Frequency of carriers for rare metabolic diseases in a Brazilian cohort of 320 patients

Caio Robledo D’Angioli Costa Quaio1,2,3,5 · Caroline Monaco Moreira2 · Christine Hsiaooyun Chung2 · Sandro Felix Perazzio2,4 · Aurelio Pimenta Dutra2 · Chong Ae Kim1

Received: 19 October 2021 / Accepted: 8 February 2022 / Published online: 28 February 2022 © The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Background Several metabolic disorders follow an autosomal recessive inheritance pattern. Epidemiological information on these disorders is usually limited in developing countries. Our objective is to assess carrier frequencies of rare autosomal recessive metabolic diseases in a cohort of Brazilian patients that underwent molecular investigation with exome sequencing and estimate the overall frequency of these diseases using the Hardy–Weinberg equation.

Methods and results We reviewed the molecular findings of 320 symptomatic patients who had carrier status for recessive diseases actively searched. A total of 205 rare variants were reported in 138 different genes associated with metabolic diseases from 156 patients, which represents that almost half (48.8%) of the patients were carriers of at least one heterozygous pathogenic/likely pathogenic (P/LP) variant for rare metabolic disorders. Most of these variants are harbored by genes associated with multisystemic involvement. We estimated the overall frequency for rare recessive metabolic diseases to be 10.96/10,000 people, while the frequency of metabolic diseases potentially identified by newborn screening was estimated to be 2.93/10,000.

Conclusions This study shows the potential research utility of exome sequencing to determine carrier status for rare metabolic diseases, which may be a possible strategy to evaluate the clinical and social burden of these conditions at the population level and guide the optimization of health policies and newborn screening programs.

Keywords Whole exome sequencing · Rare diseases · Carrier frequency · Recessive Mendelian diseases · Metabolic diseases · Inborn errors of metabolism

Background

Inborn metabolic diseases (in this article referred to as “metabolic diseases”) are a complex group of disorders of genetic etiology that affect breakdown or biosynthesis of substances within specific pathways, recognizable by specific biochemical tests [1]. The majority of these disorders present monogenic etiology, autosomal recessive inheritance and multisystemic involvement. The determination of exact etiology may be challenging and involve biochemical studies searching for abnormal metabolites or enzymatic assays, genomic approaches with next generations sequencing (NGS) or even other growing methods, such as glycomics, proteomics or lipidomics [1].

A rapid diagnosis is key for favorable long-term prognosis for several of these disorders because (1) some affected individuals can die during an acute metabolic crisis and (2) several of these disorders are treatable by vitamins, special
diets, cleansing drugs or other newly-developed treatments [2]. For this reason, the majority of countries have created universal newborn screening for several treatable metabolic disorders that are potentially recognized through this early approach. These screening programs, though, may vary widely. The Newborn Screening ACT Sheets and Algorithms (ACMG ACT Sheets) [3] are a good resource for mapping metabolic diseases that may be identified through newborn screening.

Considering the clinical relevance of these conditions, the possibility of early identification of several diseases through newborn screening and the potential to dramatically change the outcomes with early treatment, algorithms that allow standardized recognition of these diseases become crucial. A recent thorough nosology has mapped 1015 metabolic diseases in 130 groups and an additional 111 diseases lacking definite classification or poorly characterized [4]. These 130 groups of metabolic diseases are divided into nine major groups: A. disorders of nitrogen-containing compounds; B. disorders of vitamins, cofactors, metals and minerals; C. disorders of carbohydrates; D. mitochondrial disorders of energy metabolism; E. disorders of lipids; F. Disorders of tetrapyrroles; G. Storage disorders; H. Disorders of peroxisomes and oxalate and I. Congenital disorders of glycosylation.

Epidemiological information on metabolic diseases may also be useful to guide health policies and newborn screening programs, though this information is often inadequate, particularly in developing countries. Our objective is to evaluate carrier frequencies of rare autosomal recessive metabolic diseases in a cohort of Brazilian patients that underwent molecular investigation with exome sequencing and estimate the overall frequency of these diseases using the Hardy–Weinberg equation.

Methods

Selection of cases, molecular analysis, and bioinformatics

We reviewed the molecular findings of exome sequencing from 320 symptomatic patients who had carrier status for recessive diseases actively searched. These patients represent a convenience sample and were selected from an original cohort of 500 symptomatic patients who had undergone molecular analysis for suspected diseases with genetic etiology from 2016 to 2020 in facilities of the Fleury Group. Full details of the clinical features of patients, molecular analysis, bioinformatics protocols, clinical data, and molecular data for primary and secondary findings were previously published [5]; a clinical summary is available at Supplementary Table 1. Details regarding molecular analysis, bioinformatics protocols and quality parameters are available in Supplementary Table 10.

All patients or their legal guardians provided consent before exome analysis, and this study was granted ethics committee approval from both institutions involved (Plataforma Brasil; CAAE# 02617018.3.0000.5474; Fleury# 3.372.339).

Carrier status identification, variant classification and gene categories

Carrier status for autosomal recessive diseases was actively searched for all 320 patients according to the following protocol of analysis: at least two parallel analyses were performed to preselect variants considering (1) relevant reports from databases (e.g., ClinVar and HGMD) and the literature and (2) variants that were not reported in ClinVar that had an allele frequency < 1%, had relevant molecular impact (e.g., potentially leading to loss of function) or an in silico prediction of a functional impact and was located in a clinically relevant gene; these preselected variants were then discussed by a board comprising three experts. Variants were classified according to ACMG guidelines [6].

We excluded the following from our analysis: (1) variants for common autosomal recessive diseases with high allele frequencies and low penetrance; (2) variants associated with no clinical impact (such as traits); (3) variants associated with conditions in which the diagnosis depends on other tests (e.g., variants in BTD, G6PD, HFE and SERPINA1); and (4) genes with pseudogenes with high homology (such as CYP21A2).

We have used nosologies, reviews and professional guidelines available in the literature to divide all genes into clinical categories. The overall frequency of pathogenic or likely pathogenic variants associated with all clinical categories initially studied (intellectual disability, cancer, muscular diseases, ciliopathies, skeletal disorders, immune disorders, epilepsy, hearing loss, retinitis pigmentosa and others) along with overall estimations of frequencies of autosomal recessive diseases were previously published [7].

In this article, we explore novel data regarding the rare variants associated with metabolic disorders; the group of genes part of metabolic disorders was selected based on the nosology by Ferreira et al. [4]. The complete list of variants and genes for metabolic disorders along with nomenclature and classification information of all variants are available in the Supplementary Tables 2 and 3.

The nosology divides all metabolic diseases into the following categories: A. Disorders of nitrogen-containing compounds (Groups 1–24); B. Disorders of vitamins, cofactors, metals and minerals (Groups 25–45); C. Disorders of carbohydrates (Groups 26–53); D. Mitochondrial disorders of energy metabolism (Groups 54–81); E. Disorders of
lipids (Groups 82–97); F. Disorders of tetrapyrroles (Groups 98–99); G. Storage disorders (Groups 100–108); H. Disorders of peroxisomes and oxalate (Groups 109–114); I. Congenital disorders of glycosylation (Groups 115–130); and finally, Disorders with individual families described or poorly characterized. We used this division to attribute each individual gene to a specific group of metabolic disorders.

**Recessive disease frequency estimation**

We used the Hardy–Weinberg equation to estimate disease frequencies ($q^2$) based on the respective carrier frequencies (2pq) observed in this study [8]. For this purpose, we considered random mating and the approximation $p \sim 1$.

**Clinical impact**

We used the Human Phenotype Ontology (HPO) database to map possible phenotypic abnormalities encountered in genes associated with metabolic diseases [9]. We chose the following HPO terms that correspond to common clinical findings in metabolic diseases: Intellectual disability (HP:0001249), Encephalopathy (HP:0001298), Seizure (HP:0001250), Acidosis (HP:0001941), Hypoglycemia (HP:0001943), Facial dysmorphism (HP:0001999), Growth abnormality (HP:0001507), Liver disease (HP:0001392), Muscle disease (HP:0003011), Kidney disease (HP:0000077) and Eye disease (HP:0000478).

**Newborn screening diseases/genes**

We used the Newborn Screening ACT Sheets and Algorithms (ACMG ACT Sheets) [3] for mapping metabolic diseases that may be identified through newborn screening. The main classes of metabolic diseases catalogued by these documents are: amino acidemias, organic acidemias, fatty acid oxidation disorders, galactosemias and lysosomal storage diseases. We used the IEMbase platform [10] and OMIM to assess the list of genes associated with these diseases. The complete list of metabolic diseases potentially identified through newborn screening along with the genes associated with these diseases are available in the Supplementary Table 8.

**Results**

**Variant characteristics and frequency of heterozygotes**

We observed a total of 205 occurrences of 172 different variants harbored by 138 different genes associated with metabolic diseases in 156 patients; 23 variants recurred more than once, ranging from two to six times. This result implies that almost half (48.8%) of the patients were carriers of at least one heterozygous P/LP variant for rare metabolic disorders. On the other hand, P/LP variants associated with metabolic disorders were not found in 164 (51.3%) patients. Individual variant data are available in Supplementary Table 2.

In total, 79 different variants were classified as pathogenic and 93 as likely pathogenic according to ACMG criteria. The majority of the 172 different variants was found in the literature ($n = 124; 72.1\%$) and ClinVar Database ($n = 128; 74.4\%$). In contrast, 39 variants (22.7%) were exclusively found in our work: these non-reported variants were classified as likely pathogenic because all of them were considered rare in population databases (frequency lower than 0.01%; ACMG PM2 criteria) and presented a predicted loss-of-function mechanism (such as stop codon, frameshift; ACMG PVSI criteria). Table 1 shows the distribution of genes, variants and the estimated frequency of diseases ($q^2$; estimation per 10,000 individuals) per group of metabolic disorders.

**Genes and their potential clinical impact**

Considering all 138 different genes, 33 (23.9%) belong to group A (disorders of nitrogen-containing compounds), 21 (15.2%) to group D (mitochondrial disorders of energy metabolism) and 19 (13.8%) to group B (disorders of vitamins, cofactors, metals and minerals).

We summed the frequencies of all different P/LP variants per gene and the combined frequency of variants for the most frequent diseases and their related genes are as follows: (1) 1.9% for Gaucher disease (GBA), Phenylketonuria (PAH) and Aicardi-Goutieres syndrome (RNASEH2B); (2) 1.6% for Spastic paraplegia 7 (SPG7); (3) 1.3% for Hereditary fructose intolerance (ALDOB), Wilson disease (ATP7B), CPT II deficiency (CPT2), Tay-Sachs disease (HEXA), and PMM2-CDG [congenital disorder of glycosylation] (PMM2).

Considering HPO terms used to estimate the clinical impact and relevance of the genes identified in this study, we observed that the majority ($n = 125; 90.6\%$) presented two or more clinical manifestations; more than half ($n = 81; 58.7\%$) presented five or more manifestations; eleven genes (8%) presented a sole manifestation and two genes (1.4%; CUBN, FMO3) did not present any of the 11 clinical manifestation in HPO database. The complete list of genes and their corresponding HPOs are available in the Supplementary Tables 5 and 6. Table 2 shows the frequency of 11 clinical manifestations from the HPO database in the 138 genes identified in our study.
Variants

Variants harbored by genes that belong to group A (disorders of nitrogen-containing compounds) were most frequent occurrences (n = 51; 24.9%); this group is followed by group G (storage disorders; n = 31; 15.1%) and group D (mitochondrial disorders of energy metabolism; n = 29; 14.1%). On the other hand, variants of group F (disorders of tetrapyrroles) were the least frequent (n = 6; 2.9%) and we did not find variants harbored by genes associated with disorders with individual families described or poorly characterized (n = 0). Table 1 (third column) shows the distribution of occurrences of pathogenic or likely pathogenic (P/LP) variants per group of metabolic disorders.

Metabolic diseases frequency estimation

As stated before, we used the Hardy–Weinberg equation to estimate disease frequencies (q^2) based on the respective carrier frequencies observed in this study. The combined frequency considering all metabolic diseases was estimated to be 10.96/10,000 (Supplementary Table 3).

Considering only the top nine frequencies of recessive metabolic diseases, their combined frequency is estimated to be 5.2 per 10,000 people. The individual frequencies for these nine diseases per 10,000 individuals are as follow: Gaucher disease (0.88), phenylketonuria (0.88), Aicardi-Goutieres syndrome (0.88), spastic paraplegia (0.61), hereditary fructose intolerance (0.39), Wilson disease (0.39), CPT II deficiency (0.39), Tay-Sachs disease (0.39), PMM2-CDG. | Groups of metabolic disorders | Number of different genes (%) | Occurrences of P/LP variants (%) | Combined allele frequencies | Predicted frequency of affected individuals (q^2; per 10,000 people) |
|-----------------------------|--------------------------------|---------------------|-----------------------------|-----------------------------|
| A. Disorders of nitrogen-containing compounds Groups 1–24 | 33 (23.9%) | 51 (24.9%) | 0.159 | 3.10 |
| B. Disorders of vitamins, cofactors, metals and minerals Groups 25–45 | 19 (13.8%) | 27 (13.2%) | 0.084 | 1.25 |
| C. Disorders of carbohydrates Groups 26–53 | 11 (8%) | 15 (7.3%) | 0.047 | 0.71 |
| D. Mitochondrial disorders of energy metabolism Groups 54–81 | 21 (15.2%) | 29 (14.1%) | 0.091 | 1.44 |
| E. Disorders of lipids Groups 82–97 | 14 (10.1%) | 18 (8.8%) | 0.056 | 0.78 |
| F. Disorders of tetrapyrroles Groups 98–99 | 4 (2.9%) | 6 (2.9%) | 0.019 | 0.24 |
| G. Storage disorders Groups 100–108 | 17 (12.3%) | 31 (15.1%) | 0.097 | 2.08 |
| H. Disorders of peroxisomes and oxalate Groups 109–114 | 4 (2.9%) | 7 (3.4%) | 0.022 | 0.37 |
| I. Congenital disorders of glycosylation Groups 115–130 | 15 (10.9%) | 21 (10.2%) | 0.066 | 1.00 |
| Disorders with individual families described or poorly characterized | – | – | – | – |
| Total | 138 (100%) | 205 (100%) | 0.641 | 10.96 |

The first column represents the groups of metabolic disorders according to the nosology by Ferreira et al. (2019); the second column represents the number of different genes per group of disorders; the third column congregates all occurrences of variants per group per group of diseases; the fourth column represents the combined allele frequencies per group of metabolic disorders in the cohort of 320 patients; the last column represents the predicted frequency of affected individuals (q^2; p ~ 1) per 10,000 individuals (a scale of color was used to compare the distribution of frequencies). The last row congregates all groups.
Table 2  Clinical manifestations associated with metabolic diseases genes

| Clinical manifestation                  | Number of genes | Percentage (total genes = 138) |
|----------------------------------------|-----------------|--------------------------------|
| Muscle disease HP:0003011               | 111             | 80%                            |
| Eye disease HP:0000478                 | 96              | 70%                            |
| Growth abnormality HP:0001507          | 91              | 66%                            |
| Seizure HP:0001250                     | 88              | 64%                            |
| Intellectual disability HP:0001249     | 76              | 55%                            |
| Liver disease HP:0001392               | 64              | 46%                            |
| Kidney disease HP:0000077              | 45              | 33%                            |
| Acidosis HP:0001941                    | 44              | 32%                            |
| Hypoglycemia HP:0001943                | 27              | 20%                            |
| Facial dysmorphism HP:0001999          | 25              | 18%                            |
| Encephalopathy HP:0001298              | 19              | 14%                            |

The table shows the number of genes presenting each manifestation (second column) and the corresponding frequency in percentage (last column) of 11 clinical phenotypes studied (first column). The clinical phenotypes and the complete list of genes associated with them was based on the Human Phenotype Ontology (HPO) database. A scale of color was used to compare the distribution of frequencies.

Table 3  Distribution of genes, occurrences of pathogenic or likely pathogenic (P/LP) variants, combined allele frequency and predicted frequency of affected individuals per group of metabolic disorders potentially identified by newborn screening

| Metabolic disorders potentially identified by newborn screening | Number of different genes | Occurrences of P/LP variants | Combined allele frequencies | Predicted frequency of affected individuals (q^2; per 10,000 people) |
|----------------------------------------------------------------|---------------------------|------------------------------|-----------------------------|---------------------------------------------------------------|
| Amino acidemias                                                | 5                         | 10                           | 0.016                       | 0.98                                                          |
| Fatty acid oxidation disorders                                 | 5                         | 8                            | 0.013                       | 0.49                                                          |
| Galactosemias                                                  | 1                         | 1                            | 0.002                       | 0.02                                                          |
| Organic acidemias                                              | 10                        | 12                           | 0.019                       | 0.44                                                          |
| Lysosomal storage diseases                                     | 3                         | 9                            | 0.014                       | 1.00                                                          |
| Total                                                          | 24                        | 40                           | 0.063                       | 2.929                                                         |

The first column represents the groups of metabolic disorders potentially identified by newborn screening according to the Newborn Screening ACT Sheets and Algorithms (ACMG ACT Sheets); the second column represents the number of different genes per group of disorders; the third column congestates all occurrences of variants per group; the fourth column represents the combined allele frequencies in the cohort of 320 patients; the last column represents the predicted frequency of affected individuals (q^2; p ~ 1) per 10,000 individuals (a scale of color was used to compare the distribution of frequencies). The last row congestates all groups.
Table 1 (last column) contains the predicted frequency of diseases per group of metabolic disorders.

**Newborn screening diseases**

A total of 40 variant occurrences (19.5% of total occurrences) are harbored by 24 different genes (17.4% of total genes) that are associated with diseases potentially identified by newborn screening. The combined allele frequency of these diseases is 0.0625, while the predicted frequency of affected individuals ($q^2$) is 2.93 per 10,000 people. Table 3 shows the distribution of genes, variants and the estimated frequency of diseases ($q^2$; estimation per 10,000 individuals) per group of metabolic disorders potentially identified by newborn screening.

**Discussion**

We studied the frequencies of carriers for rare metabolic recessive diseases in a cohort of 320 Brazilian individuals. We identified 205 occurrences (average of 0.64 occurrence per patient) of 172 different variants harbored by 138 different genes in 156 patients. We used carrier frequency data and the Hardy–Weinberg and estimated an overall frequency of recessive metabolic diseases to be approximately 0.11% (10.96/10,000 people).

The majority of the variants are harbored by genes associated with multisystemic involvement, which widens clinical and social burden of these conditions. Several of these conditions may manifest with acute crises that often require critical care and may involve seizures (64%), acidosis (32%), hypoglycemia (20%) or encephalopathy (14%). The majority of the diseases may also present with long-term complications, such as growth abnormality (66%), intellectual disability (55%) and other complications addressed in Table 2.

It is noteworthy that almost one fifth (18%) of these genes may present with facial dysmorphisms, which is a clinical characteristic generally regarded as non-suggestive of metabolic diseases by clinicians. We also found unexpectedly high frequency of muscle (80%) and ophthalmologic involvement (70%) in the genes harboring P/LP variants. These latter involvements may require specific clinical measurements and should not be neglected in the follow-up of these patients.

We have used the HPO database to estimate the clinical impact of the genes found in our study and discussed above. This approach, though, presents some limitations: HPO tends to be highly inclusive even for rare manifestations of diseases; this fact may have overestimated the multisystemic impact of the diseases. On the other hand, some diseases may be neglected; for instance, CUBN gene was not associated with any HPO term studied, though it is known that the conditions associated with this gene may present with growth anomaly, intellectual disability, renal involvement, etc. (OMIM#261100).

We have estimated the frequency of rare recessive metabolic diseases to be 10.96 per 10,000 people. Other studies using different approaches have estimated variable frequencies of metabolic disorders: 4.0 (40/100,000) [11], 12.76 (1/784) [12] or 15.6 (66/42,257) [13] per 10,000 people. These estimates, though, are usually challenging especially if they are based on biochemical-based newborn screening programs because several factors may influence the results, including the scope of diseases screened, deaths occurring before a diagnosis is made, ease of access for screening programs, diagnosis confirmation methods, specific regional clinical interests [12]. Additionally, the frequency of these diseases may vary widely in different population groups due to founder effects, endogamy, nonrandom mating, and cultural, religious, social and/or geographical isolation [14]. For example, in endogamous societies this rate can be as high as 8.4% in highly consanguineous populations [15].

Considering the relevant clinical impact of several metabolic diseases, the relatively common combined frequencies and the possibility to improve clinical outcome with early diagnoses, several countries have established newborn screening programs aiming rapid diagnosis and treatment.
for clinically-actionable metabolic diseases. We have used the Newborn Screening ACT Sheets and Algorithms, from ACMG, to map the metabolic diseases potentially identified by newborn screening programs and we estimate the frequency of these diseases to be 2.93 per 10,000 people. Figure 1 helps to understand the magnitude of the estimated frequency of diseases potentially identified by newborn screening in the universe of predicted frequency of all metabolic disorders. If our estimates are accurate, we estimate that 26.6% (2.92/10.96) of patients with metabolic disorders may potentially be detected by neonatal screening. Amino acid studies would have the potential to identify 33.6% (0.98/2.92) of potentially detectable diseases, while acylcarnitine analysis the potential to identify 31.8% (fatty acid oxidation disorders: 0.49/2.92 + organic acidemias: 0.44/2.92) and a small fraction (0.7%; 0.02/2.92) would be detected by assays for galactosemia. An important number of detectable diseases rely on specific tests for lysosomal storage disorders (34.2%; 1.0/2.92) that are generally not widely available in neonatal screening programs.

Our study presents several limitations. The first refers to variant classification and pathogenicity presumption: several variants were classified as likely pathogenic based solely on ACMG criteria (39 variants exclusively found in this study received solely PVS1 and PM2 criteria) but were not definitively proven to be pathogenic (disease-causing); while this fact that may have overestimated the carrier frequencies by falsely classifying some variants as P/LP, we took precautions to standardize variant classification by strictly using ACMG classification criteria uniformly for all sequence variants. On the other hand, we may have excluded variants not detected by NGS, missense hypomorphic alleles that do not result in a clear and recognizable loss of function and several other variants that are not yet recognized to cause disease.

Another important bias that compromises the direct estimation of single variant frequencies or even single monogenic disease frequencies is the limited number of patients. However, we avoided estimations for single diseases and rather used the strategy of studying groups of diseases to estimate frequencies, then diluting the small sample of patients.

Our study did not consist of asymptomatic patients, which would be an ideal group for carrier screening studies (Supplementary Table 1). Precautions were taken, though, to minimize this bias by eliminating P/LP variants associated with autosomal recessive disorders reported in homozygosity or compound heterozygosity (primary findings) from this analysis because our objective was to study carriers for recessive diseases. Therefore, all patients in this study had only monoallelic variants. Nevertheless, we cannot rule out that some of the monoallelic variants studied in this work may have contributed to clinical phenotypes of some patients, for example, if the disease presents digenic or even oligogenic mechanism, an unknown codominant inheritance, or if the other allele harbored a pathogenic or likely pathogenic variant not detected by the NGS methodology (such as intronic or regulatory site variants). Another issue regarding heterozygous carriers is that some monoallelic variants may have clinical impacts and health complications, especially for some genes (including POLG) that may be associated with autosomal dominant forms that have a generally milder clinical impact.

Literature regarding carrier frequencies for autosomal recessive metabolic diseases is limited in Latin America. Most of previously published studies are limited to sole or restricted groups of diseases. Additionally, the majority rely on newborn screening findings without molecular confirmation. We hope that our study encourages other studies that aim to better understand the burden of recessive diseases in our developing countries.

Conclusions

We studied the frequencies of heterozygotes for rare pathogenic/likely pathogenic alleles of recessive metabolic diseases exome sequencing and observed a total of 205 occurrences of 172 different variants in almost half (48.8%) of the 320 patients. Most of these variants are harbored by genes associated with multisystemic involvement, most frequently muscle disease, eye disease and growth anomalies. Using the Hardy–Weinberg equation, we estimated the overall frequency for rare recessive metabolic diseases to be 10.96/10,000 people, while the frequency of metabolic diseases potentially identified by newborn screening was estimated to be 2.93/10,000 people. This study shows the potential utility to evaluate clinical and social burden of metabolic diseases and guide health policies for these conditions.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11033-022-07241-3.

Author contributions All authors provided the conception and design of the study, acquisition of data, analysis and interpretation of data, drafting the article, revised it critically for important intellectual content and final approval of the version to be submitted.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Availability of data and material All data are provided as “Supplementary Material”.

Declarations

Conflict of interest The following authors are employees (received salary and other bonuses) of Fleury Medicina e Saude: CRDCQ, CHC, SFP, APD, CMM. Author CAK declares no financial interests.

Ethical approval This study was performed in line with the principles of the Declaration of Helsinki. This study was granted ethics committee
approval from Fleury Group and Faculdade de Medicina da Universidade de São Paulo (Plataforma Brasil; CAAE# 02617018.3.0000.5474; Fleury# 3.372.339).

Informed consent  Informed consent was obtained from all individual participants included in the study. All authors are aware, consented and approved this publication.

References

1. Morava E, Rahman S, Peters V, Baumgartner MR, Patterson M, Zschocke J (2015) Quo vadis: the re-definition of “inborn metabolic diseases.” J Inherit Metab Dis 38:1003–1006. https://doi.org/10.1007/s10545-015-9893-x
2. Saudubray JM, Garcia-Cazorla À (2018) Inborn errors of metabolism overview: pathophysiology, manifestations, evaluation, and management. Pediatr Clin N Am 65:179–208. https://doi.org/10.1016/j.pcl.2017.11.002
3. ACMG ACT Sheets and Algorithms [Internet]. Bethesda (MD): American College of Medical Genetics and Genomics; 2001-. Newborn Screening ACT Sheets and Algorithms. https://www.ncbi.nlm.nih.gov/books/NBK55827/
4. Ferreira CR, van Karnebeek CDM, Vockley J, Blau N (2019) A proposed nosology of inborn errors of metabolism. Genet Med 21:102–106. https://doi.org/10.1038/s41436-018-0022-8
5. Quaio CRDC, Moreira CM, Novo-Filho GM et al (2020) Diagnostic power and clinical impact of exome sequencing in a cohort of 500 patients with rare diseases. Am J Med Genet C Semin Med Genet 184:955–964. https://doi.org/10.1002/ajmg.c.31860
6. Richards S, Aziz N, Bale S et al (2015) 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 17:405–424. https://doi.org/10.1038/gim.2015.30
7. Quaio CRDC, Chung CH, Perazzio SF et al (2021) Frequency of carriers for rare recessive Mendelian diseases in a Brazilian cohort of 320 patients. Am J Med Genet C Semin Med Genet 187:364–372. https://doi.org/10.1002/ajmg.c.31932
8. Edwards AW (2008) G. H. Hardy (1908) and Hardy-Weinberg equilibrium. Genetics 179:1143–1150. https://doi.org/10.1534/genetics.104.092940
9. Köhler S, Gargano M, Matentzoglu N et al (2021) The human phenotype ontology in 2021. Nucleic Acids Res 49:D1207–D1217. https://doi.org/10.1093/nar/gkaa1043
10. Lee JY, Wasserman WW, Hoffmann GF, van Karnebeek CDM, Blau N (2018) Knowledge base and mini-expert platform for the diagnosis of inborn errors of metabolism. Genet Med 20:151–158. https://doi.org/10.1038/gim.2017.108
11. Applegarth DA, Toone JR, Lowry RB (2000) Incidence of inborn errors of metabolism in British Columbia, 1969–1996. Pediatrics 105:e10. https://doi.org/10.1542/peds.105.1.e10
12. Sanderson S, Green A, Preece MA, Burton H (2006) The incidence of inherited metabolic disorders in the West Midlands, UK. Arch Dis Child 91:896–899. https://doi.org/10.1136/adc.2005.091637
13. Zhang W, Yang Y, Peng W et al (2020) A 7-year report of spectrum of inborn errors of metabolism on full-term and premature infants in a Chinese Neonatal Intensive Care Unit. Front Genet 10:1302. https://doi.org/10.3389/fgene.2019.01302
14. Antonarakis SE (2019) Carrier screening for recessive disorders. Nat Rev Genet 20:549–561. https://doi.org/10.1038/s41576-019-0134-2
15. Keyfi F, Nasserri M, Nayerabadi S, Alaei A, Mokhtariyeh A, Varesteh A (2018) Frequency of inborn errors of metabolism in a Northeastern Iranian sample with high consanguinity rates. Hum Hered 83:71–78. https://doi.org/10.1159/000488876

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.