Clinical and Basic Investigations into Congenital Disorders of Glycosylation

Study Title: Clinical and Basic Investigations into Congenital Disorders of Glycosylation

Principal Investigator: Eva Morava-Kozicz, M.D., Ph.D.

Protocol version number and date: Version 6 9/29/2020

Research Question and Aims

Objective: Define natural history, validate patient reported outcome and share knowledge on congenital disorders of glycosylation (CDG).

Aims:
Aim 1: Establish the prevalence and severity of specific morbid indicators of disease severity such as specific organ system involvement, degree of cognitive disability, and case-fatality associated with various CDG, and establish a dynamic platform to effectively disperse clinically relevant findings to families, non-expert clinicians and researchers, as well as provide a verified method to link these individuals to experts in CDG.

Aim 2: To validate the CDG rating scale and patient reported outcome /quality of life measures for CDG.

Background (Include relevant experience, gaps in current knowledge, preliminary data, etc.):
This clinical research study will determine the natural history of congenital disorders of glycosylation (CDG), delineate the spectrum of clinical features, develop and validate patient reported outcomes and disperse verified
clinical information to families and non-expert clinicians. In collaboration with the biomarker clinical research study, this study will also aid in biomarker development and correlation with clinical features. The following rationale for this research underscores its significance for CDG.

More than half of human proteins are glycosylated.\(^1\) The glycome defined as the sugar chains (glycans) that an organism makes, is \(\sim 10^2\)–\(10^4\) times larger than the proteome.\(^2\) Given this complexity, it is not surprising that about 2% of our genes encode for proteins known to participate in glycosylation reactions.\(^2\,3\)

A congenital disorder of glycosylation is a monogenic disorder in which proper modification of an organic molecule with a sugar or sugar structure is disrupted. PMM2-CDG, the first CDG that was clinically described in 1980 had its underlying genetic and enzymatic defect determined in the mid-late 1990s. The field has been expanding rapidly with the advent of genomic sequencing, with many novel CDG identified in the last few years.\(^3\) Although \(\sim 130\) genetic disorders are associated with a dysfunction of one of these proteins, CDG are rare both individually and as a group. The most common type of CDG involves N-glycosylation defects (Figure 1).

Figure 1: Overview of N-glycosylation, a step-by-step process. Any defect in these steps could lead to a CDG. (The sugars are depicted with standard colored symbols).

Worldwide, PMM2-CDG (CDG-Ia) remains the most common CDG with \(\sim 900\) cases reported.\(^3\) PMM2-CDG prevalence may be as high as 1 in 20,000 individuals in some populations.\(^4\) The next most common subtype is ALG6-CDG (CDG-Ic), with 50 cases.\(^5\) Congenital disorders of galactosylation are one of the largest subgroup, including more then 70 reported cases of PGM1-CDG and SLC35A2-CDG (Radenkovic, in press). PMM2-CDG and galactosylation defects are typically detectable in blood and show an abnormal N-glycan pattern.

Since so many biologic functions depend upon correct protein and lipid glycosylation, clinical CDG manifestations are widely variable within and among subtypes and across the lifespan. The clinical findings in affected individuals reflect glycosylation's role in the embryology and biology of many organ systems. Clinical findings may include developmental delay/intellectual disability with abnormalities in brain MRI, hypotonia, failure to thrive or growth impairment, hepatopathy, coagulopathy, immunodeficiency, protein losing enteropathy, osteopenia, skeletal dysplasias, renal abnormalities, functional and/or anatomic eye involvement, endocrinopathies, and peripheral neuropathy. Some CDG present with only tissue-specific symptoms such as SEC23B-dyserythropoietic anemia or Leukocyte Adhesion Deficiency Type II.\(^6\)

Of the known 137 glycan-synthesis disorders we have potential therapies for only 8 disorders and not all current therapies are fully effective. Clinical trials have not validated these treatments for safety and efficacy. MPI-CDG has been treated with oral mannose.\(^7\) In SLC35C1-CDG, some patients respond to oral fucose supplementation. This treatment is only effective for the typical recurrent infections with hyperleukocytosis and does not correct the neurological issues of these patients.\(^6\) PIGM-CDG, the mutation described in the only published case so far, causes histone hypoacetylation at the PIGM promoter. Butyrate controls the seizures.\(^8\) PGM1, a critical enzyme for the
conversion of glucose-1-phosphate to glucose-6-phosphate, is at the crossroads of the glyogen, glucose and galactose pathway. **PGM1 deficiency is treatable with D-galactose**, alleviating the hypoglycemia, coagulopathy, and endocrinopathy seen in this disorder, although not the cardiomyopathy. The glycosylation defect CAD-CDG has been treated with uridine. Magnesium can improve MAGT1-CDG. Manganese and D-galactose have been used to treat TMEM165-CDG, whereas **SLC35A2-CDG has been treated with D-galactose**. From the group of disorders of de-glycosylation preclinical trials suggested positive effect of **GlcNAc treatment in NGLY1 deficiency**.

Given the rarity and novelty of CDG, the geographic spread of affected individuals, the broad clinical spectrum, the limited retrospective attempts at natural history studies (focused on one disorder or limited to one site), there is a significant **knowledge gap** regarding the progression and clinical variation in this group of disorders. Even for PMM2-CDG, preliminary data from an unpublished study emphasized that there is a need for prospective natural history studies to validate and expand on retrospective, limited results. This gap impedes the development of novel targeted therapeutics since there is no clear outcome measure for clinical trials, and it will be difficult to determine efficacy of a therapy over natural history.

### Study Design and Methods

**Methods:**

**Enrollment:**

We will recruit and enroll patients with CDG in this study evaluating clinical variation and natural history when a patient is being seen as part of routine clinical care. We will invite eligible subjects to participate in the study to consent, enroll, and collect baseline data. We may also send consent forms in the mail to participants to consent or reconsent. All subjects for this study will be provided a consent form describing this study and providing sufficient information for subjects to make an informed decision about their participation in this study. This consent form will be submitted with the protocol for review and approval by the IRB for the study. The formal consent of a subject, using the IRB-approved consent form, must be obtained before that subject undergoes any study procedure. The consent form must be signed by the subject or legally acceptable surrogate, and the Investigator designated research professional obtaining the consent.

**Study Activities:**

Participants will then be seen at the study sites for annual follow up visits to undergo assessments, with a variable window of 2 months. We will collect data from local providers, collect relevant medical data, and complete study forms at annual follow up visits. We will strive to minimize duplicate tests. The clinical assessment data will be obtained with each visit according to the specific CDG type and diagnosis and appropriate to standard of care. Data will be collected from the patient’s medical records and through surveys for the patient and assessments given by the provider. Included in this data set is a food record recorded by the patient. This may be given by paper, or electronically (through ASA24 program). If given electronically, the patient will be sent a password via email to log into the site (this will only be sent to patients/guardians on the study who are not infants/children consuming infant/pediatric formulas and are eating normal diets). The site will ask for the patient to select which food has been consumed over a period of time and automatically calculate nutrition data that can be used for analysis. For those patients who are consuming pediatric/infant formula, we will utilize Metabolic Pro program, since formula information is not available in the ASA24 platform. Metabolic Pro will also be used for dietary calculations when food records are recorded via paper instead of electronically.
Surveys and evaluations to complete at each visit include the Nijmegen CDG rating scale (NPCRS; completed by physician through medical interview process), the PROMIS questionnaire (Patient Reported Outcomes survey, or PROs, completed by patient/family), physical examination (completed by physician), vitals (according to standard of care, completed by clinic staff), food record (completed by patient/family, on paper or electronically), and goal attainment score (three goals on a five point scale, completed by physician and family members in unison).

A list of retrospective data to be collected from the medical record is attached to the application. Optional blood, stool, and urine samples will be collected for research at each annual visit (once per year over the course of 5 years of follow up). If a participant is unable to provide urine, stool, or blood samples due to age or other logistical difficulty, we will forego research sample collection and rely purely on clinical/research data collected over the course of the study. If any clinical residual samples (urine, blood, or stool) are available for these participants unable to provide fresh research collections, we may use these instead. These clinical residuals would come from the Biochemical Genetics Lab (BGL) at Mayo or from participating sites’ laboratories, if any sample remains. Skin biopsies will be only taken for diagnostic purposes. Leftover tissue material from clinical diagnostics will be used for biomarker discovery. If a subject doesn’t complete regular follow up appointments, we will reach out to them via phone or email to obtain follow up information and complete questionnaires to ensure we are able to obtain accurate follow up information. Participants will be offered up to $700 reimbursement to cover lodging and travel expenses, per visit at all sites except for the University of Minnesota Masonic Children’s Hospital. University of Minnesota will offer $75 per visit to each participant as remuneration and will not offer travel reimbursement.

**Accrual:**

Our goal is to enroll 150 patients of all ages and from all CDG types over 5 years in the natural history study. At least 50 of the 150 enrolled subjects will have a diagnosis of PMM2-CDG. There will be no exclusion based on sex, ethnicity or socio-economic status. The sample size of 100 was based on a count of the number of affected individuals currently clinically followed and current patient census taken by CDG CARE in August of 2018.

**Participating Sites:**

We will collaborate with other sites on this study with Mayo Clinic acting as the IRB of record. The target accrual in total will remain at 150 participants. The sites below will be included in the study:

1. Children’s Hospital of Pittsburgh at University of Pittsburgh Medical Center – PI is Dr. Gerard Vockley
2. Seattle Children’s Hospital – PI is Dr. Christina Lam
3. Tulane University Medical School – PI is Dr. Hans Andersson
4. Baylor College of Medicine – PI is Dr. Fernando Scaglia
5. University of Minnesota Masonic Children’s Hospital – PI is Dr. Kyriakie Sarafoglou
6. Children’s Hospital of Philadelphia – PI is Dr. Andrew Edmondson
7. Boston Children’s Hospital – PI is Dr. Gerard Berry
8. Children’s Hospital of New Orleans Louisiana (CHNOLA) – PI is Dr. Hans Andersson

The sites listed below will collaborate on de-identified data and samples, but will not be recruiting patients:
1. NIH (Dr. William Gahl)
2. Sanford Burnham Prebys Institute for Medical Discovery (Dr. Hudson Freeze)
3. University of Alabama (Dr. Matt Might)
4. University of Utah (Dr. Nicola Longo)

Children’s Hospital of Philadelphia and Mayo Clinic will also collaborate on de-identified data and samples, but will be enrolling patients.

We will collaborate with two patient advocacy groups (PAGs) on this study, in the form of sharing de-identified relevant research data. These two PAGs are CDG Care (Colorado Springs, CO) and CDG and Allies – PPAIN (Lisbon, Portugal).

**Data:**

Data collected for this study will be analyzed and stored at the RDCRN funded Data Management and Coordinating Center. After the study is completed, the de-identified, archived data will be transmitted to and stored at the Mayo Clinic, under the supervision of Dr. Eva Morava-Kózicz, for use by other researchers including those outside of the study.

Data collection is the responsibility of the study staff at the site under the supervision of the site Primary Investigator. The Investigator is responsible for ensuring the accuracy, completeness, legibility, and timeliness of the data reported.

All source documents should be completed in a neat, legible manner to ensure accurate interpretation of data. When making changes or corrections, cross out the original entry with a single line, and initial and date the change. **DO NOT ERASE, OVERWRITE, OR USE CORRECTION FLUID OR TAPE ON THE ORIGINAL.**

The DMCC will develop the data capture system using a web-based data collection system, REDCap, as the primary source of data entry and storage. REDCap is a software toolset and workflow methodology for electronic collection and management of research and clinical trial data developed by Vanderbilt University, with collaboration from a consortium of institutional partners including the University of Cincinnati Academic Health Center. The DMCC will develop, test and maintain the REDCap data entry system, the data management plan, data quality checks and query management, and preparation of the data for analysis.

The REDCap system provides a secure, web-based application that is flexible and provides: 1) an intuitive interface for users to enter data and have real time validation rules (with automated data type and range checks) at the time of entry; 2) HIPAA-compliant and 21 CFR Part 11-ready audit trails for tracking page views, data manipulation and export procedures; 3) record locking and electronic signature functions; 4) fine grained control of user rights to view and manipulate data, and tool to sequester data access for multiple sites; 5) a report builder for reporting, monitoring and querying patient records; and 6) automated export procedures for seamless data downloads to common statistical packages (SPSS, SAS, Stata, R/S-Plus).

Quality control procedures will be implemented beginning with the data entry system and data quality control checks that will be run on the database will be generated. Any missing data or data anomalies will be communicated to the site(s) for clarification/resolution.

**Samples:**
With the participant’s approval and as approved by local IRBs, de-identified biological samples will be stored at the Mayo Clinic with the same goal as the sharing of data. These samples could be used for research into the causes of CDG, its complications and other conditions for which individuals with CDG are at increased risk, and to improve treatment. Samples may also be utilized in unrelated research with consent. The other institutions involved will also be provided with a code-link that will allow linking the biological specimens with the phenotypic data from each participant, maintaining the masking of the identity of the participant.

**Regulatory:**

A protocol deviation is any noncompliance with the human subject research protocol, GCP, or manual of procedures requirements. The noncompliance may be either on the part of the participant, the Investigator, or the study site staff. As a result of deviations, corrective actions are to be developed by the site and implemented promptly. It is the responsibility of the site to use continuous vigilance to identify and report deviations within 5 working days of identification of the protocol deviation, or within 5 working days of the scheduled protocol-required activity. All deviations must be addressed in study source documents, reported to Mayo Clinic IRB. Protocol deviations must be sent to the local IRB per their guidelines. The site Primary Investigator/study staff is responsible for knowing and adhering to their IRB requirements. Further details about the handling of protocol deviations will be included in the MOP.

In compliance with Protection of Human Subjects regulations, (45CFR Part 46), records related to the conduct of this trial, including but not limited to source documentation, case report forms, informed consent forms, essential study documentation, and documentation of IRB activities, will be retained by the Investigator for a period of 3 years following the official close of the study. Such records may be preserved in hardcopy, electronic or other media form and must be accessible for inspection and copying by authorized representatives of HHS, NIH, the study sponsor and/or their representatives at reasonable times and in a reasonable manner. These documents should be retained for a longer period, however, if required by local regulations. No records will be destroyed without the written consent of the sponsor, if applicable. It is the responsibility of the sponsor to inform the Investigator when these documents no longer need to be retained.

This study will comply with the NIH Public Access Policy, which ensures that the public has access to the published results of NIH funded research. It requires scientists to submit final peer-reviewed journal manuscripts that arise from NIH funds to the digital archive PubMed Central upon acceptance for publication.

This study will be conducted in full conformity with Regulations for the Protection of Human Subjects of Research codified in 45 CFR Part 46, 21 CFR Part 50, 21 CFR Part 56, and/or the ICH E6.

This protocol and any amendments will be submitted to a properly constituted independent Institutional Review Board (IRB), in agreement with local legal prescriptions, for formal approval of the study conduct. The decision of the IRB concerning the conduct of the study will be made in writing to the Investigator and a copy of this decision will be provided to the sponsor before commencement of this study. The Investigator should provide a list of IRB members and their affiliate to the sponsor upon request.
Confidentiality:

Participant confidentiality is strictly held in trust by the participating investigators, their staff, and the sponsor(s) and their agents. This confidentiality is extended to cover testing of biological samples and genetic tests in addition to the clinical information relating to participants. Therefore, the study protocol, documentation, data, and all other information generated will be held in strict confidence. No information concerning the study or the data will be released to any unauthorized third party without prior written approval of the sponsor. This study is funded by the NIH, and as such, has a Certificate of Confidentiality.

Study Monitoring:

Following written standard operating procedures, the monitors will verify that the human subject research is conducted, and data are generated, documented (recorded), and reported in compliance with the protocol, Good Clinical Practice, and the applicable regulatory requirements (e.g., Good Laboratory Practices (GLP), Good Manufacturing Practices (GMP)).

The investigational site will provide direct access to all study related sites, source data/documents, and reports for the purpose of monitoring and auditing by the NIH, and inspection by local and regulatory authorities.

Subject Information

Target accrual: 150 total: including a minimum of 50 subjects diagnosed with PMM2-CDG

Subject population (children, adults, groups): all ages

Inclusion Criteria: All individuals with a genetically, enzymatically, or molecularly confirmed diagnosis of CDG or NGLY1 deficiency, or individuals whose laboratory values are highly suggestive of CDG, but who do not have a verifiable diagnosis possibly due to difficulty with clinical molecular testing will be invited to enroll. Or those without confirmatory testing, they will be enrolled as “CDG highly likely/diagnosis pending” if the site PI and study PI concur. These individuals should have their diagnosis confirmed within one year of enrollment. If not confirmed they will be removed from the study. The study requires a minimum number of 50 patients with the diagnosis PMM2-CDG.

Exclusion Criteria: none

Biospecimens

Collection of blood samples.
a. **From healthy, non-pregnant, adult subjects who weigh at least 110 pounds.** For a minimal risk application, the amount of blood drawn from these subjects may not exceed 550ml in an 8 week period and collection may not occur more frequently than 2 times per week.
   - Volume per blood draw: 30ml
   - Frequency of blood draw (e.g. single draw, time(s) per week, per year, etc.) annually (up to 5 visits)

b. **From other adults and children considering age, weight, and health of subject.** For a minimal risk application, the amount of blood drawn from these subjects may not exceed the lesser of 50 ml or 3 ml per kg in an 8 week period, and collection may not occur more frequently than 2 times per week.
   - Volume per blood draw: up to 30 ml max, will draw minimum amounts not to exceed 3 mls/kg of weight. This will be calculated for each pediatric patient
   - Frequency of blood draw (e.g. single draw, time(s) per week, per year, etc.) annually (up to 5 visits)

Prospective collection of biological specimens other than blood: random urine, random stool at each annual visit (once per year over 5 year follow up)

---

**Review of medical records, images, specimens**

Check all that apply (data includes medical records, images, specimens).

☐ Only data that exists before the IRB submission date will be collected.

**Date Range for Specimens and/or Review of Medical Records:**
Examples: 01/01/1999 through 12/31/2015, or all records through mm/dd/yyyy.

Note: The Date Range must include the period for collection of baseline data, as well as follow up data, if applicable.

☒ The study involves data that exist at the time of IRB submission and data that will be generated after IRB submission. Include this activity in the **Methods** section.

Examples
- The study plans to conduct a retrospective chart review and ask subjects to complete a questionnaire.
- The study plans to include subjects previously diagnosed with a specific disease and add newly diagnosed subjects in the future.
The study will use data that have been collected under another IRB protocol. Include in the Methods section and enter the IRB number from which the research material will be obtained. When appropriate, note when subjects have provided consent for future use of their data and/or specimens as described in this protocol.

Enter one IRB number per line, add more lines as needed

- Data
- Specimens
- Data & Specimens

Data Analysis Plan: The study will be used to estimate the prevalence and distribution of scores and the precision of these estimates based on the specified sample size was assessed. For a binary variable (e.g. likelihood of moderate or severe for a given component of the severity score) the expected size of the confidence interval is a function of the sample size and true underlying prevalence.

The extensive clinical and laboratory data collected in this study will be recorded on electronic forms at each site and submitted to DMCC. Almost all of the clinical exams and laboratory studies will provide either a numerical score or categorical score for phenotypic severity, thereby allowing statistical analysis looking for longitudinal changes and exploring potential biomarkers that correlate with disease. We will utilize the statistical expertise of the DMCC for data analysis. The DMCC database will also be able to house imaging studies and genomic data collected from medical records in a de-identified way.

Clinical and laboratory measures, as well as demographics and age of onset, will be summarized using descriptive statistics. Principal component analysis (PCA) will be applied to summarize multivariate measurements. Baseline association between two factors of interest will be assessed using a general linear regression for a continuous response variable and logistic regression for a categorized response variable, after adjusting for potential confounders.

With the data collected over time, the average trend with age will be estimated using mixed effects models to account for the correlation of repeated measures from a subject. The model will include fixed and random effects for the intercept and slope (age). Maximum likelihood estimates (MLE) will be obtained for regression coefficients for fixed effects and restricted MLE for the variance component. The association of a factor of interest with a response variable will be examined using mixed effects models to account for the subject specific change. For a time constant factor, the model will include the factor, age and an interaction term between the factor and age as fixed effects. The model will also include random effects for the intercept and slope (time). For a time varying factor, the model may include baseline value of the factor and the change at the next visit as fixed effects to estimate cross-sectional and longitudinal effects simultaneously.
We will compare results of PROs to NPCRS subdomain scores, physical exam findings, and laboratory findings that should be reflected in that tool to assess the validity of each PRO measure. Partial least squares (PLS) regression analysis will be utilized to identify patterns of association between the PROs and the clinical scores. The PLS analysis and significant weights may reveal novel composite scores for CDG patients. Changes in PRO scores over time will be tested for association with Quality of Life questions using random effects generalized linear regression models incorporating a random intercept for each participant. We will also compare subsequent scores on the PROs from the same subject to determine reliability of the PRO if there has been no significant interval change in the clinical status. Reliability will be assessed by the weighted Kappa to account for the ordinal variables.

Expected Outcomes.

References:


8. He M, Chen J, Xiao, et al. Increased sensitivity and specificity of plasma protein N-glycane profiling for the


