Improving the Neonatal Research Network Annual Certification for Neurologic Examination of the 18-22 Month Child

Jamie E. Newman, PhD[1], Carla M. Bann, PhD[1], Betty R. Vohr, MD[2], Anna M. Dusick, MD[3], and Rosemary D. Higgins, MD[4], on behalf of the Follow-Up Study Group of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Neonatal Research Network*

Objective To describe the Neonatal Research Network's efforts to improve the certification process for the Follow-Up Study neurologic exam and to evaluate inter-rater agreement before and after two annual training workshops.

Study design The Neonatal Research Network Follow-Up Study is a multi-center observational study that has examined more than 11 500 infants from 1998-2010 and born ≤26 weeks gestational age at 18-22 months corrected age for neurodevelopmental outcome. The percentages of examiners who agreed with the Gold Standard examiner on 4 neurodevelopmental outcomes on the initial training video and a test video were calculated. Consistency among examiners was assessed with the first-order agreement coefficient statistic.

Results Improvements in agreement among examiners occurred between 2009 and 2010 and between initial training and test. Examiner agreement with the Gold Standard during the initial training was 83%-91% in 2009 and 89%-99% in 2010. Examiner agreement on the workshop test video increased from 2009-2010 with agreement reaching 100% for all four neurodevelopmental outcomes examined in 2010. First-order agreement coefficient values for the four neurodevelopmental outcomes on the training videos ranged from 0.64-0.82 in 2009 and 0.77-0.97 in 2010.

Conclusions We demonstrate the importance of annual certification and the benefits of evaluation and revision of certification protocols to achieve high levels of confidence in neurodevelopmental study outcomes for multi-center networks. *(J Pediatr 2012;161:1041-6)*.

Multi-center trials encounter challenges of standardizing research procedures across sites. It is important to periodically examine center differences, particularly for primary outcomes, to assess whether research procedures are being implemented uniformly. Periodical review of examiner training procedures can aid in identifying areas for improvement.

Examiner training for research protocol purposes is often achieved using video vignettes and, to a lesser extent, the examination of actual patients during centralized training. Examiner training generally improves scoring accuracy and inter-rater reliability. For example, examiner training improved the motor development classification of infants as normal or abnormal[1] as well as the reliability of upper limb function scores.[2] However, there are instances where training has not improved inter-rater reliability and accuracy.[3] Consideration has also been given to whether training and certification procedures are valid and reliable for novice examiners in addition to expert examiners.[4]

Training multiple examiners in a network for inter-rater reliability for the diagnosis of cerebral palsy (CP) and neurologic status can be particularly challenging due to substantial heterogeneity of neurologic findings ranging from very mild to severe, which is further complicated because young children may not clearly fall into a specific category. Examining the research standardization and training procedures of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) Neonatal Research Network (NRN) Follow-Up Study may be beneficial for those conducting neurodevelopmental outcomes research in neonatal medicine as well as other fields where annual examiner certification is required. The NRN Follow-Up Study evaluates infants that are born ≤26 completed weeks gestational age at 18-22 months corrected age for neurodevelopmental outcome. The Follow-Up Study assessment battery includes: demographic and medical history, physical examination, a standardized neurologic

AC1	First-order agreement coefficient
CP	Cerebral palsy
GMFCS	Gross Motor Function Classification Score
NDI	Neurodevelopmental impairment
NICHD	*Eunice Kennedy Shriver* National Institute of Child Health and Human Development
NRN	Neonatal Research Network

From the [1]Department of Statistics and Epidemiology, RTI International, Research Triangle Park, NC; [2]Department of Pediatrics, Women and Infants' Hospital, Brown University, Providence, RI; [3]Department of Pediatrics, Indiana University School of Medicine, Indianapolis, IN; and [4]*Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD

*A list of members of the Follow-Up Study Group of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development Neonatal Research Network is available at www.jpeds.com (**Appendix**).

The National Institutes of Health and the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) provided grant support for the Neonatal Research Network's (NRN) Follow-Up Study. Data collected at participating sites of the NICHD NRN were transmitted to RTI International, the data coordinating center for the network, which stored, managed, and analyzed the data for this study. J.N.'s time was supported by the NICHD Neonatal Research Network, as the Data Coordinating Center coordinator for the Follow-Up Study. The authors declare no conflicts of interest.

 http://dx.doi.org/10.1016/j.jpeds.2012.05.048

examination,[5] the Palisano et al [6,7] Gross Motor Function Classification Score (GMFCS), a developmental assessment using the Bayley Scales of Infant and Toddler Development, 3rd edition,[8] and a behavior assessment using the Brief Infant-Toddler Social Emotional Assessment.[9] From 1998 to 2010, more than 11 500 infants completed the 18 - 22 month Follow-Up Study visit.

Our objective is to describe the NRN's implementation of a protocol to improve the certification process for the neurologic exam by NRN center examiners by evaluating inter-rater agreement before and after the annual training workshops in 2009 and 2010. It was hypothesized that inter-rater agreement would improve between initial training and test as well as between 2009 and 2010.

Methods

The NICHD NRN maintains a study manual with standard definitions (Table I). A diagnosis of neurodevelopmental impairment (NDI) is implemented for both randomized trials and observational follow-up studies. The definition of NDI is the presence of any of the following: moderate to severe CP, bilateral blindness, bilateral hearing impairment requiring amplification, (GMFCS)[6,7] ≥level 2 reflecting moderate to severe motor impairment, or a Bayley Scales of Infant and Toddler Development 3rd edition cognitive or motor score <70. The NRN Follow-Up Study neurologic examination[5] includes a standard assessment of reflexes, muscle tone and strength, functional motor skills, hearing, and vision. Assessment of vision and hearing is completed by: (1) reviewing history with family; (2) review of records from subspecialty ophthalmology and audiology services; (3) determining if child is receiving services for the blind, deaf educator, speech/language, or sign language, or has amplification/FM system; (4) examination of the eyes including tracking, strabismus, glasses, and surgical procedures; and (5) determining if the child can follow a simple verbal direction provided by the examiner. Hospital records, for the most part, contain this information and parents are usually aware of the formal diagnosis of blindness or permanent hearing loss.

Based on the neurologic examination, a determination is made of whether the child is normal or abnormal, and whether the child has CP as well as the severity of CP. CP is defined with the following three criteria: (1) definite abnormalities observed in the classic neuromotor exam, which includes measurement of tone, deep tendon reflexes, coordination, and movement. Any one definite abnormality in the classic neuromotor exam, as defined, except for isolated low tone (hypotonia) or toe walking without tight ankles is sufficient; (2) a delay in motor milestones with a disorder of motor function must be present. This may or may not be reflected in a motor quotient <70. In mild cases, there may be a subtle difference in hand functioning with a fine pincer grasp in 1 hand and a raking grasp in the other hand. Some disorder of motor function must be present; and (3) aberrations in primitive reflexes and postural reactions may be present.

The hierarchical classification of CP subtypes[10] is included in the annual examiner training, with the aim of achieving inter-rater reliability in all subtypes including spastic, dyskinetic, and hypotonic CP. An assessment of gross motor function level is completed based on the Palisano et al[6,7] GMFCS to provide a standardized classification of the severity of motor disability. This method is useful for assessing a level of severity of CP.[6,7,11] Bimanual hand function is assessed in 2 ways. First, in a 3-level system that includes: (1) no problems with bimanual task; (2) some difficulty with bimanual tasks; and (3) no functional bimanual tasks. Second, in a 5-level assessment of each hand that ranges from 1-fine pincer grasp to 5-unable to grasp/makes no attempt to grasp.

Examiners

Since 2005, there have been 16 participating NRN centers. Each center designates a primary neurologic examiner who is generally the Follow-Up Study Principal Investigator of the center, based on the investigator' expertise in the field

Table I. NICHD NRN 2009-2010 definitions for the 18-22 months neurologic exam

Parameter	Definition
Normal	No abnormalities of tone, reflexes, or function noted in the neurologic exam.
CP	Abnormalities observed in tone, deep tendon reflexes, coordination, and movement. Some disorder of motor function must be present. Aberrations in primitive reflexes and postural reactions may be present.
GMFCS	Normal - walks 10 steps independently and fluently. Possible Level 1 - walks 10 steps independently but not fluently; child exhibits toe walking (mild diplegia) or asymmetric walking (possible hemiplegia). Level 1 - infants move in and out of sitting and floor-sit with both hands free to manipulate objects. Infants creep or crawl on hands and knees, pull to stand and take steps holding onto furniture. Infants walk between 18 mo and 2 y of age without holding on. Level 2 - infants maintain floor sitting but may need to use their hands for support to maintain balance. Infants creep on their stomach or crawl on hands and knees. Infants may pull to stand and take steps holding onto furniture. Level 3 - infants maintain floor sitting when the low back is supported. Infants roll and creep forward on their stomachs. Level 4 - infants have head control but trunk support is required for floor sitting. Infants can roll to supine and may roll to prone. Level 5 - physical impairments limit voluntary control of movements. Infants are unable to maintain antigravity head and trunk postures in prone and sitting. Infants require adult assistance to roll.
NDI	The presence of any one of the following: moderate to severe CP with a GMFCS score of ≥2, bilateral blind with vision 20/200 or worse, bilateral permanent hearing loss ± amplification that does not permit the child to understand instructions of the examiner, or a Bayley Scales-III cognitive or motor score <70.

Bayley Scales-III, Bayley Scales of Infant and Toddler Development 3rd edition.