Congenital disorders of glycosylation: Prevalence, incidence and mutational spectrum in the Polish population

Patryk Lipiński a,1, Anna Bogdańska b,1, Anna Tylki-Szymańska a,∗

a Department of Pediatrics, Nutrition and Metabolic Diseases, The Children’s Memorial Health Institute, Warsaw, Poland
b Department of Biochemistry, Radioimmunology and Experimental Medicine, The Children’s Memorial Health Institute, Warsaw, Poland

A B S T R A C T

Introduction: The incidence and prevalence of congenital disorders of glycosylation (CDG) have not been well established. The aim of the study was to evaluate the prevalence, incidence and genotypes of CDG patients diagnosed during the last 23 years in Poland (1997 – 30th October 2020).

Material and methods: The diagnosis was based on serum Tf IEF which is performed at The Children’s Memorial Health Institute (CMHI) in Warsaw. Based on demographic data, the prevalence of CDG among the Polish population in 2020 as well as the birth prevalence of CDG from 1990 to 2020 were estimated.

Results: 39 patients (from 35 families) were diagnosed with molecularly confirmed CDG, including 17 (44%) patients (from 16 families) with PMM2-CDG. The c.422G>A, p.Arg141His and c.691G>A, p.Val231Met pathogenic missense variants were the most common identified PMM2 variants. Eleven other patients were diagnosed with CDG based on serum Tf IEF analysis only; the molecular analysis is pending. Ten CDG patients died, including 6 with PMM2-CDG, 1 with PGM1-CDG and 1 with DPAGT1-CDG. The prevalence of CDG in the Polish population was estimated at approximately 1 per million while that of PMM2 at 0.4 per million. The annual incidence of CDG was estimated at 0.013 per 100,000 people in 2020.

Conclusions: A low frequency of CDG in our study could be underestimated.

1. Introduction

Congenital disorders of glycosylation (CDG) are genetic defects in the synthesis of glycans and their attachment to proteins and lipids [1,2]. Phosphomannomutase 2 deficiency (PMM2-CDG) is the most common entity; since its first description (1980), more than 130 CDG subtypes have been reported [1,2]. The first-line screening test for N-glycosylation defects with sialic acid deficiency is still the serum transferrin (Tf) isoelectric focusing (IEF); however, normal results do not exclude CDG [3]. Since next-generation sequencing (NGS) became more widely available, an improvement in diagnostics has been observed, with more patients as well as novel subtypes being reported [1,2].

The incidence and prevalence of CDG have not been well established. The aim of the study was to evaluate the prevalence and incidence of CDG in Poland in patients diagnosed in the last 23 years in the CMHI in Warsaw.

2. Material and methods

Since 1995, CDG selective screening based on serum Tf IEF has been performed at our Institute (CMHI) for patients from the entire country. During the years 1995–2020, a total number of 23,183 serum Tf isoform analyses have been performed, while some patients underwent repeat analyses. In this cohort, 2822 samples were also investigated through collaborative initiatives of EUROGLYCAN (the years 2000–2003, 2822 analyses) and EUROGLYCANET (the years 2005–2009, 6098 analyses). Some of the patients have been previously reported [21–25].

According to Statistics of Poland, the population of Poland in 2020 amounted to 38,354,173 people. The demographic data are publicly available and are updated every 6 months; the most recent available dataset is from 30th June 2020 [4].

All CDG patients were enrolled into this study, from the first patient diagnosed in 1997, to the patients diagnosed until 30th October 2020. The prevalence of CDG in the Polish population in 2020 as well as the annual incidence were estimated based on newly diagnosed patients and the number of Polish inhabitants.

3. Results

39 patients (from 35 families) were diagnosed with molecularly confirmed CDG including 17 (44%) patients with PMM2-CDG (Table 1).
Eleven other patients were diagnosed with CDG based on serum Tf IEF analysis only; the molecular analysis is pending.

CDG was diagnosed by serum Tf IEF analysis. However, four patients (3 with ALG1-CDG and 1 with ALG11-CDG) were missed by serum Tf analysis. ALG1-CDG patient with normal serum Tf isoform profile was diagnosed based on array comparative genomic hybridization in which chromosome 16p13.3 deletion involving the ALG1 gene was found. Eleven other patients were diagnosed with CDG based on serum Tf IEF analysis, showing an alteration in the apoC-III isoform profile (increased apoCIII-1, decreased apoCIII-2), indicative of a combined N- and O-glycosylation defect. ATP6AP1-CDG, ATP6V0A2-CDG, and VMA21-CDG were diagnosed by WES.

A detailed characteristics of CDG patients is presented in Table 1. Five patients, including three of them from one family, with CDG-IIx based on serum Tf isoform analysis, showed an alteration in the apoC-III isoform profile (increased apoCIII-1, decreased apoCIII-2), indicative of a combined N- and O-glycosylation defect. ATP6AP1-CDG, ATP6V0A2-CDG, and VMA21-CDG were diagnosed by WES.

A detailed characteristics of CDG patients is presented in Table 1. In the case of 17 PMM2-CDG patients (from 16 families), 15 different PMM2 variants were identified (Supplementary Table 1), including missense (n = 12), frameshift (n = 1), and single-nucleotide variant (n = 2). The most common variants were c.422G > A, p.Arg141His (25%) and c.691G > A, p.Val231Met (21%). All variants were present in a heterozygous state (Supplementary Table 2), the most common genotype was c.691G > A, p.Val231Met/c.422G > A, p.Arg141His (23%).

The mutation spectrum of non-PMM2 CDG is presented in Table 2. Either homozygous or compound heterozygous variants were identified in 17 patients (from 14 families) yielding a total of 18 different variants including missense (n = 11), nonsense (n = 3), and frameshift (n = 2).

10 CDG patients died, including 6 PMM2-CDG, 1 PGM1-CDG, 1 DPAGT1-CDG and 2 ALG1-CDG patients.

The period (1997–2020) prevalence of CDG in the Polish population was estimated at approximately 1 per million while that of PMM2 was at 0.4 per million. The birth prevalence of CDG was estimated as 0.14 per 100,000 live births while that of PMM2-CDG as 0.06 per 100,000 live births.

Table 3 illustrates the number of CDG patients diagnosed annually during the study period. The annual incidence of CDG per 100,000 persons was between 0 and 0.015.

### 4. Discussion

Prevalence and incidence are the two fundamental measures of disease frequency. However, the rarity of the disease often renders those estimations a challenging task. Data regarding the prevalence of CDG originates mostly from isolated reports. On the other hand, there is no information about CDG incidence.

PMM2-CDG was the most common type of CDG identified in our study, similarly as reported in the literature, with the highest mortality observed during the study period [5]. SRDSA3-CDG and ALG13-CDG were the second most frequent types. We did not identify any patient with ALG6-CDG which is the second most frequent type in the literature.

The prevalence of PMM2 could be as high as 1:20,000 [6]. Schollen et al. estimated the frequency of PMM2 based on allele frequencies with ALG6-CDG which is the second most frequent type in the literature.

Table 2 illustrates the number of CDG patients diagnosed annually during the study period. The annual incidence of CDG per 100,000 persons was between 0 and 0.015.

### Table 1
Number of patients in each CDG and serum Tf IEF profile (AR – autosomal recessive).

| CDG type       | Number of patients/families | Diagnosis |
|----------------|----------------------------|-----------|
| PMM2-CDG (AR)  | 17/16                      | CDG type I based on Tf IEF in all (17); 9 patients diagnosed using both molecular and enzyme analysis; 7 patients diagnosed only using molecular analysis |
| ALG13-CDG (X-linked) | 4/4                       | 3 patients diagnosed first using molecular analysis (WES); Tf IEF normal in 1 of them, in another, it was indicative of CDG—I, and in the third one the disialo-Tf isoform was slightly elevated; In another patient, CDG type I based on Tf IEF confirmed by molecular analysis |
| ALG1-CDG (AR)   | 3/3                        | CDG type I based on Tf IEF in all (3); Confirmed by molecular analysis |
| ALG3-CDG (AR)   | 1/1                        | CDG type I based on Tf IEF; Confirmed by molecular analysis |
| MPI-CDG (AR)    | 3/2                        | All diagnosed both using enzyme and molecular analysis |
| PGM1-CDG (AR)   | 1/1                        | Diagnosed both using enzyme and molecular analysis |
| SRDSA3-CDG (AR) | 4/3                        | CDG type I based on Tf IEF in all (4); Confirmed by molecular analysis |
| DPAGT1-CDG (AR) | 1/1                        | CDG type I based on Tf IEF; Confirmed by molecular analysis |
| ATP6AP1-CDG (X-linked) | 3/1                     | CDG type II based on Tf IEF in all (3) |
| ATP6V0A2-CDG (AR) | 1/1                      | CDG type II based on Tf IEF; Confirmed by molecular analysis |
| VMA21-CDG (X-linked) | 1/1                     | CDG type II based on Tf IEF; Confirmed by molecular analysis |
| CDG-Ix          | 7/7                        | CDG type I based on Tf IEF |
| CDG-Ix (X-linked) | 4/7                       | CDG type II based on Tf IEF |

### Table 2
Mutational spectrum in non-PMM2 patients.

| Diagnosis         | Variant | Status   |
|-------------------|---------|----------|
| ALG13-CDG (X-linked) | c.320A > G, p.Asn107Ser, de novo | Heterozygous |
|                    | c.280A > G, p.Lys94Glu, de novo | Heterozygous |
| ALG1-CDG (AR)      | c.773C > T, p.Ser258Leu | Heterozygous |
|                   | c.1182C > G, p.Phe394Leu | Heterozygous |
| MPI-CDG (AR)       | c.1193 T > C, p.Ile398Thr | Heterozygous |
|                   | c.656G > A, p.Arg219Glu | Heterozygous |
|                   | c.748G > A, p.Gly250Ser | Heterozygous |
| PGM1-CDG (AR)      | c.988G > C, p.Gly330Arg | Heterozygous |
|                   | c.1129G > A, p.Glu377Asp | Heterozygous |
| SRDSA3-CDG (AR)    | c.292_293del, p.Leu98ValX121 | Homozygous |
|                   | c.424C > T, p.Arg142X | Homozygous |
|                   | c.498C > A, p.Tyr163Ter | Homozygous |
| DPAGT1-CDG (AR)    | c.1117C > G, p.Pro373Ala | Heterozygous |
|                   | c.1197T > A, p.Tyr399X | Heterozygous |
| ATP6AP1-CDG (X-linked) | c.1284G > A, p.Met428Ile | Homozygous |
| ATP6V0A2-CDG (AR)  | c.2015T > A, p.Leu672X | Homozygous |
|                   | c.1304G, p.N436X | Heterozygous |
| VMA21-CDG (X-linked) | c.188A > G, p.Asn63Gly | Heterozygous |
The number of annually newly diagnosed CDG patients (n.a. – not analyzed).

| Year | Number of newly diagnosed patients | Number of living patients | Annual incidence per 100,000 |
|------|-----------------------------------|---------------------------|-------------------------------|
| 1997 | 1                                 | 1                         | n.a.                          |
| 1998 | 0                                 | 1                         | n.a.                          |
| 1999 | 1                                 | 2                         | n.a.                          |
| 2000 | 0                                 | 2                         | 0                             |
| 2001 | 3                                 | 4                         | 0.008                         |
| 2002 | 3                                 | 6                         | 0.008                         |
| 2003 | 3                                 | 9                         | 0.008                         |
| 2004 | 2                                 | 10                        | 0.005                         |
| 2005 | 1                                 | 11                        | 0.002                         |
| 2006 | 2                                 | 12                        | 0.005                         |
| 2007 | 0                                 | 12                        | 0                             |
| 2008 | 2                                 | 13                        | 0.005                         |
| 2009 | 3                                 | 15                        | 0.008                         |
| 2010 | 4                                 | 19                        | 0.01                          |
| 2011 | 1                                 | 19                        | 0.002                         |
| 2012 | 0                                 | 19                        | 0                             |
| 2013 | 1                                 | 20                        | 0.002                         |
| 2014 | 1                                 | 21                        | 0.002                         |
| 2015 | 6                                 | 26                        | 0.015                         |
| 2016 | 4                                 | 30                        | 0.01                          |
| 2017 | 0                                 | 29                        | 0                             |
| 2018 | 1                                 | 29                        | 0.002                         |
| 2019 | 5                                 | 34                        | 0.013                         |
| 2020 | 5                                 | 39                        | 0.013                         |

5. Conclusions

The prevalence of CDG in the Polish population was estimated in our study at approximately 1 per million while that of PMM2 at 0.4 per million. This fact depends on the number of patients that have been screened and then could be extrapolated to the national disease incidence. Serum Tf IEF is traditionally used in the selective screening of inborn errors of metabolism. Some patients were also investigated through collaborative initiatives of EUROGLYCAN and EUROGLYCANET (population screening).

PMM2-CDG has the best-defined phenotype so that clinicians are highly aware of this disease. This fact could correspond with a relatively high prevalence of PMM2-CDG among other forms of CDG. The analysis of serum Tf isofoms is still the method of choice for CDG diagnosis but only for N-glycosylation defects with sialic acid deficiency [3]. Next-generation sequencing (NGS) technology including targeted gene panels, whole-exome sequencing (WES) or even whole-genome sequencing (WGS) is necessary to diagnose a specific CDG and permits the detection of novel CDG [1,2]. Given the high clinical heterogeneity of CDG and the fact that NGS is not routinely available we speculate that a low frequency of CDG in our study could be underestimated.

Magalhães et al. have recently published results of an observational and retrospective study of individuals investigated by serum Tf IEF in a laboratory in southern Brazil, from 2008 to 2017 [19]. A total number of 1546 individuals underwent serum Tf IEF, of whom only four individuals were molecularly diagnosed with CDG.

The PMM2 gene (NM_000303.3) encodes the PMM2 protein (EC 3.2.1.28). Hundred thirty pathogenic PMM2 variants have been reported in Human Gene Mutation Database (HGMD Professional 2020.3), the large majority (85%) being missense variants [9]. Most PMM2-CDG reported patients were compound heterozygotes; the c.422G[A, p.Val231Met was most frequently found variant [6,7,10,11]. Homozygosity for c.422G[A, p.Arg141His has been shown to be absent since it is probably lethal [10]. In our study, the c.422G[A, p.Arg141His and c.691G[A, p.Val231Met pathogenic missense variants were the most common identified. Compound heterozygotes for c.422G[A, p.Arg141His and c.691G[A, p.Val231Met were reported in the literature to be associated with the moderate to severe phenotype [12]. The pathogenic variant p.Val231Met was reported to be associated with high early mortality and severe multiorgan insufficiency [12,13,16]. In our study, 4 out of 6 patients who died were heterozygous for the c.691G[A, p.Val231Met variant but all of them were heterozygous for the other deletion/insertion variant. Considering the effect on residual activity of PMM2, mutations classified as severe include p.Arg123Glu, p.Arg141His, p.Phe157Ser, p.Pro184Thr, p.Phe207Ser and p.Asp209Gly, while mild mutations include p.Leu32Arg, p.Val44Ala, p.Asp65Tyr, p.Pro113Leu, p.Thr118Ser, p.Thr237Met and p.Cys241Ser [14,16]. In our population cohort, we found 1 patient to be heterozygous for c.357C[A, p.Phe119Leu, which was reported as the second most common mutation among the South-Scandinavian population (43% allele frequency in Danish patients) [11]. We did not find c.415G[A, p.Glu139Lys, the most prevalent variant among French patients [15] nor c.95TA[C, p.Leu32Arg, the second most common mutation in the Italian population (16% of disease alleles). The c.95TA[C, p.Leu32Arg mutation was reported in 12 out of 37 Italian patients, all of them presenting with a mild neurological phenotype (preserved ambulatory ability and autonomy) [16]. Comparing with the study of Perez-Cerda et al. reporting 71 Spanish PMM2-CDG patients gathered during the last 20 years, the frequency of c.710C[T, p.Thr237Arg and c.338C[T, p.Pro113Leu were similar [17]. Regarding the mutational spectrum of PMM2-CDG in the Portuguese cohort reported by Quelhas et al., a striking similarity with Spanish population was found [18].

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Funding sources

The study was funded by the Children’s Memorial Health Institute intramural grant S190/2020.

Ethics approval and consent to participate

Ethical approval was obtained from the Children’s Memorial Health Institute Bioethical Committee, num 23/KBE/2020, Warsaw, Poland. Informed consent was obtained from all included patients.

Declaration of Competing Interest

All authors certify that they have NO affiliations with or involvement in any organization or entity with any financial or non-financial interest in the subject matter or materials discussed in this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ymgmr.2021.100726.
