Advisory Committee on Heritable Disorders in Newborns and Children

Meeting Summary
December 1, 2020

The Advisory Committee on Heritable Disorders in Newborns and Children (Committee) meeting was convened on December 1, 2020 and adjourned that same day. In accordance with the provisions of Public Law 92-463, the meeting was open for public comment

Table of Contents

Committee Members

Mei Baker, MD
Professor of Pediatrics
University of Wisconsin School of Medicine and Public Health
Co-Director, Newborn Screening Laboratory
Wisconsin State Laboratory of Hygiene

Jeffrey P. Brosco, MD, PhD
Professor of Clinical Pediatrics, University of Miami
Title V CYSHCN Director, Florida Department of Health
Associate Director, Mailman Center for Child Development
Director, Population Health Ethics, UM Institute For Bioethics and Health Policy

Kyle Brothers, MD, PhD
Endowed Chair of Pediatric Clinical and Translational Research
Associate Professor of Pediatrics University of Louisville School of Medicine

Jane M. DeLuca, PhD, RN
Associate Professor
Clemson University School of Nursing

Shawn E. McCandless, MD
Professor, Department of Pediatrics Head, Section of Genetics and Metabolism
University of Colorado Anschutz Medical Campus
Children's Hospital Colorado

Cynthia M. Powell, MD, FACMG, FAAP (Chairperson)
Professor of Pediatrics and Genetics
Director, Medical Genetics Residency Program Pediatric Genetics and Metabolism
The University of North Carolina at Chapel Hill

Annamarie Saarinen
Co-founder
CEO Newborn Foundation

Scott M. Shone, PhD, HCLD(ABB)
Director
North Carolina State Laboratory of Public Health

Ex-Officio Members

Agency for Healthcare Research & Quality
Kamila B. Mistry, PhD, MPH
Senior Advisor
Child Health and Quality Improvement

Centers for Disease Control & Prevention
Carla Cuthbert, PhD
Chief
Newborn Screening and Molecular Biology Branch
Division of Laboratory Sciences
National Center for Environmental Health

Food and Drug Administration
Kellie B. Kelm, PhD
Director
Division of Chemistry and Toxicology Devices
Office of In Vitro Diagnostics and Radiological Health

Health Resources & Services Administration
Michael Warren, MD, MPH, FAAP
Associate Administrator
Maternal and Child Health Bureau

National Institutes of Health
Diana W. Bianchi, MD
Director
Eunice Kennedy Shriver National Institute of Child Health and Human Development

Designated Federal Official

Mia Morrison, MPH
Genetic Services Branch
Maternal and Child Health Bureau
Health Resources and Services Administration

Organizational Representatives

American Academy of Family Physicians
Robert Ostrander, MD
Valley View Family Practice

American Academy of Pediatrics
Debra Freedenberg, MD, PhD
Medical Director, Newborn Screening and Genetics
Texas Department of State Health Services

American College of Medical Genetics & Genomics
Maximilian Muenke, MD, FACMG
Chief Executive Officer

American College of Obstetricians & Gynecologists
Steven J. Ralston, MD, MPH
Chair, OB/GYN
Pennsylvania Hospital

Association of Maternal & Child Health Programs
Jed Miller, MD
Director, Office for Genetics and People with Special Care Needs
Maryland Department of Health Maternal and Child Health Bureau

Association of Public Health Laboratories
Susan M. Tanksley, PhD
Manager, Laboratory Operations Unit
Texas Department of State Health Services

Association of State & Territorial Health Officials
Christopher Kus, MD, MPH
Associate Medical Director
Division of Family Health
New York State Department of Health

Association of Women's Health Obstetric and Neonatal Nurses
Shakira Henderson, PhD, DNP, MS, MPH, RNC- NIC, IBCLC
Vice President, Research Officer University of North Carolina Health Board Director, Association of Women's Health, Obstetric & Neonatal Nurses

Child Neurology Society
Jennifer M. Kwon, MD, MPH, FAAN
Director, Pediatric Neuromuscular Program American Family Children's Hospital
Professor of Child Neurology, University of Wisconsin School of Medicine & Public Health

Department of Defense
Jacob Hogue, MD
Lieutenant Colonel, Medical Corps, US Army
Chief, Genetics, Madigan Army Medical Center

Genetic Alliance
Natasha F. Bonhomme
Vice President of Strategic Development

March of Dimes
Siobhan Dolan, MD, MPH
Professor and Vice Chair for Research
Department of Obstetrics & Gynecology and Women's Health
Albert Einstein College of Medicine and Montefiore Medical Center

National Society of Genetic Counselors
Cate Walsh Vockley, MS, CGC
Senior Genetic Counselor Division of Medical Genetics
UPMC Children's Hospital of Pittsburgh

Society for Inherited Metabolic Disorders
Georgianne Arnold, MD
Clinical Research Director, Division of Medical Genetics
UPMC Children's Hospital of Pittsburg

I. Administrative Business

Cynthia M. Powell, MD, FACMG, FAAP
Committee Chair
Professor of Pediatrics and Genetics
Director, Medical Genetics Residency Program Pediatric Genetics and Metabolism
The University of North Carolina at Chapel Hill

Mia Morrison, MPH
Designated Federal Official
Health Resources and Services Administration (HRSA)

A. Welcome and Roll Call

Dr. Powell welcomed participants to the December 2020 meeting of the Advisory Committee on Heritable Disorders in Newborns and Children.

Dr. Powell then conducted the roll call. The Committee members in attendance were:

- Dr. Kamila Mistry
- Dr. Mei Baker
- Dr. Kyle Brothers
- Dr. Jane DeLuca
- Dr. Carla Cuthbert
- Dr. Kellie Kelm
- Dr. Michael Warren
- Dr. Shawn McCandless
- Dr. Melissa Parisi
- Ms. Annamarie Saarinen
- Dr. Scott Shone

Organizational representatives in attendance were:

- American Academy of Family Physicians, Dr. Robert Ostrander
- American Academy of Pediatrics, Dr. Debra Freedenberg
- American College of Medical Genetics and Genomics, Dr. Maximilian Muenke
- American College of Obstetricians and Gynecologists. Dr. Steven Ralston
- Association of Maternal and Child Health Programs, Dr. Jed Miller
- Association of Public Health Laboratories, Dr. Susan Tanksley
- Association of State and Territorial Health Officials, Dr. Christopher Kus
- Child Neurology Society, Dr. Jennifer Kwon
- Association of Women's Health and Obstetrics and Neonatal Nursing, Dr. Shakira Henderson
- Department of Defense, Dr. Jacob Hogue

- Genetic Alliance, Ms. Natasha Bonhomme
- March of Dimes, Dr. Siobhan Dolan
- Society for Inherited Metabolic Disorders, Dr. Georgianne Arnold

B. Committee Business

Dr. Powell announced that Dr. Shakira Henderson is replacing Dr. Rychnovsky as the Organizational Representative from the Association of Women's Health, Obstetric and Neonatal Nurses (AWHONN). Dr. Henderson is a hospital-based nurse researcher, health educator, and certified neonatal nurse, who serves as an elected member and neonatal expert for AWHONN. Dr. Powell thanked Dr. Rychnovsky for her service to the committee.

Dr. Powell noted that the Committee received a nomination package for mucopolysaccharidosis II (MPS II). She will keep the Committee apprised of next steps.

Dr. Powell reminded the Committee they had received two presentations during the August 2020 meeting: *1) Assessing Values as Part of the Newborn Screening Committee's Evidence Review Process;* and, *Review of Newborn Screening Timeliness*. Dr. Kemper and Dr. Lam incorporated feedback from the discussions into both reports and the Committee was provided with final drafts in the December Briefing Book.

C. Committee Vote and Review of Newborn Screening Implementation for Spinal Muscular Atrophy (SMA) Final Report

The first item on the agenda was to vote on whether or not to approve the Review of Newborn Screening Implementation for Spinal Muscular Atrophy Final Report. In February 2018, the Committee voted to recommend to the Secretary to expand the Recommended Uniform Screening Panel (RUSP) to include the addition of SMA due to homozygous deletion of exon 7 in SMN1. The Committee was notified in July 2018 that the Secretary had accepted the Committee's recommendation and requested a report, "…describing the status of implementing newborn screening for SMA in clinical outcomes of early treatment, including any potential harms for infants diagnosed with SMA." In collaboration with subject matter experts, the Committee developed a report detailing states' experiences with implementation of screening for SMA and impact for newborns diagnosed with the condition.

The final draft report incorporated feedback from Committee members and organizational representatives and includes the following substantive changes:

- In August 2020, risdiplam was approved to treat adults and children two months and older with SMA.
- The report clarifies that testing for *SMN2* copy number requires a separate assay.
- Assessing for *SMN2* copy number was pulled into its own subsection within the report and additional details were added, including that at least two newborn screening programs are using a droplet digital PCR method to report *SMN2* copy number.

Committee Discussion

- A Committee member requested a minor edit to the report: The first paragraph in the Summary and Conclusions (page 27) states SMA screening does not require new equipment or expertise. This statement contradicts with information provided on page 11 which notes that implementation of screening for SMA requires *little or no* additional equipment or expertise. Many newborn screening programs and laboratories have had to upgrade their Severe Combined Immunodeficiency (SCID) testing equipment to implement SMA. He requested that the sentence on page 27 be revised to, "SMA screening requires little to no additional equipment". The Committee voted unanimously to approve the report with the suggested revision.

II. Newborn Screening Decision Making Criteria and Matrix

Alex R. Kemper, MD, MPH, MS
Lead, Evidence-based Reviews

The Committee is in the process of reviewing its condition nomination, evidence review and decision-making processes. Dr. Kemper presented on the background of the newborn screening decision-making criteria and matrix. He framed the presentation with the following observations:

- Making decisions about which conditions to recommend for inclusion on the RUSP is challenging and complex;
- The fact that the Committee is taking a detailed look at the matrix does not mean that the matrix doesn't work. Rather, it is an indication that the Committee is continuing to review its processes including its decision-making processes.

The Committee's current/revised decision matrix was first published in a 2014 article in *Genetics in Medicine*. Dr. Kemper suggested Committee members refer to the article for detailed information on the development of the matrix.

Dr. Kemper provided a brief overview of the evidence review process and Committee deliberations using the decision matrix. The Condition Review Workgroup conducts a systematic review of the evidence that allows the Committee to deliberate on issues related to the net benefit of screening as well as potential harms associated with testing for the condition. Along with determining the degree of benefit there are also the issues of the certainty of the evidence.

The Committee also considers readiness and feasibility of newborn screening programs to implement the condition. Because the conditions are rare or ultra-rare there will always be uncertainty. Dr. Kemper highlighted that, "…a B rating of the evidence indicates moderate certainty that screening would lead to a significant net benefit. The term moderate indicates that the Advisory Committee believes further research could change the magnitude or direction of findings such that the assessment of net benefit would be small to zero or even negative."

Dr. Kemper reminded the Committee that the current decision matrix was developed in 2014 and that exploring ways to refine the Committee's deliberation processes is normal and expected. Since 2014, the Committee has used the matrix a number of times and there may be opportunities to apply lessons learned. He suggested it would be useful for the Committee to consider ratings for Conditions reviewed by the Committee since 2010 and the subsequent recommendations.

Dr. Powell posed the following questions to the Committee:

- Are there components that need clarification?
- How should the Committee approach a B rating?
- Should cost data be considered as part of state readiness or feasibility? If so, what cost data could be included?
- If a condition were to be removed, how would the matrix be used in that process?

Discussion

- An Organizational Representative asked if there should be a scalar representation of certainty as opposed to high, moderate and low and if this would help clarify the B ratings. He suggested that a scalar representation would be more helpful for certainty and would support the Committee's highly nuanced deliberations particularly when a condition is on the cusp of a rating. In his opinion, a scalar representation would be less important in terms of readiness and preparedness. He agreed that the Committee should implement a process for re-examining conditions on the RUSP. One approach would be to have a subcommittee that periodically reviews conditions.
- An Organizational Representative wondered if the concept of net benefit needs clarification. He also asked the Committee to consider if benefits versus risks are skewed towards different populations based on race, ethnicity, geography etc. It is important for the Committee to consider how to approach these types of disparities given that there may be a relatively small body of published research on a particular condition.
- An Organizational Representative recommended defining the different categories within the decision matrix. For example, is the term net benefit interpreted uniformly? She receives questions from advocacy groups and industry representatives about how closely the Committee adheres to the matrix, indicating that there is a potential area of need for additional stakeholder education. In terms of considering when a condition should be removed, she felt a separate conversation was warranted to determine the process and if the matrix would be the most appropriate tool.
- Dr. Kemper noted two important themes:
    - It would be helpful for the Committee to strengthen the way it addresses net benefit because the benefits and harms occur to different groups of people.
    - The degree to which the Advisory Committee uses the matrix as a deliberation tool versus a prescriptive tool.
- A Committee member emphasized that the principles should be the same in terms of adding or removing a condition. She recommended conducting a regular review every

five to ten years. As new evidence becomes available, there should be a systematic process for reviewing and removing conditions.

- One Committee member suggested either conditional approval of additions to the RUSP or automatic review of conditions for which long-term evidence is lacking. Two-five years may not be sufficient to answer critical public health questions such as: who is getting care; types of care; and outcomes such as who is still alive. Currently, there is no mechanism in place to collect that information.
- A Committee member recommended that the matrix should provide general guidance and not be overly prescriptive for the following reasons:
    - It is challenging to predict the relative success of screening for a condition during the evidence review process; and,
    - Given the timeframe for reviewing conditions, a large portion of the data is likely to fall into more of a moderate/B range category which leads to some intrinsic uncertainty. At some level, the Committee has to weigh whether they believe the general likelihood of benefit from the available evidence is unlikely to be overturned or to be reversed such that they feel comfortable with recommending a condition for addition to the RUSP.
- A Committee member noted that there appears to be more of a struggle with the B category and wondered if there would be a benefit to expanding the B section. It would also help if some of the sections in the decision matrix were more explicit. Regarding the question on removing a condition, she pointed out that it would be difficult for states to remove a condition after implementation. The Committee member agreed with a recommendation made earlier in the discussion to conduct a periodic, perhaps annual review of specific conditions not with the objective of determining if the condition should be removed but rather to further refine and build on what is already known about conditions on the RUSP.
- A Committee member expressed concern about ongoing challenges related to disorders in the B category. He wondered if newborn screening once thought of as addressing the diagnostic odyssey is now creating a treatment or prognostic odyssey. For example, there are instances were labs face challenges determining the *SMN2* copy number and for some states that have added it to their newborn screening panel, there are challenges linking families to care.
- A Committee member suggested that perhaps it would be helpful to conduct a systematic review of conditions on the RUSP that received a B rating. She also reminded the Committee that it is important to look at the number of children and families that have been positively impacted by newborn screening and what a removal or any substantive change to the RUSP looks like at the state level.
- An Organizational Representative noted that SMA received a B rating and has been successfully implemented by newborn screening programs across the U.S. The B rating reflects the Committee's thoughtful and engaged conversation about the data available at the time of the evidence review. X-linked adrenoleukodystrophy (X-ALD) got an A rating and there is considerable evidence supporting the benefits of early diagnosis and early screening but not necessarily newborn screening. She suggested that the

Committee should review RUSP conditions not for the purposes of removal but for the purposes of making the deliberation process more consistent. Could the Committee look at the impact of screening on states and the difficulty of identifying cases? What is the prognostic odyssey like for families who have a diagnosis but are not sure what that diagnosis might mean?

- A Committee member asked about cost data and wondered if it included more than the laboratory costs or does it include staffing costs related to follow-up consultation? He noted that there is an inaccurate assumption that states can raise newborn screening fees, however, in many cases states face limitations in raising fees and are required to undergo a legislative review or approval by the governor in order to do so.
- Two Committee members supported the idea of conducting annual reviews of one or two conditions to consider reclassification (e.g., from primary to secondary or vice versa) or redefinition. A Committee member and Organizational Representative noted that refining recommendations to ensure greater consistency should be the goal (as opposed to removal).
- Dr. Kemper responded that it is difficult to assess cost. Thus, the focus has been on trying to estimate the cost to newborn screening programs for adopting the specific laboratory screening test. He questioned the feasibility of conducting a detailed cost analysis with any degree of rigor or certainty as costs are changing constantly. He would not want to overstate the type of cost analysis that could be completed during the evidence review process.
- A Committee member agrees that cost data is important but questions the utility of incorporating granular cost data into the evidence review process. One suggestion is to have broad categories of upper and lower cost categories for adoption. Cost could be a barrier to implementation, especially if a new test might cost $20.00-$30.00 per newborn.
- Dr. Kemper asked if the issue of availability of follow-up services should be incorporated into the considerations around readiness and feasibility?
- An Organizational Representative asked if net benefit should examine the impact of benefits apart from biological interventions that come from early detection.
- A Committee member noted that an underlying agreement about the parameters to determine benefit, readiness and feasibility are missing from the conversation. What definition of benefit are they looking for both in the short- and long- term? What costs are important? In addition, what matters about the resources and the availability of resources (e.g., equal access to specialists and other health care/therapeutic services)? The Committee should define these things before the matrix can be used as a prescriptive tool. He also agrees that it is important to establish a review system for conditions for which there is not good long-term outcome data. It is unfair to expect that there would be individuals or groups of individuals that would put themselves forward to nominate for the removal of a particular condition. The bar for removal of a condition without pre-existing definitions for what would lead to removal would be so high, making it nearly impossible. Perhaps there could be a conditional approval for addition to the RUSP with a requirement for follow-up data.

- An Organizational Representative expressed her concern about the availability of the workforce. Increasing the number of conditions now and waiting for the workforce to catch up will not work well. She wondered if the Committee could establish a workforce workgroup to liaison with other medical societies to discuss these issues.
- An Organizational Representative said she was intrigued by the idea of conducting a systematic review of conditions on the RUSP every five years. They could collect simple information such as the number of infants who screened positive; number successfully followed and those lost to follow-up. This would provide a report card that captures how well they are doing with some of these conditions which will lead to conversations that include barriers to follow-up. She gave the example of Pompe disease where they know that early screening is valuable but how children do long- term is still unknown with unexpected outcomes.
- Dr. Kemper asked the Advisory Committee to consider whether or not there are unexpected disparities by region, race, ethnicity etc.
- An Organizational Representative commented on cost considerations being part of the state readiness and feasibility. Every state has a different funding mechanism some are fee-based; others are revenue-based or funded through Medicaid. Adding SMA screening may seem simple but one has to replace equipment, have more staff for follow-up etc. She liked the bucket/upper-lower limit concept to capture the relative expense for a program. She also supported the idea of a workforce workgroup and asked if they are going to see an increase in physicians specializing in metabolic conditions due to increased demand from newborn screening?
- Dr. Powell asked if the system is currently able to assess 5-10 year, longitudinal follow-up information and if not, what would be needed?
    - A Committee member answered that there are two components to consider: there is a growing body of literature for specific conditions like *Pompe* suggesting that there were unexpected outcomes and that the natural history of the disorder has significantly changed with therapy for the better in many ways. However, the research has also unmasked some unexpected aspects of the natural history of the disease that they did not see because patients did not survive long enough. Additionally, there will be a growing body of literature supporting the facts about the physical benefits to patients receiving these treatments after early diagnosis; leading to a growing body of literature supporting the timing of the initiation of the therapy. Suggesting that earlier treatment is better and may even raise questions about whether prenatal treatment is needed in some cases. The Committee member believes SMA is one condition where this will be the case.
    - Questions related to access to care are more difficult to answer. The answers to the basic public health questions of who is lost to follow-up, who is getting adequate care; who is still alive; is there growth and development and overall health for the treated condition are unknown and would probably require building some infrastructure.

III. Review of Newborn Screening Implementation for Added RUSP Conditions: *SCID, CCHD, Pompe, MPS-1, XALD*

Alex R. Kemper, MD, MPH, MS
Lead, Evidence-based Reviews

Dr. Kemper recapped an earlier presentation on the lessons learned from the implementation of conditions added to the RUSP since 2010 including SMA. He summarized the timeline between when the conditions were first nominated, re-nominated (when applicable), the dates of the Committee's vote and when the Secretary formally added a condition to the RUSP. He shared the proportion of states that are screening for SMA and asked the members to keep in mind the time period from when a condition is recommended to the RUSP to when states might begin screening. The time lag could be due to personnel issues, getting the appropriate equipment, and figuring out the process of exactly how the screening and algorithms will be implemented. Additional challenges include updating data information systems and establishing protocols for short-term follow-up. He noted that facilitators of new disorder implementation include: states utilizing Peer Resource Networks; availability of pilot funding; collaboration with patient advocacy groups and mandated state adoption of RUSP conditions.

He shared information for each newborn screening condition covered in the report and SMA, providing information including the number of infants screened, true positives and false positives. He also discussed the positive predictive value which is the proportion of newborns that screen positive that are confirmed/diagnosed with the condition. He also discussed the following outcomes published in peer-reviewed literature:

- A report on severe combined immunodeficiency (SCID) indicating a high survival rate for infants detected through newborn screening.
- A Journal of the American Medical Association (JAMA) article that looked at screening policies and infant deaths related to critical congenital heart disease (CCHD), demonstrating that states with mandated CCHD newborn screening policies had about a one-third reduction in deaths due to CCHD following the implementation of newborn screening.
- Published studies on newborn screening for Pompe disease and mucopolysaccharidosis type I (MPS I) demonstrating a high variability in the positive predictive value because they are rare diseases.
- New York State's experiences with screening 225,000 infants for SMA and the number of cases that were detected and the benefit from having this level of early identification.

Discussion

- A Committee member asked whether the data in the presentation reflected the cumulative positive predictive value of SCID across the three reports. Dr. Kemper indicated the data was cumulative but requested caution when interpreting this data as states interpret a positive screen differently, algorithms are complicated and the

denominators change. Dr. Kemper posits that looking at the true positive screens and the magnitude of public health impact are the key lessons as there is a high degree of variability in what is counted as positive predictive value.

- An Organizational Representative stressed the importance of data and hoped that the Committee will continue to look at the long-term data that has been collected for these conditions.
    - Dr. Kemper concurred and clarified that the data presented are from published research and does not include unpublished data from states. He wondered if access to state data was possible and noted that it would be ideal if states published their findings in peer reviewed literature.
- An Organizational Representative concurred with Dr. Kemper on the usefulness of being able to look at complete data from states.
- An Organizational Representative recommended the Committee look at the Follow-up and Treatment Workgroup's framework/blueprint for long-term follow-up. He suggested the Committee ask nominators to include an outline of long-term follow-up requirements in their nomination package.
- An Organizational Representative reiterated that states have a large amount of data some of which they are willing to share and some they are not. She pointed out that this type of information and funding for it tends to get siloed. It behooves the Committee to take a holistic approach and look at the overall system because newborn screening programs do not function in isolation: they partner with clinicians, metabolic disease doctors and other specialists who treat these conditions. In the Organizational Representative's state, they are beginning to examine the long-term follow-up for many of the health measures.
- A Committee member wondered if the Committee considered doing a meta-analysis of implementation for newborn screening conditions. She also stated that in the case of CCHD for example, it can be challenging to accurately capture the true number of positives or lives saved.
    - Dr. Kemper concurred and noted that as states are collecting newborn screening data and they also have a multitude of competing priorities. Thus, the bigger question is whose responsibility is it to collect and report all this information collectively?
    - Dr. Powell agreed about the importance of the issue and specifically noted that for CCHD the challenge among states is that there is limited reporting but as time goes by it will hopefully be strengthened through interoperability.
- An Organizational Representative suggested that the available long-term data tends to concentrate more heavily on more severe disorders. It is important to look at long-term outcomes with the understanding that there may be confirmation bias towards more severe disorders.

IV. Public Comment

Dylan Simon
The EveryLife Foundation

Mr. Simon thanked the Committee for the opportunity to speak and thanked the members of Evidence Review Group that spoke at the Newborn Screening Bootcamp. The Bootcamp was a five-week event designed to educate stakeholders. More than 200 individuals attended at least one week of the Bootcamp. There were a wide variety of sessions including the importance of building coalitions to the challenge of implementation. The Bootcamp closed with three discussion groups: 1) advocates self-identifying as just getting started in newborn screening; 2) advocates on the path to RUSP; and, 3) advocates with conditions on the RUSP. Participants shared ideas and best practices on how to address common challenges facing the community. The next Bootcamp will take place in Sacramento, CA next year.

Steve Holland
National MPS Society

Mr. Holland thanked the committee for the opportunity to speak. He shared that the Society was founded 46 years ago with the mission to cure, support, and advocate for the MPS and mucolipidoses (ML) diseases. They serve individuals and families by supporting research and increasing public and professional awareness. All 15 known subtypes of MPS and ML diseases are rare and caused by the body's inability to produce certain enzymes. The missing enzymes result in the storage of materials in cells throughout the body. As the disease progresses, there is systemic damage throughout the body, including the heart, bones, joints, respiratory system and central nervous system leading to widespread physical and developmental challenges and a shortened lifespan. Due to the progressive nature and irreversible damage, especially in the brain, it is imperative to diagnose these diseases as early as possible to enable more effective intervention. To date about half of the MPS and ML subtypes have U.S. Food and Drug Administration (FDA) approved treatments with over twenty companies working on new treatments.

With available markers and low-cost tests, MPS is a primary candidate for newborn screening. A few years ago he worked closely with the Committee to get MPS I on the RUSP. Since its inclusion in the RUSP, MPS I has been included in 21 state newborn screening programs and the District of Columbia, enabling newly diagnosed patients to access treatments and interventions as newborns. With the collaborative efforts from their academic and industrial partners, the Society has recently completed the nomination package for MPS II. In parallel, the nomination for mucopolysaccaridosis type VII (MPS VII) is also underway. Their goal is to work with the Committee over the new few years to submit nominations for the remaining MPS and ML diseases with effective therapies.

Ryan Colburn

Mr. Colburn thanked the Committee for having him. He was diagnosed with Pompe Disease five years ago. Mr. Colburn is passionate about data and his written submission is in support of the importance and power of data; and to advocate for making as much de-identified data as possible publicly available in order to support improved health for newborns and children.

He emphasized that data is an opportunity to use newborn screening as a vehicle to accelerate the rate of development of new knowledge of rare diseases. He stressed that when data is shared, it amplifies the opportunity for a variety of stakeholders including pre-med students, family members, advocates and researchers to study conditions identified through newborn screening.

Mr. Colburn wants to help advance progress for Pompe and other rare diseases. Currently the effort required to compile the data is a barrier and prohibits many stakeholders from contributing to the development of new knowledge in rare diseases. However, data with huge potential to be useful is collected through federally supported programs and is not publicly accessible.

Per his written statement, Mr. Colburn scoured publicly available data to find the incidence of Pompe disease and his analysis shows a clear convergence that excludes the currently most cited projections. He emphasized the importance of an open access policy to data while honoring the collective duty to protect the privacy of individuals.

V. Newborn Screening in Genomic Medicine and Public Health (NSIGHT)

Jonathan S. Berg, MD, PhD
Professor, Department of Genetics
The University of North Carolina at Chapel Hill

Dr. Berg was the Co-Principal Investigator (PI) with Dr. Powell for the North Carolina (NC) Newborn Exome Sequencing for Universal Screening (NEXUS) project. He discussed the inspiration for the NSIGHT consortium, which stems from predicting the many uses of genomic sequencing as the cost of sequencing genomes makes it increasingly accessible.

In the NC NEXUS project, study participants from two cohorts were sequenced: 1) one group was recruited prenatally from healthy pregnancies/healthy newborns; 2) the second cohort was a group of infants and children recruited from clinics at the University of North Carolina (UNC) who had a condition identified through newborn screening. Patients from the metabolic clinic and the hearing loss clinic were included in this cohort as a way to gauge how well the sequencing would have identified the diagnoses. The researchers defined several different categories for result disclosure using clinical actionability.

Extensive parental survey research was done in the study. Parents were then randomized to two groups to study whether they would be interested in additional genomic findings outside of conditions on a typical newborn screening panel. Parents were randomized to either a decision arm or a control arm. There were several categories of information which were defined based on whether they met a threshold for clinical actionability:

- Conditions of childhood onset and medically actionable were put in the NGS/NBS category.
- There was also a category of excluded information including adult onset, and non-actionable conditions that the group decided not to report to participants for ethical reasons.
- There is also a middle category that included conditions which did not quite meet the threshold of actionability or statuses such as carrier status for recessive disorders or adult onset, medically actionable conditions where the researchers wanted to address parental preferences and consider potential psychosocial implications of having to make decisions based on that information (this was in the optional category for those that were randomized to the decision arm.)

Dr. Berg shared some of the results and said he would argue for genetic testing increase in sensitivity comes from in depth analysis of the variants of uncertain significance (VUS). The specificity of the tests is also affected and those variants may not be pathogenic using genomic screening as a screen. If the variants were restricted only returning the known pathogenic variants, the true positives would be sacrificed because those variants might be rare or private and would not meet the set threshold for being classified as a known pathogenic variant. However if the set threshold was dropped for specificity and included results with anything less than 99.99% specificity, there would be more false positives than true positives in a large population screen which would have downstream costs and consequences that would be important to consider.

In metabolic conditions that are included in newborn screening, a threshold can be set so that we get a certain number of false positives which ensures we identify all of the true positives. Then a secondary test is used to confirm or refute that initial screening finding. The challenge for the molecular test is that we don't always have a secondary enzyme test or other analyte that can be used to verify the finding as being pathogenic. So it raises the question of whether using the likely pathogenic results should also be considered. Should the degree of variant return be tapered based on what types of available follow- up are there?

In NC NEXUS the first molecular analysis was done blinded to whether this was an affected or unaffected individual. The analysts looked at all genes and would report the variants that were thought to be pathogenic or likely pathogenic. The results would then be a positive or negative screen result. After the analyst registered their results from a blinded analysis, they were then unblinded so that they could know whether it was an individual from the metabolic cohort, or the hearing loss cohort. Then a second list of genes was examined that would be included for the diagnostic analysis of the relevant genes. This resulted in a positive, negative or

inconclusive result that would also include the VUS. This way the results would be returned to the patients and we could also evaluate how well the screening would have done if this had only resulted as a positive or negative screen.

There were 17 metabolic patients and the team was able to identify clearly causative variants for 15 out of the 17 without even knowing that they had the diagnosis. There were a couple of negatives and one of them was due to only being able to identify a single heterozygous pathogenic variant in the gene associated with Maple Syrup Urine Disease (MSUD). This was probably an example where the second variant is not detectable by exome sequencing. Those types of false negatives are going to depend on the sequencing technology that is being used and what the analytics are looking at.

The second false negative is thought to be a homozygous missense variant in the *MLYCD* gene in a patient with malonyl-CoA decarboxylase deficiency. It's highly likely that this variant is the cause of that patient's disease but it was a unique variant and had not been reported previously. It was a VUS. Both of these are reported as inconclusive on the diagnostic report and is appropriate if one is thinking about a patient with MSUD to report as a single heterozygous pathogenic variant. And if a patient had decarboxylase deficiency, you would want to see the homozygous missense variant. Using this method, you would find carriers of MSUD at a much higher rate than you would actual patients. Reporting all carriers for MSUD would lead to a large number of results and a high number of missense VUS if you allowed the VUS results to be returned.

In the hearing loss cohort there was not quite as high a clinical sensitivity. Only about 5 out of the 28 patients had a clear diagnostic result. There are a number of reasons for this in part because hearing loss might be non-genetic. After unblinding there were some additional results that the researchers determined were appropriate to return as an inconclusive finding. This may be a less efficient way to screen for hearing loss than actually using the phenotypic screening.

Dr. Berg explained that through their research they identified individuals with the following conditions:

- A patient with a heterozygous *LDLR* pathogenic variant for familial hypercholesterolemia that was unknown when the patient was enrolled.
- An individual with a novel splice variant in a *DSC2* gene associated with a type of cardiomyopathy.
- A patient with heterozygous factor 11 variants that probably indicates a bleeding disorder.
- Results were reported for a female study participant with phenylketonuria (PKU) who was heterozygous for an *OTC* (ornithine transcarbamylase) pathogenic variant. The *OTC* pathogenic variant was probably a hypomorphic variant and was actually found in some of her male family members who were unaffected.

There were 45 parents out of the group that were randomized to the decision arm and 41 of them requested at least one other category of information. This could indicate how motivated the parents were to participate in the study. It is unclear whether 90% of the general population would have requested at least one other category of information but it indicates there is a great interest among parents for this type of information. Thirty-four parents requested all three categories.

When thinking about genomic sequencing as a screen, the perinatal period is a difficult time to conduct informed consent. It may be different with a child with a phenotype requiring sequencing as opposed to the parent of a healthy child. What needs to be examined is tolerance for false positives and over-diagnosis. We also need to determine how to gauge what types of variants and what conditions are returned as results. Dr. Berg strongly suggested starting with a subset of the most well understood, highly actionable conditions that could be analyzed in a targeted way in newborns rather than exome or genome sequencing.

Dr. Berg's workgroup has been thinking about age-based genomic screening and how could sequencing for age relevant conditions be worked into routine wellness visits in newborns and children. He envisions that there would be several time points during pediatric care in which genomic screening for conditions relevant to those age groups could be offered and might avoid some of the challenging ethical legal and social issues (ELSI) related to sequencing in newborns and children; and, allow parents to focus on conditions relevant to their child at that time and not necessarily front-load everything on newborn screening.

The advantage would be that it gradually introduced genomic screening to children over time- and as they become adolescents and start to engage in this decision-making process, they will be more prepared to engage in making decisions about adult-onset conditions or their own health screening in adulthood.

Robert Currier, PhD
Research Associate, Department of Pediatrics
University of California San Francisco

Sequencing of Newborn Blood Spot DNA to Improve and Expand Newborn Screening (NBSeq) sought to answer whether whole exome sequencing could replace MS/MS screening for inborn errors of the metabolism and how whole exome sequencing could supplement MS/MS. Evidence from the NBSeq study was recently published in *Nature Medicine.*

To address the main question, dried blood spots obtained from the California biobank program were sequenced. The biobank includes dried blood spots from all 4.5 million newborns screened between 2006 and 2013. From these there were a total of about 1,300 affected children with an identified metabolic disorder. A comprehensive set of these was requested. Also requested were all 13 known cases where mass spectrometry failed to identify a metabolic disorder. In addition, about 400 cases where mass spectrometry screening were positive but

the newborns were unaffected were examined. The false samples were not random samples. They consisted of all the positive results for six screening tests- PKU, MSUD, isovaleric academia (IVA), very long chain acyl-CoA-dehydrogenase deficiency (VLCAD),, and long-chain 3-hydroxyl-coenzyme A dehydrogenase deficiency (LCHAD), and glutaric acidemia type II (GA II). They have large numbers of false positives or because the overlapping profiles in the MS/MS screen make differential diagnosis more challenging. Samples from the Neonatal Intensive Care Unit (NICU) were excluded to avoid the effects of prematurity in treatment.

Of the 1,728 samples, 538 were omitted because of budgetary restraints or failure of quality control on the resulting DNA. The remaining 1,190 exomes were divided into 178 and a test set of 1,012. In order to model newborn screening, it was essential that the interpretation pipeline be completely automated. The analysis of the validation set showed that it was necessary to include rare protein altering variants of a MAF of less than .005. Of the 674 affected cases in the test set, only 377 would have been identified using curated variants. Including rare and predicted variants brings the total to 571, but that still leaves 103 missed by exome screening.

Exome screening had a sensitivity of 88 percent and specificity of 95 percent, which compares unfavorably with current MS/MS screening. Restricting to only curate the variance improves the specificity.

The sensitivity varied by disorder. For many the exome was 100 percent sensitive, but for BKT analyte screening one was missed. However, even in some PKU and MCAD results, the exome pipeline failed to identify some individuals. The secondary conditions and additional conditions group showed similar overall sensitivity. Because the study included false positive results for PKU, there was the opportunity to examine the performance of sequencing as a screening test for PKU. The first point is that of the 108 variants identified in the study samples, 23 or roughly 23 percent were absent from ClinVar. Using the published guidelines only 6 could be called pathogenic with the remainder being variants of uncertain significance. Using the results of the clinical follow-up of positive PKU screens, the study cohort was divided into five groups:

- 67 classical PKU;
- 27 variant H-PHE;
- 30 benign H-PHE;
- 31 false positive screens;
- 1,023 negatives.

PKU is an autosomal recessive condition where mutations in both copies of the gene lead to disease, each of these phenotype groups can be broken down further by the number of variants they had in the phenalalanine hydroxyase (PAH) gene. Two PAH cases were flagged by the pipeline. Of the 128 diagnosed clinically with either classic PKU or H-PHE, the predominant majority, around 92 percent, had two variants. Among the 1,023 controls, there were 30 with a single PAH variant. Most interestingly, among the 31 MS/MS positive cases, 16 had two variants in the PAH gene.

For the classic PKU cases, both variants in the diplotype were most often pathogenic. Interestingly, even in classic PKU, 11 cases had one variant of uncertain significance. Among the MS/MS false positive cases with two variants approximately one-third of them had two variants in the diplotype annotated as pathogenic.

Sequencing for screening in effect reverses the process of diagnostic sequencing. Rather than identifying the genetic cause of an existing phenotype, screening seeks to infer a future phenotype from the genetic variants. One-third of the variants had not been seen before. Inferences about disease severity from pathogenicity assertions of variants in autosomal recessive disorders of individuals are not straightforward.

A feature of the NSIGHT program was an important integrated ELSI component. The NBSeq group convened the NSIGHT ethics and policy advisory board, including representatives from all four of the NSIGHT groups plus other experts in law, ethics, law, policy, and medicine. The group considered four contexts in which genetic sequencing might be applied to newborns:

1. Clinical care of sick infants, particularly in the NICU;
2. Public health newborn screening;
3. Routine pediatric care of healthy newborns; and,
4. Direct to consumer genetic testing.

The conclusion of the deliberations can be summarized by the title of the final report – "Sequencing Newborns: a Call for Nuanced Use of Genomic Technologies**.**"

Currently, whole exome sequencing alone is not suitable as a sole screening methodology for inborn errors of metabolism. For selected disorders, sequencing was as reliable as MS/MS; could reduce false positive results; and, could facilitate an accurate and timely case resolution.

Stephen F. Kingsmore, MB ChB BAO DSc FRCPath
President/CEO, Rady Children's Institute for Genomic Medicine

Dr. Kingsmore's research focused on newborns and children who are suffering with symptoms of disorders, in an intensive care unit (ICU) setting, that are suggestive of a heritable disorder. Three factors are currently coalescing to revolutionize the approach to diagnosis and screening. The first is the exponential decrease in the cost of genome sequencing. It currently costs $600 - $700 to sequence a genome. The second factor is associated with the timeliness of sequencing in the newborn period.

Dr. Kingsmore's team wondered if they could return results in two days as is necessary for unstable children with suspected heritable disorders. They were able to reduce time to result to 13.5 hours (unpublished). They start off with a typical dried blood spot punch/punches from the archived California dried blood spots. The first step is to prepare the DNA for genome sequencing which is now possible in 80 minutes. There is a new prep that involves DNA extraction and library preparation as a single automated protocol with minimal hands-on time by a laboratory specialist. The genome sequencing is then run to 100 nucleotides by stripping

away the exome steps. The library is immediately put onto a sequencer, and using the latest generation of Illumina instruments, it can now be completed in less than 12 hours. The Dragon Platform uses hyper threading to expedite the computational side of the genome-wide variant calling which currently takes about 45 minutes. It scales with the full genome coverage. In addition, since most of the children are in an intensive care unit and have a medical record, it has useful information in terms of narrowing the differential diagnosis. Natural language processing is used to extract from nonstructured text, the clinical features that summarize a child's illness and on average there are 80 to 250 clinical features describing that child's illness.

Finally, throughout the project, the researchers have prototyped and clinically validated the use of artificial intelligence to take those two inputs- roughly five million genomic variants and 150 clinical features, and parse that against approximately 6,000 genetic diseases- and in about ten minutes and provide an automated provisional diagnosis. This underscores that this technology increasingly is ready for the sickest of patients and ready for consideration technically from the newborn screening perspective.

There are two types of rapid genome sequencing: rapid and ultra-rapid. They correspond to time critical conditions and other conditions. There is no doubt that rapid whole genome sequencing meets the timeliness goals.

NSIGHT2 was a randomized group of 213 infants in neonatal intensive care units with a critical illness of unknown etiology at the time of admission. Enrollment was purposely limited to the first 96 hours of admission with a goal to have genome sequencing done as a first-tier test to minimize that window of abbreviated empirical treatment and to treat disorders from rapid disease progression.

The diagnostic rate varied dependent on the modality. It was 21 percent for rapid genome or exome. It was 46 percent for ultra-rapid because there is different selection criteria in those two groups. And those positive results were returned in a median of two days. Disease questionnaires and parental questionnaires provided insight into the usefulness of the information. Questions asked included: Does it change management? Does it change outcomes? The answer was yes because for the vast majority of positive results, it changes management and dramatically in 39 percent it changed the outcome according to physician perception. When it came to parental perceptions, the results indicated, no concerns in terms of anxiety or depression as a result of returning genomic results in an ICU setting.

The utility of negative results was surprising. Genome sequencing is very different than a screening for selected conditions. It's very different from exome sequencing because the entire genome is decoded and the entire gene region is scanned, and can rule out specific differentials. This, therefore, has huge negative predictive value in terms of a physician's perception of whether the child has a genetic disease. Seventy-two percent of clinicians reported that negative results had clinical utility. In 16 percent, it changed clinical management.

Finding that a genetic disorder was ruled out, was very good news to most parents and in many instances changed some of the decision-making in terms of care intensity.

There are five participating sites in the Baby Bear project with 183 infants enrolled. The intent of this study was to explore clinical utility and economic value in newborns with Medi-Cal insurance. The diagnosis rate was similar to other studies, approximately 43 percent. The proportion who had a change in care was also similar to previous studies, 31 percent. The median time to obtain results was three days.

There were significant cost savings both for the hospitals and for Medi-Cal with the three-day turn-around time. Dr. Kingsmore's team has modeled costs for a 7 and 14 -day turn-around time. The hospital would break even with a 7 day turn-around and would experience an increase in cost in a scenario where the results would take 14 days. Dr. Kingsmore's research appears to demonstrate that cost savings are not limited to the initial hospitalization and continue to accrue into subsequent years. It is clear that there is a net cost savings for Medi-Cal.

In summary, Dr. Kingsmore believes that the Wilson-Jungner criteria could be applied by whole genome sequencing. It becomes allele driven rather than disorder driven, and as was with traditional newborn screening, an entire system of delivery is needed and the sole focus needs to be on the diagnostic piece where the screening piece will not be sufficient. We need a strong focus on immediate implementation of interventions.

Robert C. Green, MD, MPH
Professor of Medicine, Harvard Medical School Director,
Genomes2People Research Program /Preventive Genomics Clinic Brigham Health
Broad Institute • Ariadne Labs • Harvard Medical School

Dr. Green directs the Genomes2People Research Program where he is accelerating the implementation of genomic and precision medicine. He discussed the BabySeq project, and how this might relate to global change in the health care system towards preventive genomics.

From the beginning, this was designed as a randomized trial that had two parts. It started off with a randomized trial of newborns in the ICU. But unlike Dr. Kingsmore's work, babies that were most likely to benefit from sequencing were not selected. The team was curious as to whether if one sequenced all babies, one would find unanticipated information. The answer, in the small sample size was that it was not a good idea. The team didn't find enough in a random selection of NICU babies so they focused on a randomized trial of healthy newborns.

Early on, the researchers decided to focus on childhood and adolescent onset. The team looked at the disease gene associations that had strong or moderate evidence; primarily those with high or moderate penetrants; and they also utilized a broad definition for actionablility. Later on the team added a very limited number of adult-onset conditions, mostly associated with the ACMG 59. The team reported only pathogenic and likely pathogenic variants. They did not report VUS as they often get reclassified in a benign direction.

The Babyseq team first took about 1,500 genes, separated out the ones they felt had strong or definitive evidence and then further separated out the ones they felt had high or moderate penetrants and the ones that had some age of onset below 18 years old. Finally, the team had

954 genes that met the reporting criteria. This became the BabySeq list. Other countries have been very curious about this because there was a very logical, rational way in which the list was developed.

The demographics of the enrolled parents were primarily white and primarily non-Hispanic. It is a challenging time to recruit parents (on the newborn unit in the first 24 hours after they have delivered.) There was a lengthy consent process and participation involved a blood draw. There were many parents who were not interested in research for a variety of reasons. Many parents cited privacy, discrimination and study logistics. Sixty percent of people who agreed to hear a description of the study agreed to participate.

Eleven percent of healthy babies had an unanticipated monogenic (i.e. dominant heterozygote or a bi-allelic recessive) set of findings. For about a quarter of the 11 percent, in light of the DNA evidence and reexamining the child, the team found evidence that the condition in question was already underway. There was already an abnormality. In another set of five individuals the team found family history that had not been reported. Eleven percent was higher than expected but not as high as what was found in MedSeq where there was different criteria including no limits set on actionability.

There is neither standard prenatal care nor commercial panels picking up the conditions/trait status in ostensibly healthy newborn babies. In terms of recessive carrier traits from about 566 genes. Eighty-eight percent of infants had at least one heterozygous recessive gene. In most cases, the parents of reproductive age did not know they were carriers. Among the 159 babies in the sequencing arm, there was 1 baby found to be an unaffected recessive carrier. The parents were checked and found they both were carries and used this information to use reproductive technologies for their next baby.

Parents in the control group and the group that participated in exome sequencing showed no differences in clinical rates of depression, anxiety, self-blame, or disruptions in bonding with their baby.

A substantial effort was spent on trying to assess the downstream medical impact. Of 17 newborns with unanticipated genomic findings, there were 43 services ordered. Those broke down into 23 specialty visits and 20 other resources. There was no statistically significant difference in costs between the randomized and control groups. A larger sample size is needed to address whether or not cost is a true concern with exome sequencing. A key remaining question is if a baby is sequenced at birth could the information be used to reduce morbidity and mortality and realize a cost savings throughout the lifespan.

Discussion

- Dr. Powell asked Dr. Currier the first question: Since the data is fairly clear that sequencing would not be a good substitution as a first-tier test for inborn errors of metabolism, could it potentially be used as a second or third tier test, to reduce false positive results.

- Dr. Currier responded that there is potential for two reasons: 1) the number of people who would need a second-tier test after primary MS/MS screening is much smaller than the whole population which makes the economic impact smaller; and 2) with increasingly rapid turnaround, it suggests that even in time-critical disorders, a DNA-based second tier test might be highly effective if it could be immediately incorporated into the screening workflow.

- A Committee member asked Dr. Kingsmore to elaborate on turn around time and through put.
  - Dr. Kingsmore responded that in terms of secondary testing, there is mounting evidence that sequencing is highly effective for babies with immune deficiency disorders. Genomic sequencing reduces the need for multiple tests to identify the molecular diagnosis in a highly time-sensitive setting.
  - In terms of through put: it is a rapidly evolving approach. There are now studies of half a million to a million being done in specific populations, most notably Genome England, with their half million cases.
- A Committee member asked the project leads to describe what they would have done differently and what is the next logical step for this type of research.
  - Dr. Berg said he would advocate for a study that pursues more targeted sequencing.
  - Dr. Currier wondered to what extent is a molecular diagnosis a diagnosis? There is a place for additional research on instances when variants are pathogenic in combination or pathogenicity for diplotypes for Mendelian recessive disorders.
  - Dr. Kingsmore responded that the NSIGHT RFA was extremely cutting-edge. His biggest regret is that only at the end of the award did their research team identify the appropriate follow-on project and he would have liked to have done additional research. .
  - Dr. Green indicated that he would have eliminated the NICU hypothesis from the beginning because it was incorrect. They also wished they could have had a larger sample size, and streamlined the logistics of both analysis and recruitment. In future research, they hope to expand parental recruitment within underrepresented populations.
- A Committee member directed a comment to Dr. Green, stating that he thought the biotinidase deficiency case identified was a partial case and indicated that clinical diagnosis and public health criteria may not align.
  - Dr. Green concurred and stressed that they wrote a paper on that one case because it was picked up by newborn screening. The newborn screening was repeated and the results were normal. The baby was started on a vitamin biotin and is doing well but may have never needed an intervention. This shows how difficult this task is. The family was included in every step and they are grateful to have the information.
- An Organizational Representative asked Dr. Berg and Dr. Green about counseling families for reoccurring risk. He wondered if they had learned from their published and unpublished studies practices for enhanced prediction of outcomes in a planned pregnancy.
  - Dr. Green responded that he thinks risk factor modeling and mental modeling, is better because we will gradually understand not only the other genetic factors, like

polygenic on top of monogenic but also post translational modification, metabolomics etc. that will gradually provide additional insight into nuances.

- Dr. Berg agreed with Dr. Green and added that until we understand all of the other genetic and non-genetic factors included in risk predication we will continue to rely on a Mendelian versus polygenetic framework.

VI. New Business

There was no new business.

VII. Adjourn

The next meeting will be a webinar scheduled for February 11 – 12, 2021.