A Population-Based Study to Determine the Performance of the Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale to Predict the Mental Developmental Index at 18 Months on the Bayley Scales of Infant Development-II in Very Preterm Infants

Michael J. Vincer, MD*; Heather Cake, MA‡; Michael Graven, MD*; Linda Dodds, PhD§;
Shelly McHugh, BScN‖; and Theresa Fraboni, MD*

ABSTRACT. *Objectives.* To determine optimal ages to perform the Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS) and optimal "cutoff" score of the CAT/CLAMS to screen very preterm infants (<31 weeks) for severe cognitive-adaptive delay and to ascertain the sensitivity, specificity and likelihood ratios using optimal cutoff scores compared with the Mental Developmental Index (MDI) of the Bayley Scales of Infant Development II.

Methods. A population-based cohort of very preterm infants who were born to mothers who resided in Nova Scotia or Prince Edward Island were evaluated at 4, 8, 12, and 18 months' corrected gestational age, which included a CAT/CLAMS by a physician. At 18 months' corrected gestational age, each child was assessed using the Bayley Scales of Infant Development II, the "gold standard" for developmental delay in young infants. The results of each CAT/CLAMS was compared with the 18-month MDI to identify significant developmental delay (MDI <70).

Results. Optimal scores on the CAT/CLAMS to identify correctly MDI <70 were determined by using the κ statistic for chance independent agreement. Sensitivities and specificities for optimal cutoff scores were as follows: 4-month score <109 (88% and 37%), 8-month score <98 (75% and 82%), 12-month score <81 (63% and 99%), and 18-month score <83 (88% and 98%).

Conclusion. Sensitivity and specificity of the CAT/CLAMS are high in very preterm infants at identifying major developmental delay at 12 and 18 months. For follow-up programs without psychology services, the CAT/CLAMS at 12 and 18 months is a reasonable screening tool to determine which children need expedited psychology referral for cognitive delay. *Pediatrics* 2005;116:e864–e867. URL: www.pediatrics.org/cgi/doi/10.1542/peds.2005-0447; *developmental screening, preterm infants, Capute scales, neurodevelopmental outcome.*

ABBREVIATIONS. MDI, Mental Developmental Index; PFUP, Perinatal Follow-Up Program; CAT/CLAMS, Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale; BSID, Bayley Scales of Infant Development; LMP, last menstrual period.

The rate of severe cognitive delay (Mental Developmental Index [MDI] <70) at 18 months' corrected gestational age in infants <31 weeks' gestational age, as based on data from the Perinatal Follow-Up Program (PFUP) of Nova Scotia 1997–1999, is 10.7%. In a meta-analysis, Lorenz et al[1] noted rates of developmental delay as high as 23% for the extremely immature infant. With such high rates of developmental delay, it is essential that good screening tools be available to identify cognitive delay. One such tool is the Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS), which was developed >20 years ago.[2] The CAT/CLAMS is a simple screening tool for pediatricians to learn and generally takes <10 minutes to administer, even by pediatric residents.[3] Correlation coefficients comparing the CAT/CLAMS index score with the Bayley Scales of Infant Development II (BSID-II) MDI have been reported to be as high as 95%,[4] with sensitivities for identifying MDI <70 to be 88%[4] and specificity to be 97%.[5]

It has been suggested that the pediatrician is in an important position to conduct early screening of children for developmental delay so that appropriate early intervention services can be initiated.[6] In addition, the American Academy of Pediatrics has recommended that pediatricians acquire skills in administering and interpreting reliable and valid developmental screening tools.[7]

Previous research has evaluated the CAT/CLAMS for (1) children with suspected developmental delay,[4,8,9] (2) follow-up for infants who recovered from aseptic meningitis,[10] (3) the evaluation of normal infants recruited from a well-infant clinic,[11] and (4) a high-risk follow-up clinic.[5] The results obtained from a high-risk follow-up clinic prompted us to examine our own experience of using the CAT/CLAMS in a follow-up population. Despite that Macias et al[5] identified their population as being randomly selected from a consecutively enrolled cohort that had a mean gestational age of 31 weeks and a mean birth weight of 1694 g in their follow-up clinic, it is unclear whether the inception cohort was from a population-

Birth weight, median (range)	1200 (465–2045)
Gestational age, median (range)	28 (23–30)
Single-parent families	35/171 (20%)
White	84/87 (97%)
Mothers with less than high school	25/163 (15%)
Low socioeconomic status	50/155 (32%)

based cohort or there were any specific enrollment criteria for their follow-up clinic.

Macias et al[5] noted in their work 2 additional features that are addressed further in this article: (1) an older age at performing the CAT/CLAMS is more predictive for the MDI on the BSID-II, and 2) CAT/CLAMS index scores are not equivalent to the MDI scores. Macias showed that using a higher cutoff score on the CAT/CLAMS resulted in a greater sensitivity with little loss in specificity at predicting MDI scores <70.

The objective of the present study was to determine the optimal age to perform the CAT/CLAMS and the optimal "cutoff" score to identify a birth cohort of infants of <31 weeks with MDI scores <70. Once cutoff scores are determined, the sensitivity, specificity, and likelihood ratios are ascertained for these scores.

METHODS

All infants who were <31 weeks' gestational age; born to mothers who resided in Nova Scotia or Prince Edward Island, Canada, at the time of birth; and born between November 22, 2000, and May 25, 2003, were enrolled in the PFUP of Nova Scotia. Each infant was evaluated by a pediatrician or a neonatologist (or by a member of the housestaff under the supervision of a pediatrician or a neonatologist) at 4, 8, 12, and 18 months' corrected gestational age. In addition to a history and physical examination (including a neurologic examination), each assessment included the CAT/CLAMS. The CAT/CLAMS performed at 4, 8, and 12 months' corrected age were conducted before the 18-month MDI; thus, they were blinded to MDI scores. Most (~70%) of the 18-month CAT/CLAMS scores were conducted in a similarly blinded manner. The total composite CAT/CLAMS score was determined by the age-equivalent level of function divided by the corrected gestational age multiplied by 100 to give an index score as recommended by Capute and Accardo.[12,13] Thus, if the child was functioning at his or her exact corrected gestational age, then the score would be 100.

Gestational age was assigned on the basis of 1 of the following 3 criteria: conception date, last menstrual period (LMP), early ultrasound, or a clinical assessment (either a Ballard assessment[14] performed by the housestaff or clinical estimate performed by the attending neonatologist). The conception date was used over all other estimates, but this was available in only 4% of preterm infants. The ultrasound dates were used over the LMP when the difference between the ultrasound and the LMP was >10 days as suggested by Gray et al[15]; otherwise, dates were used. When no maternal dates were available, ultrasound was used. The only time a Ballard score or other clinical assessment was used is when no other reliable dating was available.

At 18 months' corrected gestational age, the developmental psychologist administered the BSID-II, to assess for cognitive delay, which was considered the "gold standard." By comparing each child's CAT/CLAMS score with the MDI at 18 months' corrected gestational age, receiver operator curves were constructed for each age (4, 8, 12, and 18 months' corrected age) to estimate the area under the curve. To determine the optimal CAT/CLAMS score that predicted an MDI <70, the κ statistic measuring chance independent agreement between the CAT/CLAMS and the MDI <70 was calculated for each CAT/CLAMS score, and the highest κ was used.

RESULTS

From November 22, 2000, to May 25, 2003, 172 infants who were <31 weeks' gestational age were born to Nova Scotia or Prince Edward Island mothers and survived to follow-up. Because all surviving infants who are <31 weeks are invited to be enrolled in the PFUP, this represents a complete geographically defined population-based cohort of infants who are <31 weeks' gestational age. Table 1 gives some basic population characteristics of these children.

A CAT/CLAMS was performed at 4, 8, 12, and 18 months' corrected gestational age. When the CAT/CLAMS was started in the PFUP in June 2002, some children were already beyond the 4-, 8-, or 12-month ages; thus, we did not have CAT/CLAMS performed on them at these ages.

Of the 172 infants who were included in the study, 147 (85% of the eligible population) had both an MDI of the BSID-II conducted by a developmental psychologist and a CAT/CLAMS conducted by the physician at 18 months' corrected gestational age. Eight had only the CAT/CLAMS (median score: 91; range: 63–104), 4 had only the MDI (median score: 82; range: 67–103), and 13 had neither. Of the 13 with neither score, 4 had scores for both performed at 24 months, and 3 would have been identified correctly as normal and 1 incorrectly would have been called abnormal (false positive). One other child had a normal CAT/CLAMS at 24 months and a normal Differential Ability Scales at 3 years, and 1 other child had such severe disability from cerebral palsy that the MDI was impossible to obtain. None of these additional results would have had a significant impact on the 18-month outcomes noted in Table 2.

A receiver operator curve for each of the 4 ages in determining the MDI <70 is shown in Fig 1. Not surprising, the area under the receiver operator curves increased with age. The area and 95% confidence intervals were 56% (41%–71%) at 4 months, 82% (71%–93%) at 8 months, 88% (79%–98%) at 12 months, and 97% (94%–100%) at 18 months. Optimal cutoff CAT/CLAMS as determined by the highest κ statistic were <109 at 4 months, <98 at 8 months,

TABLE 2. Sensitivity and Specificity at the Optimal Cutoff Scores to Identify the MDI <70 for the Corrected Gestational Ages at Which the CAT/CLAMS Was Performed

CAT/CLAMS Age (Optimal Cutoff)	Sensitivity	Specificity	Likelihood Ratio for a Positive Test	Likelihood Ratio for a Negative Test
4 mo (score <109)	88% (15/17)	37% (23/63)	1.4	0.3
8 mo (score <98)	75% (15/20)	82% (63/77)	4.1	0.3
12 mo (score <82)	64% (14/22)	98% (93/95)	30.2	0.4
18 mo (score <84)	88% (22/25)	97% (119/122)	35.8	0.1
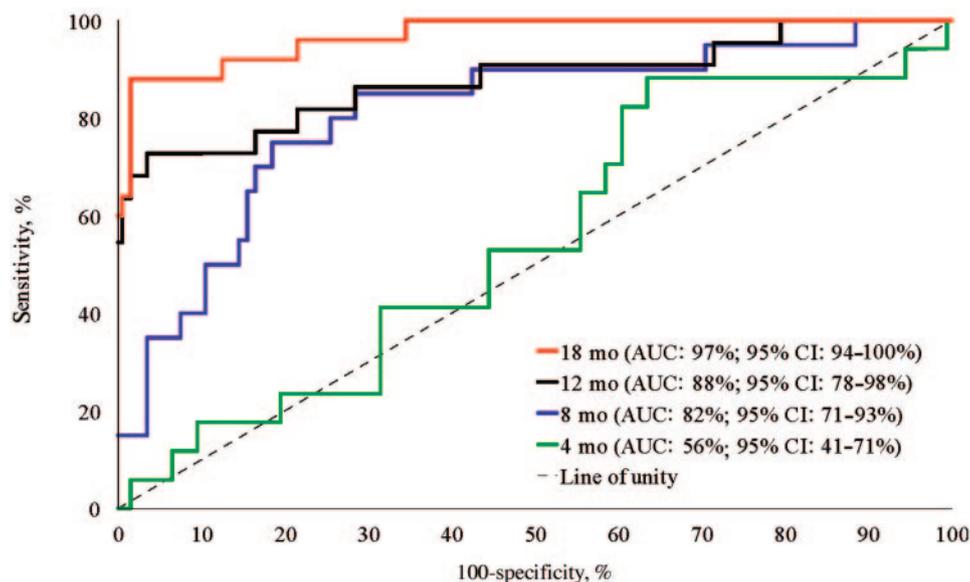
Fig 1. Receiver operator curves for each age of the CAT/CLAMS scores diagnosing MDI <70. (The point closest to the left upper hand corner for each age represents the highest "agreement" between the CAT/CLAMS and the MDI and thus gives the optimal cutoff values of the CAT/CLAMS to predict the MDI <70.)

<82 at 12 months, and <84 at 18 months. These values concurred with visual inspection of the receiver operator curves. Table 2 confirms that 12 and 18 months have the highest likelihood ratios for a positive test with the lowest values for a negative test; therefore, they are the best ages to identify the MDI score <70.

DISCUSSION

The current study has a number of strengths: (1) use of a complete population-based cohort of very preterm infants, (2) development of "optimal cutoffs" rather than arbitrary cutoffs for the CAT/CLAMS to predict the MDI, and (3) use of >1 age to determine the optimal ages to perform the CAT/CLAMS in very preterm infants. It could also be considered a strength that the current study used the BSID-II, the current standard of determining cognitive delay in infants. This helps redress a limitation in the current medical literature because the CAT/CLAMS was originally compared with the BSID-I (the 1969 version). The BSID-I is known to give different MDI scores than the BSID-II (the 1993 version); for example, in a population of children with Down syndrome, the BSID-I tended to give MDI scores ~8 to 9 points higher than the BSID-II.[16]

An area of ongoing controversy that may affect the interpretation of the current data is correction for prematurity. Some investigators stop correcting after 2 years,[17] whereas others continue until 8.5 years.[18] It might be argued that for infants at 4 and 8 months' corrected age, there is a proportionally greater degree of age correction, making the CAT/CLAMS scores "overcorrected." This may explain, to some extent, why CAT/CLAMS scores at these ages are poor predictors of the 18-month MDI. We found at 4 months' corrected age that an uncorrected CAT/CLAMS score of 61 performed marginally better than the corrected scores with a sensitivity of 71% and a specificity of 63% at predicting the 18-month MDI <70. This was not true at other ages.

Another issue to be considered is that the derivation of the CAT/CLAMS and the MDI scores is different. The CAT/CLAMS score is a simple ratio of the developmental age level ÷ corrected age × 100, whereas the MDI of the BSID is indexed to a Gaussian distribution so that 1 SD about a mean score of 100 is ±15 points. This means that a score of 85 for a child of 18 months (16th percentile) would be a developmental level of ~16 months, whereas a score of 70 (2nd percentile) would be 14 months. The CAT/CLAMS scores at these developmental performances would be 89 and 78, respectively. Use of the receiver operator curve allows us to identify the "optimal" CAT/CLAMS score to predict the MDI <70 (Fig 1).

A minor limitation of the present data is that a small number (14%) of CAT/CLAMS scores at 18 months were derived from the items performed by the developmental psychologist on the same day, which are in common with the BSID-II. To address this issue, we changed the schedule of testing so that the CAT/CLAMS was performed 2 weeks after the BSID-II. An additional refinement by this separation was that all physicians remained blinded to MDI scores until the CAT/CLAMS was performed.

Our study performed much better than Macias et al,[5] who found a very low sensitivity and a very high specificity when using the CAT/CLAMS as a screening tool to identify infants who have MDI scores <70. Our population confirmed the high specificity of the CAT/CLAMS; however, we also noted a relatively high sensitivity at 12 and 18 months' corrected gestational age in contrast to data by Macias et al. We found a similar trend of improved predictability of performing the CAT/CLAMS at higher ages at identifying children with developmental delay (see Table 2). In addition, the data of Macias et al suggested that a score of 70 on the CAT/CLAMS may

not be optimal at identifying children with MDI scores <70, an observation supported by Capute et al,[19] who examined the ability of an earlier version of the CLAMS component in which a score of 80 on the CLAMS was comparable to 70 on the MDI. We took this 1 step further by determining the test characteristics of the "optimal" scores at various ages of the CAT/CLAMS at identifying MDI score <70 at 18 months' corrected age. Our data suggested that the CAT/CLAMS tends to score higher than the MDI such that it did not make sense to assume that the CAT/CLAMS <70 was best at predicting the MDI <70. If we had done so at 18 months, then the sensitivity would be only 56% with a specificity of 100%. The linear regression curve ($r^2 = 0.74$) of the 18-month CAT/CLAMS for the 18-month MDI is $MDI_{18} = 0.9 \times CAT/CLAMS_{18} - 2.4$, where MDI_{18} is the 18-month corrected age MDI score and $CAT/CLAMS_{18}$ is the 18-month corrected age CAT/CLAMS score. Therefore, the CAT/CLAMS score is ~10% higher than the corresponding MDI score. It was not the intent of this study to adjust the MDI score for variables other than the CAT/CLAMS; however, in addition to the 18-month CAT/CLAMS, temperature on admission to the NICU accounted for a small amount of variance when entered in an adjusted model. Admission temperature adjusted the MDI downward by ~4 points for every degree below 37°C.

To summarize the performance at the 2 optimal ages, the CAT/CLAMS at 12 months moves the pretest probability of 10.7% (as in our population) to a posttest probability of 78% when the score is <82 or to a posttest probability of 4% when the score ≥82. At 18 months, the posttest probability moves to 81% when the CAT/CLAMS score is <84 or to 2% when the CAT/CLAMS is ≥84.

We believe that the CAT/CLAMS may be beneficial to some follow-up programs that do not have access to developmental psychology services on a routine basis. With the use of this simple tool, the majority of infants with developmental delay can be identified and thus give the physician the rationale for an expedited referral to a psychologist. It is also very useful when the CAT/CLAMS score is normal because this allows the physician to be reassuring to the family with regard to early cognitive delay.

Another aspect for some programs without psychology services is whether the economics justifies hiring a psychologist to provide a diagnostic service for developmental delay in the preterm infant. As an example, in Nova Scotia, the psychologist usually takes 2 hours to perform and interpret the BSID-II at 18 months' corrected age. For infants <31 weeks, there are ~60 such assessments per year. At conservative private rates of ~$500 (Canadian) per BSID assessment, the total annual expenditure of $30 000 would be required for this service. On the basis of the data presented in this study, it is estimated that the physician could reduce the need from 60 psychology referrals per year to ~10, costing ~$5000 per year (10 × $500). There may be a slight increase in physician time spent preforming the CAT/CLAMS; however, most follow-up physicians conduct a developmental assessment but generally not a standardized tool such as the CAT/CLAMS. By substituting the time spent performing a nonstandardized tool with the CAT/CLAMS, the cost for additional physician time is insignificant.

REFERENCES

1. Lorenz JM, Wooliever DE, Jetton JR, Paneth N. A quantitative review of mortality and developmental disability in extremely premature newborns. *Arch Pediatr Adolesc Med.* 1998;152:425–435
2. Capute AJ, Palmer FB, Shapiro BK. The Clinical Linguistic and Auditory Milestone Scale of infancy (CLAMS): initial validation. *Dev Med Child Neurol.* 1983;25:115
3. Pittock ST, Juhn YJ, Adegbenro A, Voigt RG. Ease of administration of the Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS) during pediatric well-child visits. *Clin Pediatr (Phila).* 2002;41:397–403
4. Hoon AH, Pulsifer MB, Gopalan R, Palmer FB, Capute AJ. Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale in early cognitive assessment. *J Pediatr.* 1993;123:S1–S8
5. Macias MM, Saylor CF, Greer MK, Charles JM, Bell N, Katikaneni LD. Infant screening: the usefulness of the Bayley Infant Neurodevelopmental Screener and the Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale. *J Dev Behav Pediatr.* 1998;19:155–161
6. King TM, Glascoe FP. Developmental surveillance of infants and young children in pediatric primary care. *Curr Opin Pediatr.* 2003;15:624–629
7. American Academy of Pediatrics. Developmental surveillance and screening of infants and young children. *Pediatrics.* 2001;108:192–195
8. Kube DA, Wilson WM, Petersen MC, Palmer FB. CAT/CLAMS: its use in detecting early childhood cognitive impairment. *Pediatr Neurol.* 2000;23:208–215
9. Wachtel RC, Shapiro BK, Palmer FB, Allen MC, Capute AJ. CAT/CLAMS: a tool for the pediatric evaluation of infants and young children with developmental delay. *Clin Pediatr (Phila).* 1994;33:410–415
10. Rossman MJ, Hyman SL, Rorabaugh ML, Berlin LE, Allen MC, Modlin JF. The CAT/CLAMS assessment for early intervention services. Clinical Adaptive Test/Clinical Linguistic and Auditory Milestone Scale. *Clin Pediatr.* 1994;33:404–409
11. O'Connor-Leppert ML, Shank TP, Shapiro BK, Capute AJ. The Capute scales: CAT/CLAMS—a pediatric assessment tool for the early detection of mental retardation and communicative disorders. *Ment Retard Dev Disabil Res Rev.* 1998;4:14–19
12. Capute AJ, Accardo PJ. The infant neurodevelopmental assessment: a clinical interpretive manual for CAT-CLAMS in the first two years of life, part 1. *Curr Probl Pediatr.* 1996;26:238–257
13. Capute AJ, Accardo PJ. The infant neurodevelopmental assessment: a clinical interpretive manual for CAT-CLAMS in the first two years of life, Part 2. *Curr Probl Pediatr.* 1996;26:279–306
14. Ballard JL, Khoury JC, Wedig K, Wang L, Eilers-Walsman BL. New Ballard score, expanded to include extremely premature infants. *J Pediatr.* 1991;119:417–423
15. Gray PH, Jones P, O'Callaghan MJ. Maternal antecedents for cerebral palsy in extremely preterm babies: a case-control study. *Dev Med Child Neurol.* 2001;43:580–585
16. Glenn SM, Cunningham CC, Dayus B. Comparison of the 1969 and 1993 standardizations of the Bayley Mental Scales of Infant Development for infants with Down's syndrome. *J Intellect Disabil Res.* 2001;45:56–62
17. Blasco PA. Preterm birth: to correct or not to correct. *Dev Med Child Neurol.* 1989;31:816–821
18. Rickards AL, Kitchen WH, Doyle LW, Kelly EA. Correction of developmental and intelligence test scores for premature birth. *Aust Paediatr J.* 1989;25:127–129
19. Capute AJ, Shapiro BK, Wachtel RC, Gunther VA, Palmer FB. The Clinical Linguistic and Auditory Milestone Scale (CLAMS). Identification of cognitive defects in motor-delayed children. *Am J Dis Child.* 1986;140:694–698

A Population-Based Study to Determine the Performance of the Cognitive
Adaptive Test/Clinical Linguistic and Auditory Milestone Scale to Predict the
Mental Developmental Index at 18 Months on the Bayley Scales of Infant
Development-II in Very Preterm Infants

Michael J. Vincer, Heather Cake, Michael Graven, Linda Dodds, Shelly McHugh and
Theresa Fraboni

Updated Information & Services	including high resolution figures, can be found at: http://pediatrics.aappublications.org/content/116/6/e864
References	This article cites 16 articles, 1 of which you can access for free at: http://pediatrics.aappublications.org/content/116/6/e864#BIBL
Subspecialty Collections	This article, along with others on similar topics, appears in the following collection(s): **Developmental/Behavioral Pediatrics** http://www.aappublications.org/cgi/collection/development:behavior al_issues_sub **Growth/Development Milestones** http://www.aappublications.org/cgi/collection/growth:development_ milestones_sub **Administration/Practice Management** http://www.aappublications.org/cgi/collection/administration:practice _management_sub
Permissions & Licensing	Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at: http://www.aappublications.org/site/misc/Permissions.xhtml
Reprints	Information about ordering reprints can be found online: http://www.aappublications.org/site/misc/reprints.xhtml

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™

PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

A Population-Based Study to Determine the Performance of the Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale to Predict the Mental Developmental Index at 18 Months on the Bayley Scales of Infant Development-II in Very Preterm Infants

Michael J. Vincer, Heather Cake, Michael Graven, Linda Dodds, Shelly McHugh and Theresa Fraboni

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/116/6/e864

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™