

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, PhD⁵; Humphries SE, PhD⁶; Bourbon M, PhD^{1,2}.

AFFILIATIONS

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

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

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

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

⁵Centre for Heart Muscle Disease, Institute of Cardiovascular Sciences, University College London, London, United Kingdom.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/cge.13697

⁶Centre 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

Accepted Article

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, especially Suza Almeida and Marta Alvim. Cibelle Mariano was supported by a PhD student grant [SFRH/BD/52494/2014]. Marta Futema is supported by the Fondation Leducq Transatlantic Networks of Excellence Program grant (no 14 CVD03). The work of Marilia Antunes is partially funded by UID/MAT/000016/2019. SEH acknowledges funding from the British Heart Foundation grant (BHF PG08/008) and from the NIHR UCLH BRC. MB acknowledges funding from Science and Technology Foundation (PIC/IC/83020/2007) (PIC/IC/83333/2007) for the e_COR and Portuguese FH Study and from Portuguese Cardiology Society for the Portuguese FH Study.

Conflict of Interest Statement

Authors declares no conflict of interest in relation to this manuscript.

Data Availability Statement

Authors declares absence of shared data.

ABSTRACT

Background: Familial Hypercholesterolaemia (FH) is a monogenic disorder characterised by high LDL-C concentrations and increased cardiovascular risk. However, in clinically defined FH cohorts worldwide, an FH-causing variant is only found in 40-50% of the cases. The aim of this work was to characterise the genetic cause of the FH phenotype in Portuguese clinical FH patients.

Methods and Results: Between 1999 and 2017, 731 index patients (311 children and 420 adults) who met the Simon Broome diagnostic criteria had been referred to our laboratory. *LDLR*, *APOB*, *PCSK9*, *APOE*, *LIPA*, *LDLRAP1*, *ABCG5/8* genes were analysed by PCR amplification and Sanger sequencing. The 6-SNP LDL-C genetic risk score (GRS) for polygenic hypercholesterolaemia was validated in the Portuguese population and cases with a GRS over the 25th percentile were considered to have a high likelihood of polygenic hypercholesterolaemia. An FH-causing mutation was found in 39% of patients (94% in *LDLR*, 5% *APOB* and 1% *PCSK9*), while at least 29% have polygenic hypercholesterolaemia and 1% have other lipid disorders.

Conclusions: A genetic cause for the FH phenotype was found in 503 patients (69%). 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.

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 atherosclerosis¹. FH is one of the most common inherited disorders associated with premature coronary heart disease (pCHD), with a frequency around 1:250 in most populations².

The genetic causes of FH are pathogenic variants mainly in the LDL receptor gene (*LDLR*)¹ or apolipoprotein B gene (*APOB*)³⁻⁵, and gain-of-function mutations in the proprotein convertase subtilisin/kexin 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, ATP-binding cassette sub-family G member 5 or 8 (*ABCG5/8*)⁸ or lysosomal acid lipase (*LIPA*)⁹. Recently, studies have reported that a significant proportion of clinically-diagnosed FH patients where a disease-causing mutation was not found are likely to have a polygenic cause for their hypercholesterolaemia, due to the inheritance of a greater-than-average number of common LDL-C raising single nucleotide polymorphisms (SNPs) with a cumulative effect, leading to an increase in LDL-C above the recommended FH diagnostic criteria^{10,11}. Even in patients with a disease-causing mutation, this polygenic contribution to the FH phenotype is also present, contributing to a variable phenotype in patients with the same FH-causing mutation¹¹. While originally 12 SNPs were used to define the polygenic contribution, a reduced 6-SNPs score has been found to provide adequate information¹⁰⁻¹².

In terms of cardiovascular risk assessment, it is of particular clinical value to distinguish between monogenic and polygenic dyslipidaemia, since for all LDL-C 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-C concentrations¹³. This highlights the concept of length of time of LDL-C exposure, or LDL-C “burden”¹⁴. Because of this, monogenic FH patients warrant treatment with high intensity and effective lipid-lowering therapy 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 only 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, in order to offer appropriate management and prognosis advice 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)^{10,11} in the Portuguese population.

2. MATERIALS 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 731 index patients (311 children and 420 adults) with a clinical diagnosis of FH according to the Simon Broome (SB) criteria, as previously described¹⁸, 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. All adult index cases included in this study fulfilled SB criteria prior to inclusion, 410 possible FH and 10 definite FH. The Dutch Lipid Clinic Network Score (DLCNS) for FH was applied to 390 of the 420 adults from the Portuguese FH Study. For the remaining 30 it was not possible to estimate the DLCNS value, due to missing LDL-C values. A total of 382 adults fulfilled DLCNS of definite (>8) (n=82), probable (>6) (n=169) or possible (>3) (n=131) FH and 8 are unlikely FH (<3). 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 reported¹⁹. Selected patients (Supplementary Table 1), 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 software²⁰ 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) nomenclature²¹ with nucleotide c.1 being A of the ATG initiation codon p.1. For one case the molecular study of the albumin gene (*ALB*) was performed, but in an external laboratory²².

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 uncertain significance (VUS), likely benign or benign, according to the American College of Medical Genetics and Genomics (ACMG) guidelines²³ following specific adaptations described in Chora *et al.*²⁴ The variants reported in the present study were considered novel if they were not described before in public databases^{25,26} or in PUBMED, 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 before²⁷.

Since 2012, our laboratory has participated in an external quality assessment (EQA) programme managed by the European Molecular Genetics Quality Network (EMQN) that provide accreditation on basis of the ISO 17043.

2.3. Polygenic hypercholesterolaemia analysis

A total of 1318 genomic DNA samples from the e_COR Study²⁸ and 410 index cases from the Portuguese FH Study from which quality DNA was available (168 children and 242 adults) were sent to aScidea Computational Biology Solutions Company (Barcelona, Spain) to be genotyped for a set of 6 SNPs, using the OpenArray™ technology (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)¹¹ (Supplementary Table 2).

The e_COR Study population²⁸ 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. The GRS cut-offs were 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 described²⁷.

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. Therapy is expected to decrease the values of TC, LDL-C^{13,29,30} and apoB³¹ by 20%, 30% and 23.7%, respectively. Untreated TG, HDL-C and apoA1 values were not estimated, since the effects of lipid-lowering therapy with statins are not significant on these traits^{32,33}.

2.6. General statistical analysis

Statistical analyses were performed using R (version 3.1.2) software. For comparison of mean concentrations 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 did not fail 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. For comparison of proportions, the 95% confidence intervals (CI) were used and when 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 deviated 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 3, 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 (\pm 3.69) for children and 45.67 (\pm 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 over all regions of Portugal.

3.2. Monogenic dyslipidaemia

A total of 731 index cases were analysed for *LDLR*, *APOB* and *PCSK9* genes. In 282 (39%) patients (128 children and 154 adults) a pathogenic or likely pathogenic variant was

found, including 3 true homozygous and 6 compound heterozygous (2 children and 7 adults); these will be referred as FH mutation positive (FH/M+) subjects (Supplementary Table 4). In 399 (54%) patients (159 children and 239 adults) no potential pathogenic variants were found; these will be referred as FH mutation negative (FH/M-) subjects. Additionally, 18 children and 26 adults were found to have a VUS. Based on phenotype (Supplementary Table 1), selected FH/M- patients were analysed for other possible causes of monogenic dyslipidaemia; 6 cases were found to have another 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), had a VUS.

Since our last report in 2015²⁷, 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 4 and 5). 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).

For both children and adults, the demographic, clinical and biochemical profile of the FH/M+ group were compared to the FH/M- group, and are 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 have 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 \geq 200 mg/dL) was higher than in the FH/M+ patients (20% (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 \geq 120 mg/dL compared to FH/M- (64.94% (n=150/231), P<0.001). No significant differences were observed for Lp(a) values. In adults, there are no statistically significant differences for CVD, premature CVD or family history of CVD between FH/M+ and FH/M- individuals (Table 1).

The mean lipid concentrations 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 6).

3.2.2. Other monogenic disorders

In selected FH/M- patients (Supplementary Table 1), 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 published⁹, one had sitosterolemia (due to homozygosity for the c.1974C>G variant in *ABCG8*) and one had congenital analbuminemia (due to heterozygosity for the 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 in the homozygote state with dysbetalipoproteinemia³⁴, so it was not considered as causative (Supplementary table 5). In the adult cohort, one individual was heterozygous for a variant in *APOE* (c.487C>T), previously associated with autosomal dominant hypercholesterolaemia³⁵ (Supplementary table 5). Additionally, one relative (child) was also found to have LALD and one (adult) sitosterolaemia⁹.

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).

3.3. Polygenic hypercholesterolaemia

The mean value of the LDL-C GRS calculated in the e_COR population was 0.62 (± 0.22) with a mean LDL-C of 135.8 mg/dL (± 46.34) [3.51 mmol/L (± 1.20)] (supplementary Figure 1). The distribution of the LDL-C concentration values by GRS percentiles in e_COR is shown in supplementary Table 7. Individuals above the P75th of LDL-C GRS (0.76) had significantly higher LDL-C ($P \ll 0.001$) than individuals below the P25th (LDL-C GRS 0.51) (supplementary Figure 2). When comparing the mean GRS values between Portuguese e_COR and UK Whitehall II (WHII) controls¹⁰, no significant differences were found (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 410 individuals from whom DNA was available, 168 children and 242 adults. Compared with e_COR controls, both FH/M- (0.73 \pm 0.18) and FH/M+ (0.69 \pm 0.19) patients had higher mean LDL-C GRS ($P \ll 0.001$) (Supplementary Figures 2-3). In this small sample, no statistically

Accepted Article

significant differences were found comparing FH/M- and FH/M+ patients. In the paediatric cohort, the mean value was 0.73 (\pm 0.17) for FH/M- (n=92) and 0.71 (\pm 0.19) for FH/M+ (n=76), and in the adult cohort was 0.72 (\pm 0.19) for FH/M- (n= 151) and 0.68 (\pm 0.20) for FH/M+ (n=91).

The distribution of the GRS in the paediatric FH/M- was: 13% (n=12) below the P25th, 21.7% (n=20) between P25th and P50th; 22.8% (n=21) between P50th and P75th and 42.4% (n=39) above the P75th (Figure 2). In the adult FH/M- cohort the distribution of the GRS was: 10.6% (n=16) below the P25th, 19.9% (n=30) between P25th and P50th; 29.1% (n=44) between P50th and P75th and 40.4% (n=61) above the P75th (Figure 2).

3.4. 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 39% (n=282/731) of patients. We also found 6 patients with other monogenic disorders: 3 with lysosomal acid lipase deficiency (LALD), 1 with sitosterolaemia, 1 with albuminuria and 1 with dysbetalipoproteinemia, representing 0.8% (n=6/731) of all patients. All adults that were tested for polygenic dyslipidaemia fulfilled both the DLCN and SB clinical criteria, so genetic diagnosis rate was the same.

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 15.7% of overall (41 children and 74 adults) had an intermediate polygenic score (Figure 3 and supplementary Figure 4). Of the remaining FH/M- patients, 28 (12 children and 16 adults) presented a score below P25th and are therefore most likely to have an unknown or unidentified cause of monogenic dyslipidaemia (Figure 3 and supplementary Figure 4). By including LDL-C GRS analysis and sequencing of genes involved in other monogenic dyslipidemias we increased the identification rate from 41% to 68.5% in the paediatric cohort and from 37% to 69% in the adult cohort (Figure 3).

4. DISCUSSION

Portuguese FH Study

For the index cases referred to the Portuguese FH Study, in 69% we have identified the cause of hypercholesterolaemia: 39.4% with monogenic hypercholesterolaemia (38.6% have FH and 0.8% other monogenic causes), 16% with intermediate polygenic hypercholesterolaemia score and 14% with high polygenic hypercholesterolaemia score. The variants of unknown significance will need further characterization to be confirm of refute their causality of disease, if all VUS are indeed pathogenic (found in 6%), the total number of positives cases with a monogenic disorder will increase to 45% and the total identification rate to 75%. Cascade screening was performed for relatives of FH cases and 430 relatives were identified with heterozygous FH, one was a compound heterozygous, and 38 relatives with VUS. For individuals with an intermediate or high polygenic score, cascade screening was not performed. Futema and colleagues³⁶ described that a cascade screening for

Accepted Article

individuals with an intermediate or high polygenic score is less cost effective, since, as compared to the 50% of mutation carriers seen in first degree relatives of monogenic families, only ~30% of relatives are likely to have LDL-C elevated above the diagnostic threshold.

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 the cause of FH in Portugal is heterogeneous, with more than 140 different variants (mainly in the *LDLR*), and that new variants are still being found, as reported for other populations.

In the Portuguese FH Study only if a variant is classified as pathogenic or likely pathogenic (ACMG), or if the variant is designated as a VUS, only after functional *in vitro* assessment with proof of affected function, the clinician is informed that the variant found confirms or is consistent with the clinical diagnosis. Due to the establishment of functional studies in our lab^{37,38}, the majority of the mutations found in the Portuguese population (120/142) have a classification of pathogenic or likely pathogenic following ACMG guidelines²⁴, and this results in more than 80% of our cohort of FH/M+ patients having a molecularly confirmed FH diagnosis. The implementation of functional studies as an adjunct to work in diagnostic laboratories improves the genetic diagnosis of FH, and should be encouraged worldwide.

The presence of xanthomas following SB criteria attributes a definite FH diagnosis. Interestingly we have observed a low rate of xanthomas even in adults (7.5%) when compared to other European populations, including our neighbours the Spanish population

(19%). In our clinical FH children, so far no xanthomas have been identified, not even in the homozygotes. The short exposure time may be a justification, but considering that in adults the percentage is low, we believe that the Mediterranean diet may be acting as a protective factor against the development of xanthomas.

Correctly identifying the cause of dyslipidaemia

In studies from many different countries, in 50-60% of individuals with a clinical diagnosis of FH the molecular cause of their hypercholesterolaemia could not be identified¹. 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 colleagues³⁹ 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 and polygenic causes of hypercholesterolaemia in patients with a clinical diagnosis of FH, but without an identified FH causing variant. This has led to an increase in the genetic diagnosis of dyslipidaemia in additional 30% of patients (from 38.6% to 68.8%).

In the great majority of the index patients, the cause of the dyslipidaemia is explained by a functional *LDLR* variant (36%), 1.9% by a functional *APOB* variant and 0.4% by a functional *PCSK9* variant. Additionally, in at least 29% of patients, a polygenic dyslipidemia is the probable cause and in 0.8% other monogenic disorders are the cause of the phenotype. These other causes are more prevalent in children, representing 1.6% 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, so as to allow a more complete diagnosis and management. There are considerable differences in the treatment of these different molecular causes of the FH phenotype.

Higher Lp(a) has also been described as a possible FH phenocopy, but in our sample, although the FH/M+ patients had higher median concentrations of Lp(a) than the FH/M- group, the differences were not statistically significant, so Lp(a) concentrations do not seem to contribute to the clinical phenotype of FH in our sample. However, it is possible that increasing the number of individuals with measures of Lp(a) concentrations may identify a group of patients where this trait is the cause of the clinical diagnosis of FH.

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. In this paper we have used the 6-SNP GRS and used a cut-off of greater than the 25th percentile as being highly likely for an individual to have a polygenic cause of their FH phenotype. This is based on statistical modelling that was carried out previously^{10,11} and as has recently been described in more detail³⁶. The analysis shows that the probability of a polygenic cause for elevated LDL-C in FH/M- individuals with a score above the 25th percentile is >98% for a range of frequencies of an unknown monogenic cause (modelled as between 0.001-0.01). This fully justifies that, except for those with a SNP score in the lowest 25%, all other FH/M subjects have a high probability of a polygenic case explaining their FH phenotype.

A limitation of our study is that not all FH/M- patients were genotyped for the polygenic score (only 243/398 were analysed) due to DNA constraints, so the proportion of patients with polygenic hypercholesterolaemia could only be estimated. 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^{10,11}. Also, especially for adults, the untreated TC, LDL-C and apoB values had to be estimated for those undergoing lipid-lowering therapy and this 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^{13,29,31}.

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 previously¹⁰. 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 did not find a statistical difference in LDL-C values, partly because of the selection criteria and partly due to the small sample size.

Unknown dyslipidaemia in the FH mutation negative patients

Twenty eight FH/M- patients presented a low LDL-C GRS (below the P25th) and are likely to have other unknown or unidentified 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 phenotype³⁹, 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, as well as being modulated by non-genetic factors as life style⁴⁰. 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 value¹³, 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 of addressing other cardiovascular risk factors in childhood, to reduce CVD rates later in adulthood.

CONCLUSIONS

With the strategy presented, it was possible to identify the cause of the hypercholesterolaemia in 503 patients representing 69% of our cohort. This work shows that 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. Our data support the view that the genetic diagnosis of clinical FH patients would benefit from the inclusion of all the genes studied in this work and to include the LDL-C GRS in a next-generation target panel. Such a panel is already implemented in our laboratory, and will not have a significant increase in cost. This has also been already recommended in the last consensus paper authored by international FH experts⁴¹. Investigation of other genes causing the FH phenotype in established FH cohorts should be encouraged.

AUTHORS' CONTRIBUTIONS

CM performed all statistical analysis under MA supervision. CM, ACA, 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, AMM, JRC, MF and SEH critically revised the manuscript. MB coordinated and designed the study. All authors approved the final version of the paper and agree to be accountable for all aspects of work ensuring integrity and accuracy.

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(45):3478-3490. doi:10.1093/eurheartj/eh273
2. Akioyamen LE, Genest J, Shan SD, et al. Estimating the prevalence of heterozygous familial hypercholesterolaemia: a systematic review and meta-analysis. *BMJ Open*. 2017;7(9):e016461. doi:10.1136/bmjopen-2017-016461
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(8):1337-1349.
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(11):1360-1366. doi:10.1093/eurheartj/ehs010
5. Alves AC, Etxebarria A, Soutar AK, Martin C, Bourbon M. Novel functional APOB mutations outside LDL-binding region causing familial hypercholesterolaemia. *Hum Mol Genet*. 2014;23(7):1817-1828. doi:10.1093/hmg/ddt573
6. Abifadel M, Varret M, Rabès J-P, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. *Nat Genet*. 2003;34(2):154-156. doi:10.1038/ng1161
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(1):83-87. doi:10.1002/humu.22215

8. Rios J, Stein E, Shendure J, Hobbs HH, Cohen JC. Identification by whole-genome resequencing of gene defect responsible for severe hypercholesterolemia. *Hum Mol Genet.* 2010;19(22):4313-4318. doi:10.1093/hmg/ddq352
9. 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(2):477-484.e2. doi:10.1016/j.jacl.2016.11.002
10. 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(9874):1293-1301. doi:10.1016/S0140-6736(12)62127-8
11. 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(1):231-238. doi:10.1373/clinchem.2014.231365
12. Hegele RA, Ban MR, Cao H, McIntyre AD, Robinson JF, Wang J. Targeted next-generation sequencing in monogenic dyslipidemias. *Curr Opin Lipidol.* 2015;26(2):103-113. doi:10.1097/MOL.0000000000000163
13. 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(22):2578-2589.

doi:10.1016/j.jacc.2016.03.520

14. Starr B, Hadfield SG, Hutten BA, et al. Development of sensitive and specific age- and gender-specific low-density lipoprotein cholesterol cutoffs for diagnosis of first-degree relatives with familial hypercholesterolaemia in cascade testing. *Clin Chem Lab Med*. 2008. doi:10.1515/CCLM.2008.135
15. Futema M, Plagnol V, Whittall RA, Neil HAW, Humphries SE. Use of targeted exome sequencing as a diagnostic tool for Familial Hypercholesterolaemia. *J Med Genet*. 2012;49(10):644-649. doi:10.1136/jmedgenet-2012-101189
16. Benn M, Watts GF, Tybjaerg-Hansen A, Nordestgaard BG. Familial hypercholesterolemia in the danish general population: prevalence, coronary artery disease, and cholesterol-lowering medication. *J Clin Endocrinol Metab*. 2012;97(11):3956-3964. doi:10.1210/jc.2012-1563
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(2):342-348. doi:10.1016/j.atherosclerosis.2013.01.007
18. Bourbon M, Rato Q. Portuguese Familial Hypercholesterolemia Study: presentation of the study and preliminary results. *Rev Port Cardiol*. 2006;25(11):999-1013. <http://www.ncbi.nlm.nih.gov/pubmed/17274457>.
19. Medeiros AM, Alves AC, Francisco V, Bourbon M. Update of the Portuguese Familial Hypercholesterolaemia Study. *Atherosclerosis*. 2010;212(2):553-558. doi:S0021-9150(10)00554-X [pii]10.1016/j.atherosclerosis.2010.07.012

- Accepted Article
20. Bonfield JK, Smith K f, Staden R. A new DNA sequence assembly program. *Nucleic Acids Res.* 1995;23(24):4992-4999.
 21. den Dunnen JT, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat.* 2000;15(1):7-12.
doi:10.1002/(SICI)1098-1004(200001)15:1<7::AID-HUMU4>3.0.CO;2-N
 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(3):479-483.
doi:10.1016/j.ymgme.2011.12.009
 23. Richards S, Aziz N, Bale S, et al. 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. doi:10.1038/gim.2015.30
 24. Chora JR, Medeiros AM, Alves AC, Bourbon M. 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(6):591-598. doi:10.1038/gim.2017.151
 25. Leigh S, Futema M, Whittall R, et al. The UCL low-density lipoprotein receptor gene variant database: pathogenicity update. *J Med Genet.* November 2016;jmedgenet-2016-104054. doi:10.1136/jmedgenet-2016-104054
 26. Iacocca MA, Chora JR, Carrié A, et al. ClinVar database of global familial hypercholesterolemia-associated DNA variants. *Hum Mutat.* 2018;39(11):1631-1640.

doi:10.1002/humu.23634

27. 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(4):316-324. doi:10.1038/gim.2015.71
28. Instituto Nacional de Saúde doutor Ricardo Jorge. *Prevalência de Fatores de Risco Cardiovascular Na População Portuguesa (E_COR)*. Lisboa: INSA IP; 2019. doi:ISBN: 978-989-8794-60-4
29. 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(9493):1267-1278.
30. 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(2):223-232. doi:10.1016/j.ajhg.2014.01.009
31. 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(1):36-42. doi:10.1016/j.jacl.2007.12.006
32. Cannon CP. The IDEAL Cholesterol. *JAMA*. 2005;294(19):2492. doi:10.1001/jama.294.19.2492
33. Kastelein JJP, Akdim F, Stroes ESG, et al. Simvastatin with or without Ezetimibe in Familial Hypercholesterolemia. *N Engl J Med*. 2008;358(14):1431-1443. doi:10.1056/NEJMoa0800742

- Accepted Article
34. Ghiselli G, Schaefer EJ, Gascon P, Breser HB. Type III hyperlipoproteinemia associated with apolipoprotein E deficiency. *Science*. 1981;214(4526):1239-1241.
 35. Rall SC, Weisgraber KH, Innerarity TL, Mahley RW. Identical structural and receptor binding defects in apolipoprotein E2 in hypo-, normo-, and hypercholesterolemic dysbetalipoproteinemia. *J Clin Invest*. 1983;71(4):1023-1031.
 36. Futema M, Bourbon M, Williams M, Humphries SE. Clinical utility of the Polygenic LDL-C SNP score in Familial Hypercholesterolemia. *Atherosclerosis*. 2018:(accepted).
 37. Silva S, Alves AC, Patel D, et al. In vitro functional characterization of missense mutations in the LDLR gene. *Atherosclerosis*. 2012;225(1):128-134. doi:S0021-9150(12)00566-7 [pii] 10.1016/j.atherosclerosis.2012.08.017
 38. Benito-Vicente A, Alves AC, Etxebarria A, Medeiros AM, Martin C, Bourbon M. The importance of an integrated analysis of clinical, molecular, and functional data for the genetic diagnosis of familial hypercholesterolemia. *Genet Med*. 2015;17(12):980-988. doi:10.1038/gim.2015.14
 39. Wang J, Dron JS, Ban MR, et al. Polygenic Versus Monogenic Causes of Hypercholesterolemia Ascertained Clinically Highlights. *Arterioscler Thromb Vasc Biol*. 2016;36(12):2439-2445. doi:10.1161/ATVBAHA.116.308027
 40. Mymin D, Salen G, Triggs-Raine B, Waggoner DJ, Dembinski T, Hatch GM. The natural history of phytosterolemia: Observations on its homeostasis. *Atherosclerosis*. 2018;269:122-128. doi:10.1016/j.atherosclerosis.2017.12.024
 41. Sturm AC, Knowles JW, Gidding SS, et al. Clinical Genetic Testing for

Familial Hypercholesterolemia. *J Am Coll Cardiol.* 2018;72(6):662-680.

doi:10.1016/j.jacc.2018.05.044

Accepted Article

Legends of Figures

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.

Figure 2 -Distribution of general population (e_COR) and FH/M- adults and children by GRS.

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 or likely pathogenic variants in either *LDLR*, *APOB* or *PCSK9*, intermediate GRS (genetic risk score) to patients with a LDL-C GRS between P25th and P75th, high GRS to patients with GRS above P75th. 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).

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

Clinical and demographic profile	Paediatric cohort (n=285) ^a			Adult cohort (n=386) ^a		
	FH/M-	FH/M+	P value ^c	FH/M-	FH/M+	P value ^c
n (%)	159 (51.1)	126 (40.5)	ND	239 (56.9)	147 (35.0)	ND
Age, year (SD)*	9.77 (3.5)	10.0 (3.9)	NS	47.2 (11.9)	43.73 (14.9)	0.008
Male gender, n (%)	63 (39.6)	63 (50.00)	NS	118 (49.4)	57 (38.8)	NS
BMI, kg/m ² (SD)*	20.3 (4.3)	18.9 (3.5)	0.012	26.1 (4.2)	26.1 (4.9)	NS
Smoking, n (%)	0	1 (0.8)	ND	56 (23.4)	21 (14.3)	0.025
Alcohol consumption, n (%)	0	0	ND	86 (35.9)	40 (27.2)	NS
Hypertension, n (%)	3 (1.89)	0	ND	61 (25.5)	41 (27.9)	NS
Diabetes, n (%)	1 (0.63)	0	ND	9 (3.8)	5 (3.4)	ND
Personal history of CVD, n (%)	0	0	ND	51 (21.3)	31 (21.1)	NS
Personal history of pCVD, n (%)	0	0	ND	41 (17.2)	20 (13.6)	NS
Family history of pCVD, n (%)	26 (16.4)	24 (19.1)	NS	80 (33.5)	43 (29.3)	NS
Pharmacological treatment (e.g. statins), n (%)	18 (11.3)	36 (28.6)	<0.001	183 (76.6)	109 (74.2)	NS
On diet, n (%)	86 (54.1)	50 (39.7)	0.021	30 (12.6)	14 (9.5)	NS
Physically active, n (%)	119 (74.8)	86 (68.3)	NS	81 (33.9)	52 (35.4)	NS
Tendon xanthoma, n (%)	0	0	ND	0	10 (6.8)	ND
Corneal arcus, n (%)	0	0	ND	1 (0.4)	13 (8.8)	ND
Other xanthomas, n (%)	0	1 (0.8)	ND	13 (5.4)	16 (10.9)	ND
Lipid profile^b						
Total cholesterol, mg/dL (IQR)†	239.2 (211-259)	273.4 (245-302)	<<0.001	285.0 (242-321)	331.3 (284-395)	<<0.001
LDL-C, mg/dL (IQR) †	156.9 (135-178)	203.6 (179-238)	<<0.001	201.7 (158-246)	263.6 (207-318)	<<0.001
HDL-C, mg/dL (IQR) †	57.2 (49-68)	50.9 (42-58)	<0.001	55.0 (44-64)	51.0 (44-62)	NS
Triglycerides, mg/dL (IQR) †	84.5 (65-119)	65.0 (54-88)	<0.001	129.0 (100-178)	112.4 (83-148)	<0.001
ApoB, mg/dL (IQR) †	99.5 (85-122)	129.5 (113-151)	<<0.001	136.8 (136-179)	177.6 (141-211)	<<0.001
ApoA1, mg/dL (IQR) †	150.0 (136-174)	135.0 (115-147)	<<0.001	155.7 (136-179)	146.5 (127-167)	0.005
Non-HDL-C, mg/dL (IQR) †	172.0 (153-194)	216.1 (192-253)	<<0.001	222.1 (182-270)	274.9 (224-328)	<<0.001
ApoB/ApoA1 ratio, mg/dL (IQR) †	0.7 (0.5-0.8)	1.0 (0.85-1.18)	<<0.001	0.9(0.68-1.17)	1.2 (0.9-1.6)	<<0.001
Lp(a), mg/dL (IQR) †	32.0 (10-80)	24.6 (10-44)	0.081	30.0 (9-72)	35.2 (9-92)	NS

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; NS, Not statistically significant (P value >>0.05) Age and BMI are expressed as mean (SD), while biochemical lipid values are expressed as median (IQR). ^aOnly 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. ^bLipid 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. ^cStatistical significance (P value <0.05), in bold.



