led to the additional identification of 116 HeFH children, 314 HeFH adults and 1 HoFH adult. Additionally, 38 relatives (7 children and 31 adults), also had a VUS.

Since our last report in 2015[36], 8 *LDLR* novel variants have been identified in our cohort (5 never described before and 3 described in other countries, but novel for Portugal). All variants have already been submitted to ClinVar (Supplementary Table 3). From these, only 3 variants are considered pathogenic or likely pathogenic: c.2214del/p.(Gln739Serfs*26), c.941-2A>C, and c.1897C>T/p.(Arg633Cys).

The demographic, clinical and biochemical profile of the FH/M+ group were compared to the FH/M- group, for children and adults, and is presented in Table 1. Patients with a VUS, homozygous patients and patients with other monogenic causes (discussed in the next section) were not included in this analysis. Although all FH/M- patients have a clinical phenotype compatible with a diagnosis of FH, they present lower mean levels of TC, LDL-C, non-HDL-C, apoB, and apoB/apoA1 ratio, and higher levels of HDL-C and TG, than the FH/M+ group. These differences are more evident in the paediatric cohort. Also, the percentage of adult FH/M- patients with hypertriglyceridaemia (defined as TG ≥200 mg/dL) was higher than in the FH/M+ patients (20.00% (n=47/235) *versus* 8.33% (n=12/44), P=0.002). The opposite was observed for apoB, a higher number of FH/M+ patients (87.85% (n=123/140) had apoB ≥120 mg/dL compared to FH/M- (64.94% (n=150/231), P<0.001). No significant differences were observed for Lp(a) values.

The lipid values of the FH/M+ index and FH/M+ relatives presented significant differences: TC, LDL-C, apoB, non-HDL-C and apoB/apoA1 ratio were statistically higher in index cases, except in the paediatric cohort where the apoB values and apoB/apoA1 ratio did not differ between these two groups (Supplementary Table 4).

Table 1– Clinical, demographic and biochemical profile (Mean ±SD) of children and adult index cases in the Portuguese Familial Hypercholesterolaemia Study with (FH/M+) and without (FH/M-) an identified mutation:

	Paediatric cohort (n=285)[a]			Adult cohort (n=386)[a]		
Clinical and demographic profile	**FH/M-**	**FH/M+**	P value[c]	**FH/M-**	**FH/M+**	P value[c]
n (%)	159 (51.13)	126 (40.51)	ND	239 (56.90)	147 (35.00)	ND
Age, years (SD)	48)	10.03 (3.86)	0.575	47.20 (11.91)	43.73 (14.91)	**0.008**
Male gender, n (%)	63 (39.62)	63 (50.00)	0.103	118 (49.37)	57 (38.78)	0.054
BMI, kg/m^2 (SD)	**20.27 (4.29)**	**18.90 (3.51)**	**0.012**	26.14 (4.21)	26.10 (4.86)	0.434
Smoking, n (%)	0	1 (0.79)	ND	**56 (23.43)**	**21 (14.29)**	**0.025**
Alcohol consumption, n (%)	0	0	ND	86 (35.98)	40 (27.21)	0.125
Hypertension, n (%)	3 (1.89)	0	ND	61 (25.52)	41 (27.89)	0.572
Diabetes, n (%)	1 (0.63)	0	ND	9 (3.77)	5 (3.40)	ND
Personal history of CVD, n (%)	0	0	ND	51 (21.34)	31 (21.09)	0.999
Personal history of pCVD, n (%)	0	0	ND	41 (17.15)	20 (13.61)	0.433
Family history of pCVD, n (%)	26 (16.35)	24 (19.05)	0.662	80 (33.47)	43 (29.25)	0.452
Pharmacological treatment (e.g. statins), n (%)	**18 (11.32)**	**36 (28.57)**	**<0.001**	183 (76.57)	109 (74.15)	0.678
On diet, n (%)	**86 (54.09)**	**50 (39.68)**	**0.021**	30 (12.55)	14 (9.52)	0.457
Physically active, n (%)	119 (74.84)	86 (68.25)	0.323	81 (33.89)	52 (35.37)	0.696
Tendon xanthoma, n (%)	0	0	ND	0	10 (6.80)	ND
Corneal arcus, n (%)	0	0	ND	1 (0.42)	13 (8.84)	ND
Other xanthomas, n (%)	0	1 (0.79)	ND	13 (5.44)	16 (10.88)	ND
Lipid profile[b]						
Total cholesterol, mg/dL (IQR)	**239.21 (211-259)**	**273.38 (245-302)**	**<<0.001**	**285.00 (242-321)**	**331.25 (284-395)**	**<<0.001**
LDL-C, mg/dL (IQR)	**156.88 (135-178)**	**203.64 (5=179-238)**	**<<0.001**	**201.71 (158-246)**	**263.57 (207-318)**	**<<0.001**
HDL-C, mg/dL (IQR)	**57.19 (49-68)**	**50.98 (42-58)**	**<0.001**	55.00 (44-64)	51.00 (44-62)	0.517
Triglycerides, mg/dL (IQR)	**84.50 (65-119)**	**65.00 (54 -88)**	**<0.001**	**129.00 (100-178)**	**112.39 (83-148)**	**<0.001**
ApoB, mg/dL IQR)	**99.50 (85-122)**	**129.50 (113-151)**	**<<0.001**	**136.84 (136-179)**	**177.63 (141-211)**	**<<0.001**
ApoA1, mg/dL (IQR)	**150.00 (136-174)**	**135.00 (115-147)**	**<<0.001**	**155.65 (136-179)**	**146.50 (127-167)**	**0.005**
Non-HDL-C, mg/dL (IQR)	**172.00 (153-194)**	**216.06 (192-253)**	**<<0.001**	**222.13 (182-270)**	**274.88 (224-328)**	**<<0.001**
ApoB/ApoA1 ratio, mg/dL (IQR)	**0.66 (00.52-0.82)**	**0.99 (0.85-1.18)**	**<<0.001**	**0.88(0.68-1.17)**	**1.16 (0.92-1.56)**	**<<0.001**
Lp(a), mg/dL (IQR)	32.00 (10-80)	24.64 (10-44)	0.081	30.00 (9-72)	35.15 (9-92)	0.339

FH, Familial Hypercholesterolaemia; FH/M-, FH mutation negative; FH/M+, FH FH/M+; BMI, body mass index; CVD, cardiovascular disease; pCVD, premature cardiovascular disease; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; ApoB, apolipoprotein B; ApoA1, apolipoprotein A1; Lp(a), lipoprotein(a); SD, standard deviation; IQR, interquartile range; ND, Not determined.
Age and BMI are expressed as mean (SD), while biochemical lipid values are expressed as median (IQR).
[a]Only heterozygous patients with pathogenic or likely-pathogenic variants were considered as FH/M+. Homozygous FH and patients with variants of uncertain significance (VUS) were excluded.
[b]Lipid profile determined in house. Correction factors were used to estimate untreated values for TC, LDL-C, and apoB, as described in the Material and Methods section of our study.
[c]Statistical significance (P value <0.05), in bold.

3.2.2. Other monogenic disorders

From analysis in XXXX ~~selected~~ FH/M- patients, disease-causing variants in the *APOE*, *LIPA*, *ABCG8* and *ALB* genes were found in 6 index patients, five children and one adult. Three children had lysosomal acid lipase deficiency (LALD) (due to homozygosity for the c.894G>A variant in *LIPA*), as previously published[11], one had sitosterolemia (due to homozygosity for a c.1974C>G variant in *ABCG8*) and one had congenital analbuminemia (due to heterozygosity for a c.1289+1G>A variant in *ALB*)[22]. An additional child had an heterozygous stop variant in *APOE* (c.683G>A) that has been associated only in the homozygote state to dysbetalipoproteinemia[38], so it was not considered as causative (Supplementary table 3). In the adult cohort, one individual had a variant in *APOE* (c.487C>T found in heterozygosity), previously associated with autosomal dominant hypercholesterolaemia[37] (Supplementary table 3). Additionally, one relative (child) was also found to have LALD and another sitosterolaemia (adult)[11].

3.2.3. Overall monogenic dyslipidaemia

A molecular cause of monogenic dyslipidaemia was found in 39% (n=288/731) of all index cases with an FH phenotype. In the monogenic dyslipidaemia paediatric group, 91% had an FH-causing variant in *LDLR*, 5% in *APOB*, 1% in *PCSK9*, 2% in *LIPA*, 1% in *ALB* and 1% in *ABCG8*. In the adult group, 93% had an FH-causing variant in *LDLR*, 5% in *APOB*, 1% in *PCSK9* and 1% in *APOE*. Altogether, other monogenic causes represent 2% (6/288) of monogenic hypercholesterolaemia index cases in our cohort, with a non-significantly (p=0.32) higher proportion in children (4%, 5/133) (Figure 1).

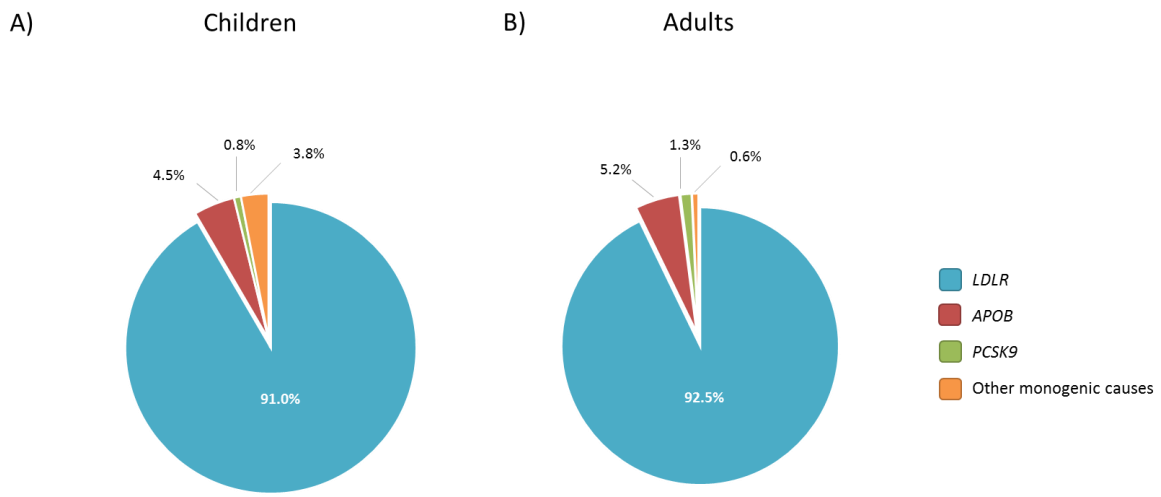


Figure 1 – Percentage of the different monogenic causes of dyslipidaemia in children (**A**) and adults (**B**). (**A**) 91% of children had an FH-causing variant in *LDLR*, 5% in *APOB*, 1% in *PCSK9*, and 4% of all children had other monogenic causes. (**B**) 93% of adults had an FH-causing variant in *LDLR*, 5% in *APOB*, 1% in *PCSK9* and 1% of adults had other monogenic causes.

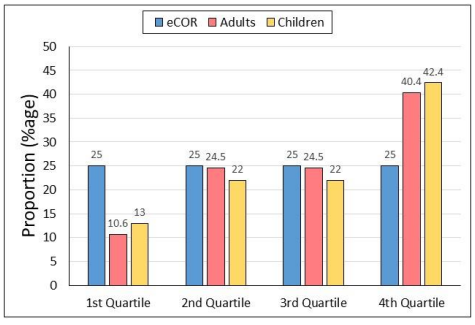
3.3. Polygenic hypercholesterolaemia

The mean value of the LDL-C GRS calculated in e_COR population was 0.62 mmol/L (SD 0.22) with a mean LDL-C of 135.75 mg/dL (SD 46.34) [3.51 mmol/L (SD 1.20)]. Distribution of the LDL-C concentration values by percentiles showed that individuals above the P75th (LDL-C GRS 0.76) had significantly higher LDL-C (P<<0.001) than individuals below the 25th percentile (P25th) (LDL-C GRS 0.51) (Figure 2). When comparing the mean GRS values between Portuguese e_COR and UK Whitehall II (WHII) controls[12], no significant differences were seen (0.62 (SD 0.22) CI=[0.61-0.63] *versus* 0.63 (SD 0.22) CI=[0.62-0.64], respectively).

Of all 731 clinical FH index cases, the LDL-C GRS was calculated for 455 individuals from whom DNA was available, 186 children and 269 adults. Compared with e_COR controls, both FH/M- and FH/M+ patients had higher LDL-C GRS (0.XX +P<<0.001) (Supplementary Figures 1-2). In this small sample, no statistically significant differences were seen comparing FH/M- and FH/M+ patients. In the paediatric cohort, the mean value was 0.73 (SD 0.17) for FH/M- (n= XX) and 0.71 (SD 0.19) for FH/M+ (n=YY), and in the adult cohort was 0.72 (SD 0.19) for FH/M- (n= AA) and 0.69 (SD 0.20) for FH/M+ (n=BB).

In the paediatric FH/M- cohort, about 42.4% (n=39) of patients are above the P75th, 13.0% (n=12) were below the P25th (P<0.001) and 44.6% (n=41) had an intermediate LDL-C GRS. In the adult FH/M- cohort, 40.4% (n=61) are above the P75th (GRS>0.76), 10.6% (n=16) are below the P25th (GRS≤0.51) and 49.0% (n=74) had an intermediate LDL-C GRS (between P25th and P75th).

Distribution of general population, and FH/M- Adults and children by GRS



If you give me the figures for the intermediate quartiles I can change this
IF YOU GIVE ME ACTUAL NUMBERS I CAN DO A CHI[2]

~~For FH/M+ paediatric patients, 39.5% (n=30) are above the P75th and 15.8% (n=12) are below the P25th. For FH/M+ adult patients, 34.1% (n=31) are above the P75th and 19.8% (n=18) are below the P25th.~~

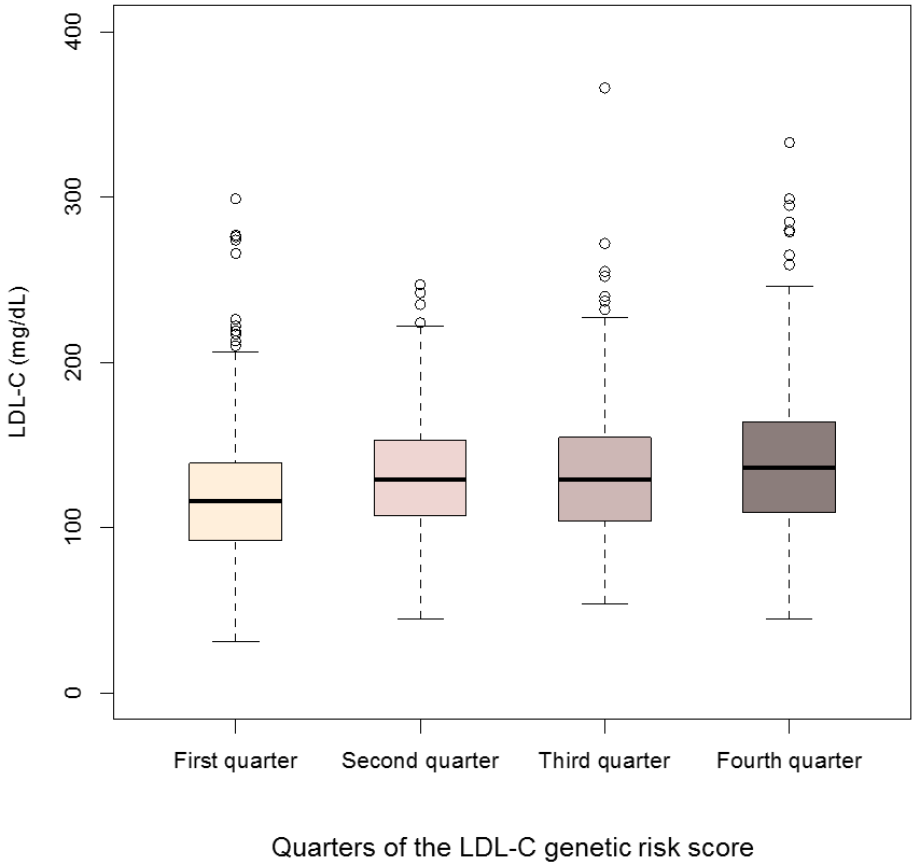


Figure 2 – Boxplot of the LDL-C values distribution by e_COR population in the first (≤25th percentile), second (>25th≤50th percentile), third (>50th≤75th percentile), and fourth (>75th percentile) quarters of the LDL-C genetic risk score (GRS). Correction factors were used to estimate untreated values for LDL-C (a 30% of LDL-C reduction was considered for patients under treatment with statins). Individuals in the fourth quarter of the LDL-C GRS distribution had significantly higher LDL-C concentration than individuals in the first quarter (P<<0.001).

3.4. The FH phenotype OR Overall proportion with an identified genetic cause for the FH phenotype

Overall, including both patients with homozygous (n=9) and heterozygous FH (n=273) we have identified a monogenic cause for the FH phenotype in XXX% of patients with the monogenic disorders lysosomal acid lipase deficiency (LALD), sitosterolaemia, albuminuria and autosomal dominant hypercholesterolaemia, representing 2% of all patients.

For the FH/M- patients for whom the LDL-C GRS was determined (92 children and 151 adults), 39 children and 61 adults had a polygenic score (above P75th), which represents 14% of overall patients in the study (n=731), while an additional X had an intermediate score. Both of these patients can bee included in the polygenic hypercholesterolaemia group (Figure 3). Of the remaining FH/M- patients

28 (12 children and 16 adults) presented a score below P25th and are therefore most likely to have other known or unknown cause of monogenic dyslipidaemia.

Studying all known causes of inherited hypercholesterolaemia increased the identification rate from 41% to 55% in the paediatric cohort and from 37% to 51% in the adult cohort (Figure 3).

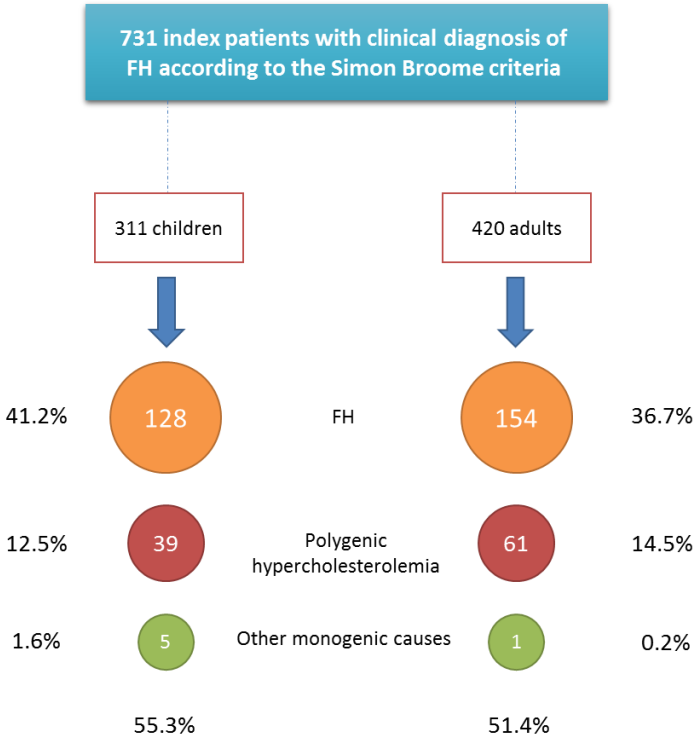


Figure 3 – Number of index cases, children and adults, referred to the Portuguese Familial Hypercholesterolaemia (FH) Study with Simon Broome FH clinical criteria divided by the different causes of the FH phenotype and percentages of identification rate by group and total. FH refers to patients with pathogenic variants in either *LDLR*, *APOB* or *PCSK9*, polygenic hypercholesterolaemia to patients with a high LDL-C genetic risk score (GRS) (> P75th) (0.76). Other monogenic causes are: in children, 3 with lysosomal acid lipase deficiency (mutation in the *LIPA* gene), 1 with sitosterolaemia (mutation in the *ABCG8* gene) and 1 with analbuminemia (mutation in *ALB* gene); in adults, 1 with an autosomal dominant hypercholesterolaemia (mutation in the *APOE* gene).

4. DISCUSSION

The Portuguese FH Study

Within index cases referred to the Portuguese FH Study, in 53% (n=388) we have identified the cause of hypercholesterolaemia: 39.4% (n=288) with monogenic hypercholesterolaemia (38.6% (n=282) have FH and 0.8% (n=6) other monogenic causes) and 14% (n=100) with polygenic hypercholesterolaemia. If all VUS are indeed pathogenic (found in 6% (n=44)), the total number of positives cases with a monogenic disorder will increase to 45% and the total identification rate to 59%.

From all individuals included in the Portuguese FH Study, we have identified 803 patients with disease-causing variants in *LDLR*, *APOB* or *PCSK9* genes: 793 heterozygous patients and 10 homozygous patients (9 index cases and 1 relative). Additionally, 82 patients with VUS in one of the three genes associated with FH were also identified. These variants will need further characterization to be considered cause of disease. From 2014 to 2017, a total of 10 novel variants were identified in our cohort, 5 never described before in association with FH, showing that although our population already has more than 140 different, new variants are still found as reported for other populations.

At the Portuguese FH Study only after functional characterisation is performed, or if a variant reaches a classification of pathogenic or likely pathogenic (ACMG), the clinician is informed that the variant found confirms or is consistent with the clinical diagnosis. A total of 52 variants have been already characterised by our group, leading to a correct FH diagnosis in more than 80% of our cohort of FH/M+ patients, improving this way the genetic diagnosis of FH. In other European countries these numbers are smaller since the majority did not implemented the functional studies in their workflow[23]. This should be encouraged worldwide.

Based on the more conservative estimation of the frequency of HeFH in Europe, 1:500, up to date our study identified 4% of the estimated HeFH patients to exist in the Portuguese population and 100% of all homozygous patients (1: 1 000 000). These numbers although small for heterozygous FH identification, place Portugal in the top 10 of countries with a higher percentage of patients identified[1].

Correctly identifying the cause of dyslipidaemia

All over the world in 50-60% of all individuals with the FH phenotype the cause of their hypercholesterolaemia could not be identified. Talmud et al raised the possibility that the majority of those where no mutation could be found in LDLR/APOB/PCSK9 were likely to have a polygenic aetiology while Wang and colleagues[39] suggested that in some cases, the FH phenotype could be due to variants in other genes yet to be described, other genes of lipid metabolism, interactions between known genes, variants inaccessible by the currently sequencing techniques, epigenetics or even environmental factors *per se*. Taking this into consideration, our group started to analyse other monogenic andr polygenic causes of hypercholesterolaemia in patients with a clinical diagnosis of

FH, but without an identified FH causing variant. This has increased the patient identification rate from 38.6% to 53%,

In the great majority of the index patients where a monogenic cause is found (36%) the cause of the dyslipidaemia is explained by a functional *LDLR* variant, 1.9% by a functional *APOB* variant and 0.4% by a *PCSK9* functional variant. Additionally, in ~60% a polygenic dyslipidemia is the probable cause and in 0.7% other monogenic disorders are the cause of the phenotype. These other causes are more prevalent in children representing 1.6% (n=6) of all cases. It is worth noting that other causes are more prevalent than *PCSK9* mutations (0.4%), reinforcing the need to study these FH phenocopy genes for a more accurate patient diagnosis and management. There are considerable differences in the treatment of these several phenocopies, which if identified, can lead to a better patient prognosis.

Our results are also consistent with previous studies showing that FH negative patients have higher mean LDL-C GRS than individuals from the general population, meaning that their LDL-C plasma levels is most likely to be due to the influence of a combination of several LDL-C variants, each with modest effect. A limitation of our study is that not all FH/M- patients were genotyped for the polygenic score (only 243 in 398 were analysed) due to DNA constraints, so the proportion of patients with polygenic hypercholesterolaemia could only be estimated. Interestingly, a similar proportion of patients above the P75th was found in the FH/M+ and the FH/M- group, a finding which was reported previously (Talmud et al 2013). This suggests that the FH phenotype in FH/M+ patients could be modulated by the modest effect of these LDL-C raising variants, at least at some level, however we could not see a statistical difference in LDL-C values, partly because of the selection criteria and partly due to the small sample size~~. Also possible is that in FH/M+ patients, where a single causative mutation has been already identified, the great mutation effect on LDL-C values could overlap the effect of several LDL-C raising variants with modest effect not allowing to detect differences in phenotype.~~

An unknown dyslipidaemia in the FH mutation negative patients

Of the 398 FH/M- patients with no identified cause for their hypercholesterolaemia, 7% (n=28) presented a low LDL-C GRS (below P25th) and are likely to have other known or unknown cause of monogenic dyslipidaemia. It is likely that in a small fraction of these patients, a new gene causing FH is yet to be discovered, although other possibilities should be considered, such as interactions between known genes or epigenetics. However, and since the mean triglyceride levels are statistically higher in FH/M- patients, in these cases environmental factors *per se* could be the cause of the phenotype [39] or these patients could have another dyslipidaemia related to triglycerides metabolism and not FH.

Importance of distinguishing the different causes of dyslipidaemia

Monogenic dyslipidaemias present a severe phenotype and are associated with an elevated CVD risk *per se*, like FH; while mild to severe dyslipidaemias are mostly due to polygenic hypercholesterolaemia, as a result of various genetic alterations that may interact, being modulated by non-genetic factors as life style (REF?). The distinction between these two types of dyslipidaemia is important for patient cardiovascular risk assessment and therapeutic management. It has been shown that FH patients with a pathogenic variant have 16 times greater cardiovascular risk compared to another individual with the same LDL value[14], but this risk can be reduced if FH patients are identified early in life and treated accordingly. This shows the importance of correctly identifying the cause of dyslipidaemia in early age and to address other cardiovascular risk factors in childhood, to reduce CVD rates later in adulthood. Different dyslipidaemias have different management approaches that can improve patient prognosis.

In this context, it is reasonable to say that the genetic diagnosis of clinical FH patients could benefit from the inclusion of all the genes studied in this work and the LDL-C GRS in a next-generation target panel, without a great increase in cost. Such a panel with the three proven FH-causing genes, *LDLR*, *APOB*, and *PCSK9*, and also the five known phenocopies, *LIPA*, *APOE* and *LDLRAP1* genes, *ABCG5*, *ABCG8* plus the 6-SNP LDL-C GRS is already implemented in our lab. This has also been already recommended in the last consensus paper authored by international FH experts[43].

6. Study Limitations

For some comparison analysis in the polygenic dyslipidaemia, the small sample size could imply bias. Although it is considered as a limitation of our study, our results are in line with the previously reported[12,13]. Another limitation of our study is that we have used the adult controls values for the paediatric cohort comparison. To confirm our results, comparison should be made against child controls. Also, the reduction in TC, LDL-C and apoB that we accounted for in those undergoing lipid-lowering therapy might imperfectly estimate the untreated values due to the heterogeneity in drug response, dosing and variability in baseline lipid values. However, the 30% reduction in LDL-C and 20% in TC was implemented in previous studies[14,28,30].

5. CONCLUSIONS

The FH phenotype can be caused by several different genotypes, especially in paediatric cohorts. The correct identification of the cause of the dyslipidaemia is important for patient management and implementation of the best therapeutic measures for the best patient prognosis. We recommended that the genetic test for the identification of the genetic cause of the hypercholesterolaemia in clinical FH patients should include all the genes described here for the most effective patient diagnosis. Investigation of other genes causing the FH phenotype should be encouraged.

The LDL-C GRS was validated in the Portuguese population and revealed that almost 80% of the FH negative patients could have polygenic hypercholesterolaemia, while a small part have a low score, which could mean that these patients are most likely to have an unknown variant in a new gene. Since the characterization of polygenic traits has also contributed for patient diagnosis, it is also recommended to include the LDL-C GRS in FH diagnosis panels.

ACKNOWLEDGEMENTS

The authors would like to thank all patients and clinicians for participating in the Portuguese Familial Hypercholesterolaemia Study and to all participants of the e_COR Study. The authors gratefully acknowledge Ana Raimundo (MSc) for performing the Lp(a) biochemical determinations, and all technicians of the Unidade de Diagnóstico e Referência (DPSPDNT−INSA) for performing the remaining lipid profile parameters. Cibelle Mariano was supported by a PhD student grant [SFRH/BD/52494/2014].

AUTHORS' CONTRIBUTIONS

CM performed all statistical analysis under MA supervision. CM, MA, MF, SEH and MB performed data analysis and interpretation. ACA and AM performed molecular analysis. JRC performed variant classification. CM and MB drafted the manuscript. MA, ACA, JRC, MF and SEH critically revised the manuscript. MB coordinated and designed the study. All authors contributed to

and approved the final version of the paper and agree to be accountable for all aspects of work ensuring integrity and accuracy.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to disclose.

FINANCIAL DISCLOSURE

The Portuguese FH Study has been funded by Science and Technology foundation, Portuguese Cardiology Society and the National Institute of Health. The e_COR Study was funded by Science and Technology Foundation (PIC/IC/83020/2007) and the National Institute of Health. SEH acknowledges funding from the British Heart Foundation grant (BHF PG08/008) and from the NIHR UCLH BRC.

SUPPLEMENTARY MATERIAL

Supplementary information is available at Genetics in Medicine website.

REFERENCES

1. Nordestgaard BG, Chapman MJ, Humphries SE, et al. Familial hypercholesterolaemia is underdiagnosed and undertreated in the general population: guidance for clinicians to prevent coronary heart disease: Consensus Statement of the European Atherosclerosis Society. *Eur Heart J* 2013; 34: 3478–90.
2. Stenson PD, Mort M, Ball E V, et al. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet* 2014; 133: 1–9.
3. Innerarity TL, Mahley RW, Weisgraber KH, et al. Familial defective apolipoprotein B-100: a mutation of apolipoprotein B that causes hypercholesterolemia. *J Lipid Res* 1990; 31: 1337–49.
4. Motazacker MM, Pirruccello J, Huijgen R, et al. Advances in genetics show the need for extending screening strategies for autosomal dominant hypercholesterolaemia. *Eur Heart J* 2012; 33: 1360–6.
5. Alves AC, Etxebarria A, Soutar AK, et al. Novel functional APOB mutations outside LDL-binding region causing familial hypercholesterolaemia. *Hum Mol Genet* 2014; 23: 1817–28.
6. Abifadel M, Varret M, Rabès J-P, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet* 2003; 34: 154–156.
7. Marduel M, Ouguerram K, Serre V, et al. Description of a large family with autosomal dominant hypercholesterolemia associated with the APOE p.Leu167del mutation. *Hum Mutat* 2013; 34: 83–7.
8. Cenarro A, Etxebarria A, de Castro-Orós I, et al. The p.Leu167del Mutation in APOE Gene Causes Autosomal Dominant Hypercholesterolemia by Down-regulation of LDL Receptor Expression in Hepatocytes. *J Clin Endocrinol Metab* 2016; 101: 2113–21.
9. Rios J, Stein E, Shendure J, et al. Identification by whole-genome resequencing of gene defect responsible for severe hypercholesterolemia. *Hum Mol Genet* 2010; 19: 4313–4318.
10. Hegele RA, Ban MR, Cao H, et al. Targeted next-generation sequencing in monogenic dyslipidemias. *Curr Opin Lipidol* 2015; 26: 103–13.
11. Chora JR, Alves AC, Medeiros AM, et al. Lysosomal acid lipase deficiency: A hidden disease among cohorts of familial hypercholesterolemia? *J Clin Lipidol* 2017; 11: 477–484.e2.
12. Talmud PJ, Shah S, Whittall R, et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study. *Lancet* 2013; 381: 1293–301.
13. Futema M, Shah S, Cooper JA, et al. Refinement of variant selection for the LDL cholesterol genetic risk score in the diagnosis of the polygenic form of clinical familial hypercholesterolemia and replication in samples from 6 countries. *Clin Chem* 2015; 61: 231–8.
14. Khera A V, Won H-H, Peloso GM, et al. Diagnostic Yield and Clinical Utility of Sequencing Familial Hypercholesterolemia Genes in Patients With Severe Hypercholesterolemia. *J Am Coll Cardiol* 2016; 67: 2578–89.
15. Futema M, Plagnol V, Whittall RA, et al. Use of targeted exome sequencing as a diagnostic tool for Familial Hypercholesterolaemia. *J Med Genet* 2012; 49: 644–9.
16. Benn M, Watts GF, Tybjaerg-Hansen A, et al. Familial hypercholesterolemia in the danish general population: prevalence, coronary artery disease, and cholesterol-lowering medication. *J Clin Endocrinol Metab* 2012; 97: 3956–64.
17. Bertolini S, Pisciotta L, Rabacchi C, et al. Spectrum of mutations and phenotypic expression in patients with autosomal dominant hypercholesterolemia identified in Italy. *Atherosclerosis* 2013; 227: 342–348.
18. Bourbon M, Rato Q. Portuguese Familial Hypercholesterolemia Study: presentation of the study and preliminary results. *Rev Port Cardiol* 2006; 25: 999–1013.
19. Medeiros AM, Alves AC, Francisco V, et al. Update of the Portuguese Familial Hypercholesterolaemia Study. *Atherosclerosis* 2010; 212: 553–558.
20. Bonfield JK, Smith K f, Staden R. A new DNA sequence assembly program. *Nucleic Acids Res* 1995; 23: 4992–9.

21. den Dunnen JT, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat* 2000; 15: 7–12.

22. Caridi G, Dagnino M, Di Duca M, et al. A novel splicing mutation causes analbuminemia in a Portuguese boy. *Mol Genet Metab* 2012; 105: 479–483.

23. Chora JR, Medeiros AM, Alves AC, et al. Analysis of publicly available LDLR, APOB, and PCSK9 variants associated with familial hypercholesterolemia: application of ACMG guidelines and implications for familial hypercholesterolemia diagnosis. *Genet Med* 2018; 20: 591–598.

24. Medeiros AM, Alves AC, Bourbon M. Mutational analysis of a cohort with clinical diagnosis of familial hypercholesterolemia: considerations for genetic diagnosis improvement. *Genet Med* 2016; 18: 316–24.

25. Applied BioSystems. TaqMan® OpenArray™ Genotyping System. User Guide.

26. Instituto Nacional de Saúde doutor Ricardo Jorge. *Prevalência de fatores de risco cardiovascular na população portuguesa (e_COR)*. Lisboa: INSA IP, 2019.

27. Medeiros AM, Alves AC, Aguiar P, et al. Cardiovascular risk assessment of dyslipidemic children: analysis of biomarkers to identify monogenic dyslipidemia. *J Lipid Res* 2014; 55: 947–55.

28. Baigent C, Keech A, Kearney PM, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet* 2005; 366: 1267–1278.

29. Peloso GM, Auer PL, Bis JC, et al. Association of Low-Frequency and Rare Coding-Sequence Variants with Blood Lipids and Coronary Heart Disease in 56,000 Whites and Blacks. *Am J Hum Genet* 2014; 94: 223–232.

30. Sniderman AD. Differential response of cholesterol and particle measures of atherogenic lipoproteins to LDL-lowering therapy: implications for clinical practice. *J Clin Lipidol* 2008; 2: 36–42.

31. Schulzeck P, Bojanovski M, Jochim A, et al. Comparison between simvastatin and bezafibrate in effect on plasma lipoproteins and apolipoproteins in primary hypercholesterolaemia. *Lancet (London, England)* 1988; 1: 611–3.

32. Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet (London, England)* 1994; 344: 1383–9.

33. Cannon CP. The IDEAL Cholesterol. *JAMA* 2005; 294: 2492.

34. Kastelein JJP, Akdim F, Stroes ESG, et al. Simvastatin with or without Ezetimibe in Familial Hypercholesterolemia. *N Engl J Med* 2008; 358: 1431–1443.

35. R: The R Project for Statistical Computing.

36. Medeiros AM, Alves AC, Bourbon M. Mutational analysis of a cohort with clinical diagnosis of familial hypercholesterolemia: considerations for genetic diagnosis improvement. *Genet Med* 2015; 18: 1–9.

37. Rall SC, Weisgraber KH, Innerarity TL, et al. Identical structural and receptor binding defects in apolipoprotein E2 in hypo-, normo-, and hypercholesterolemic dysbetalipoproteinemia. *J Clin Invest* 1983; 71: 1023–31.

38. Ghiselli G, Schaefer EJ, Gascon P, et al. Type III hyperlipoproteinemia associated with apolipoprotein E deficiency. *Science* 1981; 214: 1239–41.

39. Wang J, Dron JS, Ban MR, et al. Polygenic Versus Monogenic Causes of Hypercholesterolemia Ascertained ClinicallyHighlights. *Arterioscler Thromb Vasc Biol* 2016; 36: 2439–2445.

40. Stitziel NO, Fouchier SW, Sjouke B, et al. Exome sequencing and directed clinical phenotyping diagnose cholesterol ester storage disease presenting as autosomal recessive hypercholesterolemia. *Arterioscler Thromb Vasc Biol* 2013; 33: 2909–14.

41. Wintjens R, Bozon D, Belabbas K, et al. Global molecular analysis and APOE mutations in a cohort of autosomal dominant hypercholesterolemia patients in France. *J Lipid Res* 2016; 57: 482–91.

42. Sharifi M, Futema M, Nair D, et al. Genetic Architecture of Familial Hypercholesterolaemia. *Curr Cardiol Rep* 2017; 19: 44.

43. Sturm AC, Knowles JW, Gidding SS, et al. Clinical Genetic Testing for Familial Hypercholesterolemia. *J Am Coll Cardiol* 2018; 72: 662–680.

SUPPLEMENTARY DATA

Supplementary Tables

Supplementary Table 1 – The 6 LDL-C genetic risk score SNPs.

Gene	Chr	SNP	Haplotype[a]	Risk allele	Effect size[b]
CELSR2	1	rs629301	..	T	0.15
APOB	2	rs1367117	..	A	0.10
ABCG8	2	rs4299376	..	G	0.071
LDLR	19	rs6511720	..	G	0.18
APOE	19	rs429358	ε2ε2	..	-0.90
		rs7412	ε2ε3	..	-0.40
			ε2ε4	..	0.20
			ε3ε3	..	0
			ε3ε4	..	0.10
			ε4ε4	..	0.20

Chr, chromosome; SNP, single nucleotide polymorphism.

[a]The *APOE* SNPs (rs7412 and rs429358) determined the *APOE* haplotype, an important genetic determinant of LDL-C levels, by resulting in different isoforms of the apolipoprotein E (ApoE): ε2, ε3, and ε4.

[b]Effect sizes are the beta coefficients reported by the Global Lipid Genetic Consortium for each minor allele, taken from previous studies[1,2].

Suplementary Table 2 – Baseline characteristics of all index cases included in the Portuguese Familial Hypercholesterolaemia Study.

Clinical and demographic profile at referral	Children (n=311)	Adults (n=420)
Age, years (SD)*	9.94 (3.69)	45.67 (13.32)
Male gender, n (%)	133 (42.77)	191 (45.48)
BMI, kg/m^2 (SD)*	19.67 (4.05)	26.06 (4.47)
Smoking, n (%)	1 (0.32)	82 (19.52)
Alcohol consumption, n (%)	0	135 (32.14)
Hypertension, n (%)	3 (0.97)	108 (25.71)
Diabetes, n (%)	1 (0.32)	15 (3.57)
Personal history of CVD, n (%)	0	93 (22.14)
Personal history of pCVD, n (%)	0	71 (16.90)
Family history of pCVD, n (%)	55 (17.69)	137 (32.62)
Pharmacological treatment (e.g. statins), n (%)	62 (19.94)	317 (75.48)
On diet, n (%)	145 (46.62)	48 (11.43)
Physically active, n (%)	223 (71.70)	145 (34.52)
Tendon xanthoma, n (%)	1 (0.32)	10 (2.38)
Corneal arcus, n (%)	0	15 (3.57)
Other xanthomas, n (%)	3 (0.97)	32 (7.62)
Lipid profile 1[a*]		
Total cholesterol, mg/dL (IQR)	265.93 (242-290)	321.25 (296-362)
LDL-C, mg/dL (IQR)	189.00 (170-223)	243.00 (210-287)
HDL-C, mg/dL (IQR)	53.00 (45-62)	52.00 (44-63)
Triglycerides, mg/dL (IQR)	76.00 (57-109)	133.00 (96-185)
Lipid profile 2[b*]		
Total cholesterol, mg/dL (IQR)	247.44 (227-284)	298.75 (259-348)
LDL-C, mg/dL (IQR)	178.00 (152-219)	222.00 (177-279)
HDL-C, mg/dL (IQR)	53.00 (45-63)	53.00 (44-64)
Triglycerides, mg/dL (IQR)	75.00 (56-108)	123.00 (91-164)
ApoB, mg/dL (IQR)	115.00 (94-140)	152.63 (119-191)
ApoA1, mg/dL (IQR)	143.00 (125-162)	152.00 (131-172)
Non-HDL-C, mg/dL (IQR)	189.69 (168-233)	244.00 (200-295)
ApoB/ApoA1 ratio, mg/dL (IQR)	0.84 (0.62-1.05)	0.98 (0.72-1.35)
Lp(a), mg/dL (IQR)	29.00 (10-70)	33.00 (9-75)

BMI, body mass index; CVD, cardiovascular disease; pCVD, premature cardiovascular disease; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; ApoB, apolipoprotein B; ApoA1, apolipoprotein A1; Lp(a), lipoprotein(a); SD, standard deviation; IQR, interquartile range.
*Age and BMI are expressed as mean (SD), and biochemical values are expressed as median (IQR).

[a]Lipid profile considered for inclusion in the Portuguese Familial Hypercholesterolaemia (FH) Study. Values used for the analysis were preferably values without medication. When untreated values were not available, TC, LDL-C and apoB values under medication were corrected as described in the Material and Methods section.

[b]Lipid profile at referral. The biochemical analysis was performed in our department as described in Material and Methods section, however when patients were medicated correction factors were applied to estimate untreated values for TC, LDL-C, and apoB, as described in the Material and Methods section.

Supplementary Table 3 – Novel rare variants identified in the Portuguese Familial Hypercholesterolaemia Study between 2014 and 2017.

Gene	DNA alteration	Protein alteration	dbSNP	Functional studies		Pathogenicity[a]	ClinVar ID	Reference
APOE	c.487C>T	p.(Arg163Cys)	rs769455	No	..	LP	17851	Rall et al., 1983[5]
APOE	c.683G>A	p.(Trp228*)	rs121918396	Yes	Lohse et al., 1992[3]	P	17862	Ghiselli et al., 1981[6]
LDLR	c.313+5G>A	p.(Leu64_Pro105delinsSer)	rs879254467	Yes	Liguori et al., 2001[4]	VUS	251136	Liguori et al., 2001[4]
LDLR	c.941-2A>C	..	rs112366278	No	..	P	251553	Chmara et al., 2010[7]
LDLR	c.1186+56_1186+64del	p.=	..	Yes	Graça et al., *in preparation*	LB	*submitted*	Novel
LDLR	c.1434G>A	p.(Gly478Gly)/p.(=)	rs886039832	No	..	VUS	265902	Novel
LDLR	c.1499T>C	p.(Val500Ala)	rs886039833	No	..	VUS	265903	Novel
LDLR	c.1897C>T	p.(Arg633Cys)	rs746118995	No	..	LP	226379	Day et al., 1997[8]
LDLR	c.2214del	p.(Gln739Serfs*26)	..	No	..	P	*submitted*	Novel
LDLR	c.*13A>G	..	rs72658871	No	..	VUS	265909	Novel

dbSNP, Single Nucleotide Polymorphism database; ID, identification number; VUS, variant of unknown significance; LP, likely pathogenic; P, pathogenic; LB, likely benign.

[a]American College of Medical Genetics and Genomics (ACMG) classification for *APOE* variants according to Richards et al., 2015[9] and *LDLR* variants according to Richards et al., 2015[9] and Chora et al., 2017[10].

Supplementary Table 4 – Comparison of the lipid profile between index cases and relatives Familial Hypercholesterolaemia mutation positive patients of the Portuguese Familial Hypercholesterolaemia Study: adult and paediatric cohorts.

Paediatric cohort	Index	Relatives	
Lipid profile[a]	**FH/M+(n=311)**	**FH/M+(n=121)**	P value[b]
Total cholesterol, mg/dL (IQR)	**273.38 (245-302)**	**257.48 (219-294)**	**0.004**
LDL-C, mg/dL (IQR)	**203.64 (179-238)**	**196.00 (154-231)**	**0.002**
HDL-C, mg/dL (IQR)	50.98 (42-58)	50.00 (42-56)	0.608
Triglycerides, mg/dL (IQR)	65.00 (54-88)	71.50 (59-102)	0.081
ApoB, mg/dL (IQR)	129.50 (113-151)	123.76 (102-150)	0.102
ApoA1, mg/dL (IQR)	135.00 (115-147)	132.00 (114-143)	0.219
Non-HDL-C, mg/dL (IQR)	**216.06 (192-253)**	**210.44 (165-243)**	**0.021**
ApoB/ApoA1 ratio, mg/dL (IQR)	0.99 (0.85-1.18)	0.94 (0.75-1.16)	0.957
Lp(a), mg/dL (IQR)	24.64 (10-44)	27.00 (12-53)	0.609
Adults cohort	Index	Relatives	
Lipid profile[a]	**FH/M+ (n=420)**	**FH/M+ (n=309)**	P value[b]
Total cholesterol, mg/dL (IQR)	**331.25 (284-395)**	**299.00 (256-339)**	**<0.001**
LDL-C, mg/dL (IQR)	**263.57 (207-318)**	**225.71 (1867-267)**	**<0.001**
HDL-C, mg/dL (IQR)	51.00 (44-62)	51.00 (43-62)	0.354
Triglycerides, mg/dL (IQR)	112.39 (83-148)	103.27 (71.39-156)	0.601
ApoB, mg/dL (IQR)	**177.63 (141-211)**	**156.29 (129-183)**	**<<0.001**
ApoA1, mg/dL (IQR)	146.50 (127-167)	143.00 (125-164)	0.442
Non-HDL-C, mg/dL (IQR)	**274.88 (224-328)**	**248.25 (205-294)**	**0.021**
ApoB/ApoA1 ratio, mg/dL (IQR)	**1.16 (0.92-1.56)**	**1.06 (0.84-1.33)**	**0.005**
Lp(a), mg/dL (IQR)	35.15 (9-91)	30.00 (14-77)	0.314

FH, Familial Hypercholesterolaemia; FH/M+, FH mutation positive; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; ApoB, apolipoprotein B; ApoA1, apolipoprotein A1; Lp(a), lipoprotein(a); SD, standard deviation; IQR, interquartile range.

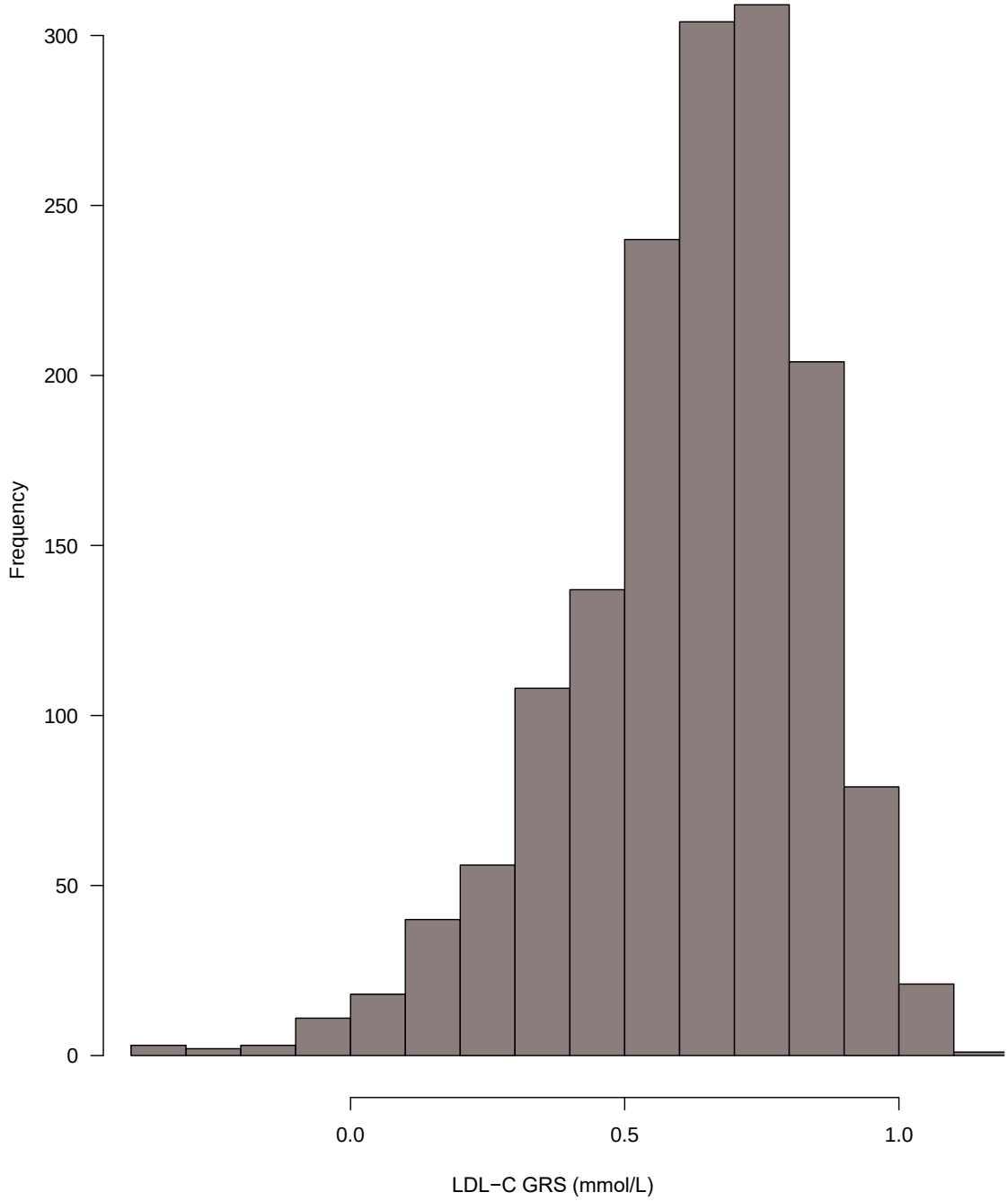
Bochemical lipid values are expressed as median (IQR).

[a]Lipid profile determined in house. Correction factors were used to estimate untreated values for TC, LDL-C, and apoB, as described in the Material and Methods section of our study.

[b]Comparison of FH/M+ patients (HeFH) between index cases and relatives; Statistical significance (P value <0.05), in bold.

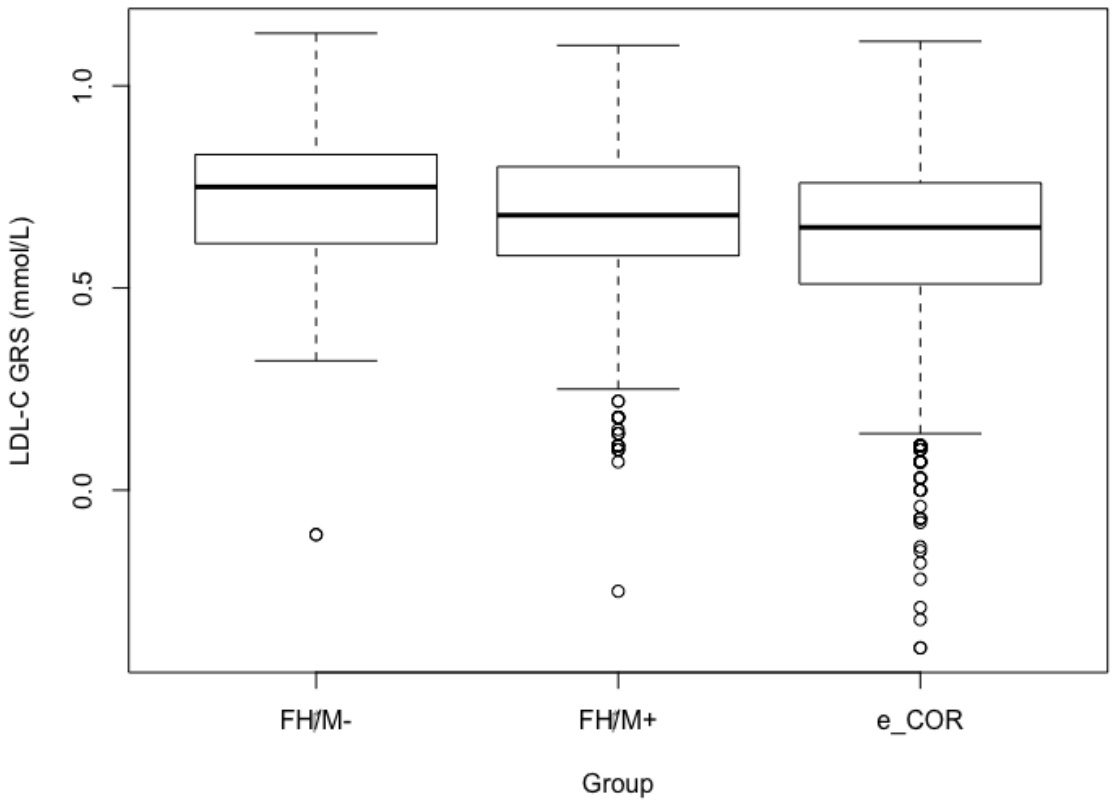
Supplementary Figures

Supplementary Figure 1



Supplementary Figure 1 – Distribution of the weighted sample LDL-C genetic risk score (GRS) in the e_COR sample (n=1318). The LDL-C GRS values ranged -0.30-1.11(mmol/L)

Supplementary Figure 2



Supplementary Figure 2 – LDL-C genetic risk score (GRS) distribution by FH/M- and FH/M+ patients and e_COR sample. Compared with e_COR controls, FH/M- and FH/M+ patients had higher LDL-C GRSP<<0.001).

For FH/M+ paediatric patients, 39.5% (n=30) are above the P75th and 15.8% (n=12) are below the P25th. For FH/M+ adult patients, 34.1% (n=31) are above the P75th and 19.8% (n=18) are below the P25th.

References

1. Talmud PJ, Shah S, Whittall R, et al. Use of low-density lipoprotein cholesterol gene score to distinguish patients with polygenic and monogenic familial hypercholesterolaemia: a case-control study. *Lancet* 2013; 381: 1293–1301.

2. Futema M, Shah S, Cooper JA, et al. Refinement of variant selection for the LDL cholesterol genetic risk score in the diagnosis of the polygenic form of clinical familial hypercholesterolemia and replication in samples from 6 countries. *Clin Chem* 2015; 61: 231–238.

3. Lohse, P.; Brewer, H. B.; Meng, M. S.; Skarlatos, S. I.; LaRosa, J. C.; Brewer, H. B. Familial Apolipoprotein E Deficiency and Type III Hyperlipoproteinemia due to a Premature Stop Codon in the Apolipoprotein E Gene. *J. Lipid Res.* 1992, 33 (11), 1583–1590.

4. Liguori, R.; Bianco, A. M.; Argiriou, A.; Pauciullo, P.; Giannino, A.; Rubba, P.; De Simone, V. LDL Receptor cDNA Sequence Analysis in Familial Hypercholesterolemia Patients: 5 Novel Mutations with High Prevalence in Families Originating from Southern Italy. *Hum. Mutat.* 2001, 17 (5), 433–433.

5. Rall, S. C.; Weisgraber, K. H.; Innerarity, T. L.; Mahley, R. W. Identical Structural and Receptor Binding Defects in Apolipoprotein E2 in Hypo-, Normo-, and Hypercholesterolemic Dysbetalipoproteinemia. *J. Clin. Invest.* 1983, 71 (4), 1023–1031.

6. Ghiselli, G.; Schaefer, E. J.; Gascon, P.; Breser, H. B. Type III Hyperlipoproteinemia Associated with Apolipoprotein E Deficiency. *Science* 1981, 214 (4526), 1239–1241.

7. Chmara, M.; Wasąg, B.; Żuk, M.; Kubalska, J.; Węgrzyn, A.; Bednarska-Makaruk, M.; Pronicka, E.; Wehr, H.; Defesche, J. C.; Rynkiewicz, A.; et al. Molecular Characterization of Polish Patients with Familial Hypercholesterolemia: Novel and recurrentLDLR Mutations. *J. Appl. Genet.* 2010, 51 (1), 95–106.

8. Day, I.; Whittall, R.; O'Dell, S.; Haddad, L.; Bolla, M.; Gudnason, V.; Humphries, S. Spectrum of LDL Receptor Gene Mutations in Heterozygous Familial Hypercholesterolemia. *Hum. Mutat.* 1997, 10 (2), 116–127.

9. Richards, S.; Aziz, N.; Bale, S.; Bick, D.; Das, S.; Gastier-Foster, J.; Grody, WW; Hegde, M; Lyon, E; Spector, E; et al. ACMG Laboratory Quality Assurance Committee. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 2015 17(5), 405–423.

10. Chora, J.; Medeiros, A.; Alves, A.; Bourbon, M. Analysis of Publicly Available LDLR, APOB and PCSK9 Variants Associated with Familial Hypercholesterolemia: Applications of ACMG Guidelines and Implications for Familial Hypercholesterolemia Diagnosis. *Genet. Med.* 2017a, gim2017151.