The inborn errors of metabolism information system: A project of the Region 4 Genetics Collaborative Priority 2 Workgroup

Susan A. Berry, MD, Anne M. Jurek, PhD, Carolyn Anderson, PNP, MSPH, Kristi Bentler, MS, RN, CNP, and the Region 4 Genetics Collaborative Priority 2 Workgroup

Abstract: The Region 4 Genetics Collaborative has brought together metabolic clinicians and follow-up specialists from state departments of health in the region (Illinois, Indiana, Kentucky, Michigan, Minnesota, Ohio, and Wisconsin) in a workgroup to create a dynamic registry, the Inborn Errors of Metabolism Information System, to facilitate gathering information about long-term follow-up for individuals identified by newborn blood spot screening. With the concept that by developing a core series of agreed-on data elements and general treatment strategies, differences in treatment plans could yield evidence about optimal treatment choices, data elements for initial intake, and interval follow-up were selected based on a paradigm condition, medium-chain acyl-CoA dehydrogenase deficiency. Demographic elements that will be used as a common data set for all conditions were identified along with condition-specific elements and general information to be obtained at intervals. Subjects were enrolled after obtaining prospective informed consent; data entry began in January 2007. Additional conditions have had data sets defined and data entry initiated; 21 disorders as of July 2009. Web-based data entry has been employed using DocSite® as the platform for data entry. With continued collaboration among members of the workgroup, we hope to extend the intellectual questions that can be explored using this data, expand the spectrum of the registry and number of patients engaged, and integrate the registry into additional domains. ***Genet Med*** **2010:12(12):S215–S219.**

Key Words: *newborn blood spot screening, long-term follow-up, inborn errors of metabolism*

Comprehensive newborn screening (NBS) using dried blood spots has become a national priority. This essential public health measure is designed to improve outcomes for children and save lives of affected infants. NBS is more than the event of completing the screening test. It should be designed as a comprehensive program that encompasses all aspects of such testing from obtaining the blood spot through provision of comprehensive care and treatment for the affected child. This necessitates collaboration between those who undertake screening, those who provide short-term testing and interpretation, and clinical service providers, primarily metabolic specialists, who convey information, confirm diagnoses, and provide long-term care in partnership with primary care providers in medical homes. This necessitates complex data sharing at many levels. Considerable effort has been devoted by public health laboratories to optimize performance of testing and initial interpretation of results. However, designing strategies for comprehensive monitoring of long-term outcomes after NBS has been a more difficult task, and the best practices for treating identified children based on evidence are not established. The Health Resources and Services Administration (HRSA)-funded Region 4 Genetics Collaborative has brought together metabolic disease clinicians and representatives of state public health departments to address these needs. HRSA provided support for two larger scale projects in Region 4: to facilitate improvements in laboratory performance and begin the foundation of long-term follow-up data collection. The first project has provided substantial and detailed information to investigators and laboratories all over the world with respect to improvement of laboratory standards and interpretation of NBS. The second project, described in this study, is our effort to provide a pilot strategy for data collection for long-term NBS follow-up.

Within the Region 4 Collaborative (Illinois, Indiana, Kentucky, Michigan, Minnesota, Ohio, and Wisconsin), all seven states screen newborns using tandem mass spectrometry (MS/MS) to identify a number of rare, serious inborn errors of metabolism (IBEM). NBS by MS/MS is a relatively new technology first implemented by States in 1998. It has ultimately been added to NBS programs across the nation. Region 4 screens approximately 740,000 babies per year by MS/MS, resulting in an estimate of 265 cases confirmed with a metabolic disorder each year assuming an incidence of approximately 1:2,800 for all IBEM combined. Although long-term follow-up is critical for monitoring health outcomes and evaluating the effectiveness of NBS, standards of clinical care for most screened conditions have never been subjected to evidence-based study. More information about outcomes for these disorders is essential to a better understanding of the natural history of the conditions and development of best practice models for treatment. Over time, mechanisms for storing information about the progress of children identified by NBS will build the foundation for evidence-based medical practice and care for rare disorders ascertained through NBS because they will provide data to support treatment decisions based on larger cohorts of affected children than can be seen by an individual practitioner or specialty center. With the collaboration of multiple states over time, long-term follow-up databases will have the power to provide a foundation for evidence-based clinical practice and care for rare disorders ascertained through NBS.

To meet this challenge, in 2005, the Long-term Follow-up and Clinical Outcomes Workgroup of the Region 4 Genetics Collaborative cooperated to focus on both the initial diagnostic phase and long-term follow-up component of NBS by creating a disease registry to document a paradigm disorder (medium-chain acyl-CoA dehydrogenase deficiency [MCADD]) that would permit collection of accurate clinical data that could be used to assess outcomes and define the prospective history of MCADD and later a variety of other IBEM. Made up of clinical

From the Region 4 Genetics Collaborative, Michigan Public Health Institute, Okemos, Michigan.

Susan A. Berry, Department of Pediatrics, University of Minnesota, 420 Delaware St SE MMC 75, Minneapolis, MN 55455. E-mail: berry002@umn.edu.

Disclosure: The authors declare no conflict of interest.

DOI: 10.1097/GIM.0b013e3181fe5d23

genetic-metabolic disease specialists, state Departments of Health staff, parents, epidemiologists, and care coordinators, the group agreed to start with this specific condition because it is relatively common. Moreover, most treatment centers had protocols in place for management, permitting comparison of treatment plans for key data elements and differences in management, with the rationale that planning for a single disorder could be readily generalized to other conditions. Our concept is that by developing a core series of agreed-on data elements and general treatment strategies, examination of the differences in treatment plans could yield evidence about optimal treatment choices.

CREATING THE IBEM INFORMATION SYSTEM

Metabolic disease clinicians from each state in the region met to establish a collaborative effort. Three fundamental aspects of the process emerged: agreements about data sharing, agreements about initial data elements to form the framework for subsequent study, and an agreement that data would be collected after prospective informed consent was obtained. This latter decision fundamentally shapes the nature of the data as it acknowledges the premise that we commit to research endeavors while accepting that by its nature ascertainment may not be complete as some potential subjects may choose not to participate.

DATA USE AND SHARING

The workgroup agreed from the beginning that all who contributed would share in the ability to access the information and all would receive credit for their contributions. We agreed that together we would make conscious decisions about the use of the information and access to it. Subsequently, the group met again and formulated a formal strategy for access to the information. The principles in governing data access are described in Table 1. Two levels of data access were defined: self-contained data, necessitating only access to the data stored in the data set and expanded data, providing access to the cohort to recruit patients for additional projects.

Table 1 Participation agreement for research and use of materials

1. All participating centers that are entering data will have access to nonprotected health information; any participant may propose a project for consideration.
2. All publications will credit both this Priority 2 project and the Region 4 Genetics Collaborative partners as a group. The lead author/presenter will be the person(s) who does the majority of work to initiate the project and who prepares the first draft. The workgroup will generally agree on the work plan at the initiation of the project, anticipating designation of authorship at this point of effort, with the understanding that this may change as the work progresses. The fundamental concept is planned expectations for allocation of credit for work performed.
3. Individual clinical projects will be reviewed/approved through a scientific advisory group selected from Region 4 and partner clinician participants, and will then be brought before the entire Region 4 IBEM-IS workgroup for final approval.
4. Two primary types of access to the data set for research studies are anticipated:
   a. IBEM-IS data-mining only (self-contained data)
   b. Access to the IBEM-IS cohort of subjects who have provided informed consent for recontact to recruit patients for additional projects (expanded data)
6. The Region 4 Project Epidemiologist accesses the database for data analysis, downloading de-identified data for projects after they have been reviewed using the method previously described and have received workgroup endorsement. The Region 4 Project Epidemiologist will provide special reports summarizing available patients who have agreed to be contacted for further research as needed for expanded data projects endorsed by the workgroup.

DEFINING THE DATA ELEMENTS

The concept used with respect to decisions about inclusion of data elements was to develop a matrix of required information for long-term follow-up. We started with consideration of long-term follow-up for MCADD. In a 1-day meeting, we developed a list of critical demographic and diagnostic-related elements that we wished to collect about affected individuals. These demographic elements would serve as the common demographic database for all disorders to be included subsequently. This group of elements, referred to as our "enrollment survey," included general demographic information, a specific query about whether families wish to be contacted for future information about studies, socioeconomic information, information about NBS results, and information about initial diagnostic testing. Most importantly, initial questions about clinical presentation are also recorded, and a record of the initial care plan is established. The information encompassed in the general enrollment survey is described in Table 2. This information is ascertained for each individual enrolled in the IBEM information system (IS). Although the enrollment survey contains some disease-specific elements, for example, which analytes are affected, to a large extent, these data will encompass information allowing us to draw very general conclusions about the whole cohort of screened children.

The unique quality of this data set and its potential utility as a potent means for establishing a research agenda and defining outcomes for long-term follow-up are encompassed in the paired "interval survey" developed for encounters for each subsequent visit by the enrolled subject. After each child is enrolled and given a unique patient identifying number, an interval survey is collected at each subsequent clinic visit. Again, the group worked together to define which elements may be important for long-term outcomes for a given disorder. The general outline for data collected at each visit is presented in Table 3. Generally, we attempted to define elements that would collect information about general status of the child, the frequency and type of medical encounters, laboratory and other clinical monitoring parameters, ongoing dietary and medication management, developmental outcomes, and coordination of care. This set of follow-up interval elements was carefully defined for our initial condition, MCADD.

To ensure that we captured a comprehensive view for data elements for both interval and enrollment surveys, we performed a comprehensive literature review to look for treatment suggestions. We also asked each center to provide their current care plans and protocols, reasoning that the professional experience of this expert group would yield important evidence about management strategies and their differences. We also were grateful to receive a copy of the Centers for Disease Control-funded Newborn Screening Long-Term Follow-up Data Collection System developed at Oregon Health and Sci-

Table 2 Initial enrollment data elements

Demographics (common to all disorders)
Unique registry ID number
Patient name[a]
Date of birth[a]
State newborn screen serial number[a]
Medical record number[a]
Is patient followed-up by more than one metabolic center?
Gender
Race of patient
Special ethnic group
Birth weight
Birth length
Occipital-frontal circumference at birth
Maternal educational level
Paternal educational level
Affected siblings?
Presentation (includes disease-specific data)
Pregnancy history
Means of initial diagnosis
Days of age at time family was notified of diagnosis
Days of age at time abnormal screen was reported to the primary care provider
Days of age at the time the abnormal newborn was reported to the metabolic provider
Days of age from birth to physician notification of abnormal screen result
Days of age from birth to treatment
Days of age at time of initial newborn screen collection
Days of age at time of initial face to face metabolic consultation with family
Method of diagnosis
Analyte levels on newborn screen
Symptoms and laboratory findings present at initial metabolic consultation
Was prenatal testing done during this pregnancy?
Diagnostic tests obtained
Confirmatory tests obtained
Disorder-specific genotype
Initial care plans
Provision of genetic counseling
Family was given a written emergency plan?
Family was given the 24-hr on-call contact information for a metabolic provider?
Patient was enrolled in a web-based emergency medical alert plan?
Internet/written support information was provided?

[a]Identifiable by enrolling center only.

Table 3 Interval data elements

Follow-up status
Is the patient still alive?
Date of death or date of last contact
Cause of death
Weight
Height
Occipital-frontal circumference
Laboratory testing
General and condition-specific laboratory tests collected and the interpretation of those results (qualitative summary or quantitative value depending on the test performed)
Imaging tests performed
Emergency care/hospitalizations
Number of emergency visits since the last metabolic visit
Number of metabolic-related emergency visits since the last metabolic visit
Number of hospital admissions since last metabolic visit
Number of metabolic-related hospital days since last metabolic visit
Disorder-specific complications
Disorder-specific monitoring used
Patient has a sick day plan?
Pharmacotherapy
Medication prescribed
Family reports compliance with medication
Nutrition intervention
Disorder-specific nutritional intervention
Family reports compliance with nutrition intervention
Developmental evaluation
Developmental milestones achieved
If no, which developmental milestones are not achieved?
Patient referred for further evaluation?
Developmental screening occurred during the metabolic center visit?
If screened were there abnormal findings?
Was patient referred for further evaluation?
Are behavioral concerns suspected?
If yes, patient was referred for further evaluation?
Referral for special education evaluation?
Neuropsychological assessment completed since the last metabolic visit?
Educational services currently received
Care coordination
Current insurance coverage
Community resource referrals
Health care referrals

ence University; data elements from this database were included in our organization of our initial prototype condition elements planning and subsequently.

HUMAN SUBJECTS PROTECTION

Information about long-term follow-up is certainly necessary for other interested parties. For example, Departments of Health should have information about long-term follow-up to judge the success of their programs. However, the IBEM-IS has been designed primarily as a platform for research. To that end, we decided that we would gather information only from individuals for whom we had recorded informed consent. Because we decided to limit general access to the data set and have designed it so only individual centers have access to their own data, we invoked the process of expedited review for evaluation by institutional review boards at each center's institution. To expedite the process of obtaining approval for initiation of research data gathering, uniform materials were prepared, providing sample protocols and consent forms. Shared access to these documents is available for all centers participating. Thus far, 10 centers in seven states have received approval to participate in the IBEM-IS. Materials used in this process are available at http://region4genetics.org/region4/ibem_is_irb.aspx.

We separated the decision of families/subjects to participate in the IBEM-IS registry activity only from a conscious decision on their part to allow further contact. Although we hope that families will agree to further contact if new research projects emerge, we understand that not every family wishes to have further contact even if they want us to gather their long-term follow-up data. This specific element is captured on the initial enrollment survey. This does not preclude the family changing their mind subsequently, but such contact would have to be initiated by the family if they wish to change their mind about further research activities sponsored through the IBEM-IS.

BUILDING THE IBEM-IS

Although we started with one disorder, the intention remained to provide comprehensive long-term follow-up information about all newborn-screened disorders. To that end, the workgroup mapped out a strategy for adding additional disorders. After completion of data elements for MCADD, the next two diseases that were added were very long-chain acyl-CoA dehydrogenase deficiency and maple syrup urine disease. Our general strategy was to build a matrix of disorders, thus the rationale for the above conditions was to extend the group of fatty acid oxidation disorders, demonstrating the utility of adaptation from the paradigm disorder, and extending to another MS/MS disorder group, an aminoacidopathy. Subsequently, we characterized elements for disorders diagnosed using the C3-acylcarnitine and C5OH-acylcarnitine species, initiating addition of the organic acidemia group of disorders. This also resulted in our decision to characterize elements for biotinidase deficiency, a non-MS/MS disorder as on occasion these infants may have C5OH-carnitine as a metabolite. Subsequently, an opportunity to participate with experts in fatty-acid oxidation disorders to form a consortium for the study of those conditions evolved, prompting the group to define data elements for all fatty acid oxidation disorders.

For each disorder, we followed the same procedure. We obtained appropriate literature as available. We asked participating metabolic disease clinicians to contribute care protocols followed in their centers. We consulted the Centers for Disease Control-funded Oregon database. In our further work, we were supported by a project initiated by the HRSA-funded Mountain States Genetics Collaborative. A work group headed by Janet Thomas, MD, University of Colorado, brought together expert metabolic disease clinicians from that region's academic centers to define minimal long-term follow-up data sets for each NBS disorder. We took this information and incorporated it into each of our groups of data elements. Subsequent work by the Region 4 Workgroup has defined disease-specific aspects for data entry of complete enrollment and interval surveys for 21 disorders as of July 2009.

INITIATION OF DATA COLLECTION AND DOCSITE® AS HOST FOR THE DATA

Having mutually agreed-on data elements for collection and having received IRB approval sequentially in metabolic centers, data collection was possible. To accomplish this, we reasoned that building our own electronic database from scratch would be expensive and beyond our expertise. To facilitate rapid initiation of the project, we turned to a commercial vendor familiar with quality assurance activities, DocSite®. DocSite® works in partnership with health plans, clinicians, patients, and other customers to offer point of care, outcomes tracking, and informed decision-making tools for population health and chronic disease management. They have experience in population-based cross-organization data exchanges from the health record and, critical to this project, have created a secure server system that will protect confidentiality and maintain the privacy of enrolled human subjects. The DocSite® web server uses 128-bit encryption, the highest level of internet security available. In addition, each individual user has password-protected access and viewing/editing privileges with differing levels of access to protected health information. Data security and privacy are also supported through a variety of other electronic mechanisms. After exploration of other database alternatives, we determined that DocSite® would best serve the clinical setting workflow and allow facile reconfiguration to accommodate additional disorders and research data tools for incorporation into the scope of this project over the next 5 years. We plan to continue our use of DocSite® as a tool for data entry and maintenance. We also anticipate their participation in data integration activities to be undertaken in this project.

INITIAL PROGRESS IN DATA COLLECTION

Data entry proceeded from the first center beginning in January 2007. A second center began entry in April and the third in June 2007. Eight centers representing six of the seven states in Region 4 are now enrolling patients and submitting data. As of July 2009, data entry is proceeding on 138 patients. Progress in data entry is noted in Figure 1. Not surprisingly, data entry for subjects with MCADD and other fatty acid oxidation disorders comprise the majority of enrolled subjects about whom data are being collected as these were among the first of the defined disorders. In Figure 1, data about enrollment includes only total numbers of patients that have data entry based on reports for the last 6 months. With additional centers receiving IRB approval, continued NBS ascertainment of those affected with IBEM in all participating states, and additional disorders now available for data collection, we anticipate continued growth of data entry. Continued addition of centers from regions beyond Region 4 should further enhance this capability.

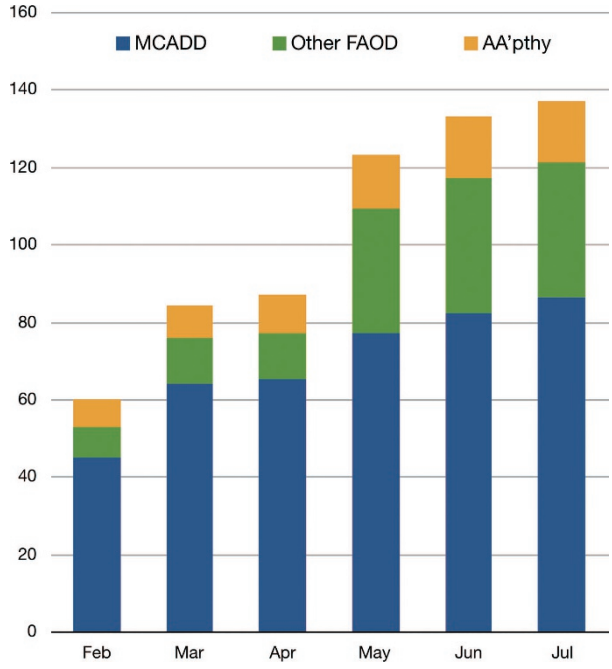

Fig. 1. Data collection progress, February to July 2009. MCADD, medium chain acyl-CoA dehydrogenase deficiency; FAOD, fatty acid oxidation disorder.

LIMITATIONS AND FUTURE CHALLENGES

Sustaining data collection and providing adequate incentives for centers to maintain data entry as a part of their workflow are among critical challenges to the long-term success of this project. For long-term, sustained data entry, it is likely that data entry will need to generate a work product useful to the clinical activities of the practitioner, for example, a billable patient chart note. In this initial effort, we provided some incentive for data entry by providing $50 per case enrollment to centers initiating data entry on a given patient. This will not be sufficient on a long-term basis to promote and sustain data entry over a prolonged period of time. All information entered in our database will ultimately need to be integrated into data sets generated beyond our region. Data entry will also need to serve the interests of a variety of stakeholders as it is unlikely that clinicians entering data will be willing or able to enter similar information more than once. Because our data set is based on subjects giving prospective informed consent, unless strategies are found to provide a complete denominator, there will be unavoidable gaps in the information we gather.

SUMMARY

Despite these potential challenges, our workgroup believed strongly that the knowledge to be gained was worth the effort we would expend. Continued collaboration among members of the workgroup should allow us to meet the objectives of our project:

1. Refine the research and intellectual questions that can be explored using the data in the IBEM-IS.
2. Expand both the spectrum of IBEM encompassed in the registry and the power of the data for each condition by adding other regions to the project.
3. Integrate the registry in additional domains, creating a true information system by increasing its interoperability and connections with other information and users to extend its functions and broaden its use; for example, web-based emergency medical plans, electronic medical records, and state NBS follow-up data systems.

ACKNOWLEDGMENTS

Supported by HRSA MCHB U22MC03963, Priority 2 Project: Follow-up Activity.

Region 4 Genetics Collaborative Priority 2 Workgroup: Allison Bannick, Ayesha Ahmad, Jan Askren, Barbara Burton, Troi Cunningham, Abby Donovan, Barbara DeLuka, Sandy Fawbush, Gerald Feldman, Gordon Gowans, Stephanie Gurnon, Bryan Hainline, Sally Hiner, Rachel Katz, Nancy Leslie, Claudia Nash, Kathryn Platky, Kimberly Regis, Carol Reid, William Rhead, Susan Romie, Jill Shuger, Kathy Stagni, Iris Stone, Sandy VanCalcar, and William Young.