Use of Developmental Milestones in Pediatric Residency Training and Practice: Time to Rethink the Meaning of the Mean

Laura Sices, MD, MS
Division of Behavioral Pediatrics and Psychology, Department of Pediatrics, Rainbow Babies and Children’s Hospital, Case Western Reserve University, Cleveland, Ohio

Abstract

Objective—Pediatricians frequently report the use of developmental milestones in monitoring young children’s development, despite evidence that use of screening tools improves detection of developmental delays.

Methods—Core texts in the field of pediatrics and developmental-behavioral pediatrics were reviewed for content and presentation on child development. Most texts included and many focused on developmental milestones, many with an emphasis on 50th percentile milestone data. Problems and limitations in the use of 50th percentile milestones to monitor young children’s development and to identify children whose development is suspicious for delay, include questionable utility in clinical decision making and the potential to increase parental anxiety.

Results—The recommendation is made to reconsider a focus on 50th percentile milestone data in pediatric training and practice, in favor of measures that have better clinical utility and are more psychometrically sound.

Conclusion—A conceptual approach to the presentation of developmental milestones differentiates the use of the 10th, 50th, and 90th percentiles of age of achievement of skills, based on the clinical purpose of surveillance.

Index terms
devolutional monitoring; developmental screening; developmental milestone
Studies suggest that while certain, less common types of developmental conditions, specifically those associated with neonatal risk factors, motor delays, and genetic conditions, are typically identified in a timely way by pediatricians, more prevalent conditions including speech and language delay, milder cognitive disabilities, and atypical behaviors may only be identified later in childhood or even at school entry, often outside the medical system. Use of validated developmental screening tools can increase detection of developmental delays. There has been limited adoption of the systematic use of developmental screening tools in practice, however, with multiple barriers identified to their use. Comparison of a child’s current developmental skills to milestone data remains the most frequently reported method of developmental surveillance for physicians in practice, in conjunction with the clinical assessment and the physical examination of the child. It is therefore critical to understand the validity, utility, and limitations of developmental milestones as a tool in developmental surveillance. In this commentary, I recommend moving away from a traditional focus on presentation of the 50th percentile for developmental skills to a varied presentation of information with criteria based on clinical and training purposes. This information is important for developmental-behavioral pediatricians who are charged with teaching pediatric residents about child development and developmental screening and who may author chapters on child development in pediatric textbooks.

In order to evaluate how developmental milestones are currently presented in academic texts, I undertook a review of chapters on child development or developmental disabilities in standard pediatric texts; a popular pediatric reference guide; texts in developmental-behavioral pediatrics; guidelines on health supervision from the AAP and Bright Futures; and a parent guide published by the AAP. The texts were chosen as a representative rather than exhaustive sample of published educational materials available for training and practice in the United States, based on my experiences as a trainee and faculty member in three pediatric training programs, and the recommendations of colleagues in the field. Chapters on child development or developmental disabilities were reviewed for content, with an emphasis on presentation of age-skill material (milestones) and original sources for milestone data. Original sources of milestone data cited in the references were reviewed.

**DEFINITION OF DEVELOPMENTAL MILESTONE**

This review found that milestone lists are a consistent feature of chapters on child development in reference texts reviewed. Many texts present information on the mean age of acquisition of specific developmental skills (50th percentile) as a point estimate. A few texts present information on the upper range of normal (e.g., the 75th or 90th percentile) or the typical age range of acquisition of skills in certain tables. In a number of cases, however, the percentile represented by the data is not indicated, either in the table or the accompanying text. Tables may include ambiguous labels such as “emerging patterns of behavior,” “developmental expectations,” “landmarks/guidelines,” “approximate ages of skill attainment,” and “milestones by the end of the period” that make it difficult for the reader to understand what the data represent. In certain cases, the source of the original milestone information provided in some tables is not noted, so that it is not possible to assess the quality of the data. Strategies and information relating to developmental monitoring and screening are presented in many of these texts, but the milestone lists feature prominently.

In contrast, certain texts do not emphasize presentation of milestones or point estimates for skills. One text limits the presentation of milestones to appendices, and focuses on discussion of monitoring and screening strategies.
also includes multiple examples of developmental screening tools. A third adopts a diverse strategy for presenting milestone information, with some tables listing point estimates of the mean age of acquisition and others the outer limit of normal and some data on typical ranges (5th to 95th percentiles) for certain skills in addition to inclusion of examples of multiple screening tools in the appendices.

Although knowledge of 50th percentile milestones is no longer emphasized in guidance for the certification examination in general pediatrics, this knowledge continues to be emphasized for the recertification examination. Guidance available on the topic of child development on the American Board of Pediatrics’ website for the 2005 recertification examination focused almost exclusively on lists of 50th percentile developmental milestones, and this knowledge was emphasized at a Boards review course that I attended last year.

LIMITATIONS OF STUDY POPULATIONS ON WHICH MILESTONE INFORMATION IS BASED

It is important for physicians to be aware of the source of the milestone data that they are using as part of developmental surveillance. Of concern, the original sources for milestone data cited in certain of the texts reviewed are from samples that have significant limitations because of their small size or lack of diversity. In some cases, the seminal studies of Arnold Gesell and colleagues are cited as a source. Gesell’s studies were conducted in the late 1920s and early 1930s on a population of 107 infants and children from families chosen to be most representative of the average, from a single city in Connecticut; all the children were white. Another source of milestone information in the references are the Child Development Inventories. Irton developed this well-regarded measure in the 1990s based on a population of 568 children drawn from a single community in Minnesota; the sample is described as “without extremes of wealth or poverty” and was 95% white.

Other texts cite sources with larger, more representative population samples. The work of Gesell was revised by Knobloch, who collected additional data from 927 evaluations of children in the Albany, NY area during 1975–1977 and included black as well as white children (15% and 85% of the sample, respectively). Continuing in the tradition of 50th percentile milestones, Accardo and colleagues recently standardized the Capute Scales (previously the Clinical Adaptive Test/Clinical Linguistic and Auditory Milestone Scale [CAT/CLAMS]) on a larger, more diverse sample of 1055 children from five metropolitan areas; the sample was 57% white, 32% black, and 11% “other” race. The population for the Denver-II, drawn from the state of Colorado, is also large (N = 2096) and relatively diverse.

LIMITATIONS OF THE MILESTONE CONCEPT

There are a number of theoretical concerns about the use of 50th percentile developmental milestones in developmental surveillance. Knowing the mean age at which a child is expected to demonstrate a skill provides limited information to assess an individual child and limited information about child development in general: one reason is the considerable variability in the normal acquisition of developmental skills in early childhood. The 50th percentile is a point estimate that provides little guidance about what is “normal,” except for children whose skills are at the mean. As an example, based on data from the Denver-II, the 50th percentile for the skill of saying “dada/mama” nonspecifically is 6.5 months, and the 90th percentile is 9.1 months. The time interval between these two points (the mean and the “red flag”) is therefore 2.6 months. In contrast, the 50th percentile for the acquisition of two-word combinations is 19.8 months, and the 90th percentile is 25.2 months. The time interval between the mean and the “red flag” for this latter skill is 5.4 months, more than twice that for the former skill. This
example illustrates that the 50th percentile alone is inadequate to estimate the outer limit of normal for the acquisition of a skill.

Residents are not taught to use mean values in monitoring or screening other aspects of a child’s health, yet this tradition developed and has persisted in the area of child development. Take as an analogy reference values for a common laboratory test such as the hematocrit. The lab reports the range of normal values as a reference to assist the clinician in making a determination about the significance of the patient’s result. Knowing the normal mean of the hematocrit is of limited value in clinical decision making, whereas knowing the upper and lower bounds of normal is useful in ruling out or diagnosing polycythemia or anemia. Similarly, knowledge of the mean age of acquisition of developmental skills provides little information to determine whether a child’s skills are typical for age or fall outside of the expected range.

**USE OF THE DEVELOPMENTAL QUOTIENT IN CONJUNCTION WITH DEVELOPMENTAL MILESTONES**

Another concern is the temptation to use individual 50th percentile milestones to obtain a crude developmental quotient (DQ). The concept of a DQ used in combination with developmental milestones was introduced by Gesell as “a shorthand device for expressing the rate of development.” The DQ was defined as “the relationship between maturity age [based on comparison to the 50th percentile: clarification mine] and actual age, expressed as a ratio.”

The concept of the DQ is at the foundation of the Clinical Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS), recently standardized as the Capute Scales, and is featured in the *Harriet Lane Handbook*. Using the CAT/CLAMS or Capute Scales, the provider elicits information about an extensive, standardized list of developmental milestones from parents combined with observation of skills in a standardized way. Information based on data from an aggregate of multiple milestones produces an age equivalent based on comparison to 50th percentile milestones that is divided by the child’s chronological age and multiplied by 100 to produce the DQ. A DQ of <70%–75% is indicative of delay in the affected area of development.

The concept of a DQ is appropriate when used with the Capute Scales, a standardized measure that aggregates data on multiple milestones, or other validated measures that provide an age-equivalent for a child’s functioning. The psychometric validity of a DQ used in combination with individual skills or milestones in isolation, however, is not known, but is probably extremely limited. This is because the age of acquisition of many skills is not normally distributed in the population, and the standard deviation (SD) for the age of acquisition of individual skills is quite variable. Individual skills are likely to vary significantly in their prediction of actual developmental problems. In addition, normative data on milestones after the first year of life are sparse, so it is difficult to estimate a child’s “age equivalent” based on individual skills. The concern, therefore, is that physicians may be extending the concept of DQ, which is valid when used with a standardized tool such as the Capute Scales or other standardized tool that provides an estimate of the child’s developmental age, to use with individual milestones, a practice for which there is no scientific evidence.

**PARENTAL ANXIETY ABOUT CHILDREN WHOSE DEVELOPMENT FALLS BELOW THE 50TH PERCENTILE**

A focus on 50th percentile milestones in clinical practice has another drawback: the potential to raise parental anxiety. By definition, the developmental skills of half of children fall below
the 50th percentile. If residents are taught to ask parents about 50th percentile milestone skills and this is the information commonly available in reference books, approximately half of parents, depending on the similarity of their backgrounds to the standardization population, might perceive that their child is behind in skill acquisition. In other words, after reviewing 50th percentile milestone information, as many as half of parents could conclude that their child is possibly “delayed.” This could unduly raise parents’ anxiety and concern and create the need for additional parent education and reassurance. It is conceivable that the widespread availability and use of 50th percentile milestone information, rather than presentation of information about the range of typical development, may increase parental anxiety unnecessarily, reducing the predictive value of parental concerns.

Conversely, if residents learn from experience that about half of children are not meeting “milestones,” they may observe that many “normal” children do not meet milestones and reassure parents about a child’s development unless it is severely delayed. The risk is that the parents of young children with mild to moderate delays may receive false reassurance.

In summary, my review found that milestone lists continue to be a key feature of the presentation of educational material on child development in many published sources in the field of pediatrics in the United States. In some cases, the original sources for milestone data are not indicated: this is concerning since it is not possible to ascertain the quality of the information. In many cases, tables of milestone data do not clearly indicate the percentile represented. The use of 50th percentile developmental milestones appears to have a long tradition in pediatric training and practice, dating to the work of Gesell in the 1920s. Despite the prevalence of milestone lists, the use of developmental milestones, 50th percentile milestones in particular, is not endorsed by policy statements on developmental monitoring and screening,1,2 nor backed by scientific evidence.

Accurate monitoring and screening of young children’s development in pediatric practice is a complex process requiring reliable, valid, and multimodal approaches. Ideally, residents are taught to elicit parental concern about specific areas of a child’s development in an evidence-based manner during serial preventive care visits,46,47 combined with clinical observations, and periodic use of a standardized developmental screening tool, with prompt referral for further assessment and treatment when significant concerns arise.1,2 A range of effective screening tools exists to aid in developmental monitoring and screening, including a tool to elicit parental concerns using a brief validated questionnaire,10 standardized parent developmental screening questionnaires,9,36 and brief provider-directed tools.12,13,39 Information on available screening tools is reviewed in recently revised guidelines from the American Academy of Pediatrics.2

The routine use of validated developmental screening tools during health supervision visits has the potential to improve the early detection of developmental delays. Given the limited adoption of structured developmental screening tools in practice, however, it is imperative to consider alternative strategies to maximize the effectiveness of developmental surveillance. Because age-skill (milestone) information continues to be a focus of pediatric training and practice, we must consider ways to maximize the utility of this type of information.

**ALTERNATIVE APPROACH TO MONITORING CHILDREN’S DEVELOPMENT**

An approach for the use of age-skill information, based on the multiple goals of developmental monitoring, is presented in Table 1. Tasks related to child development during preventive care visits with infants and young children can be conceptualized as having four major objectives. The first is to provide anticipatory guidance to help parents prepare for the acquisition of new developmental skills, particularly emerging skills that place infants and young children at risk of physical injury. In this case, knowledge of the 10th percentile, the age at which some children
may start demonstrating the skill, is useful. For example, a clinician would want to discuss rolling over in infants before the age at which infants may start to demonstrate the skill to prevent falls in infancy.

The second goal is to provide information to parents about typical behaviors at different developmental stages, to address and normalize developmentally appropriate behaviors, or to help parents anticipate emerging skills and behaviors. Information on what is “typical,” the 50th percentile, is helpful in this case. The third objective is to promote the development of young children in general. For this goal, knowledge of an individual child’s developmental skills or a general knowledge of activities that promote development at different developmental stages is needed.

Last, a major goal of developmental monitoring is to identify in a timely manner children whose development is suspicious for delay or is atypical and to respond proactively to parents’ developmental concerns. In this case, information on the 90th percentile (the “red flag”), and/or information from a structured screening tool, is required. If a child is not yet demonstrating an important skill observed in most children that age, the child requires further screening and/or assessment. This screening can be conducted by the clinician using a validated measure or through the state’s Early Intervention program, which in many states is set up to provide such screening as part of the “child find” mission of the program.

This conceptual structure clarifies the utility of different types of age-skill information for the multiple tasks related to child development during preventive care visits.

**A RECOMMENDATION FOR CORE TEXTS TO MOVE AWAY FROM 50TH PERCENTILE MILESTONES**

Given that many texts in pediatrics focus on presenting 50th percentile milestone data, it is not surprising that physicians continue to use milestones in practice. To effect a change, it will be more useful to present information on the “typical range” or “expected range” of skill acquisition (or, alternatively, mean age and SD) rather than 50th percentile milestones, as was seen in many of the texts reviewed. Such information provides guidance on the normal range of acquisition, on which to base a clinical assessment of a child’s skills. Presentation of this information will be most effective if it is coupled with strategies for the monitoring, screening, and assessment of child development, based on policy recommendations and scientific evidence. Data on the 90th percentile of attainment are available for the larger population samples studied for the Denver-II and the Capute Scales (in the latter, the lower limit is termed the “10th percentile”).

It will be important for the recertification examination in pediatrics to move away from emphasizing knowledge of 50th percentile milestones in the guidance provided to examinees, as has occurred for the certification examination.

**TOWARD A MORE PRODUCTIVE USE OF AGE-SKILL DATA IN DEVELOPMENTAL MONITORING**

In response to evidence that children with speech and language and cognitive delays may not be identified in a timely way by pediatricians, it may make sense to set clinical goals for the age by which key developmental conditions are identified. For example, goals can be set for the theoretical ages at which children with suspected autism, cognitive disability, or speech and language disorders are typically identifiable to increase awareness of the early manifestation of these conditions, and highlight the importance of monitoring and screening.
of early speech and language development and early referral for assessment and treatment for these relatively common conditions.

The goal of these recommendations is to contribute to a dialogue on training and practice in the complex area of child development, and to suggest that it is time to break with the long-standing tradition of 50th percentile milestones in pediatrics in favor of more comprehensive presentation of age-skill information, including the range of normal acquisition of skills (10th–90th percentiles). Developmental-behavioral pediatricians who develop curricula for residency training and write textbook chapters in this area are in a position to effect a change by considering these recommendations.

ACKNOWLEDGMENTS

Thanks to Stephanie Serna, M.S., for assistance in reviewing texts and references and in formatting the table, and to James Bales, B.S., for assistance in reviewing texts. Thanks to Dennis Drotar, Ph.D., Ed Schor, M.D., and H. Gerry Taylor, Ph.D., for their thoughtful comments.

Author disclosure: L.S. was supported by grant K23 HD047713 from the National Institute of Child Health and Human Development.

REFERENCES


### Table 1
Objectives Related to Developmental Monitoring During Preventive Care Visits with Young Children

<table>
<thead>
<tr>
<th>Objective</th>
<th>Method(s)</th>
<th>Information Needed</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevention: promote safety and knowledge/understanding of child development</td>
<td>Anticipatory guidance</td>
<td>10th percentile for skill, for children who demonstrate skills earlier than the mean, or in anticipation of future skill acquisition (focus is on motor skills in the first 2 yrs of life)</td>
<td>Rolling over at 2 mo, discussed at the 1- or 2-mo well-child visit (to anticipate the risk of rolling off a couch or table in infancy)</td>
</tr>
<tr>
<td>Explanation and normalization; counsel about typical behaviors, skills and activities observed in children at different ages</td>
<td>“Milestones”; discuss typical behaviors for age, and normalize behavior/skills for age; identify child’s and family’s strengths</td>
<td>50th percentile for skill; typical/“average” skills and behaviors for age (primarily behaviors and social and play skills)</td>
<td>Temper tantrums in 2 year olds; Parallel play in 2 year olds; Use of distraction and redirection to manage behavioral issues with the young child</td>
</tr>
<tr>
<td>Developmental promotion: provide guidance to parents on ways to promote parent-child interaction and promote a child’s development</td>
<td>Provide brief parent education during the preventive care visit; provide written materials</td>
<td>Knowledge of an individual child’s current skills or 50th percentile skills/typical skills for age</td>
<td>Encouraging parents to look at books/read to their young children; providing guidance on limiting television/media viewing for children</td>
</tr>
<tr>
<td>Identification of delays: monitor/screen development over time, to detect problems or delays and refer to appropriate assessment and treatment resources in a timely way</td>
<td>“Red flag;” absence of an important skill at a certain age; elicit and respond to parental concerns in multiple developmental and behavioral areas (e.g., PEDS10); screening/monitoring with the use of standardized or validated tools (e.g., ASQ, CDI, BINS)</td>
<td>90th percentile for skill (alternative: 1.5 SDs below the mean); parental questionnaire such as the PEDS10; asking similar questions during the visit; pass or fail score on a screening instrument (all areas of development, with a focus on language, cognitive and social development)</td>
<td>Absence of novel two-word combinations by 24 mo (concerning for language delay); absence of pointing coordinated with eye contact to share interest or excitement by 18 mo (joint attention) (concerning for delay in social development/PDD)</td>
</tr>
</tbody>
</table>

*The identification of strengths, resources, and resilience factors that will allow a child and family to address any developmental issues is an important part of monitoring and screening at each stage. PEDS, Parents’ Evaluation of Developmental Status; ASQ, Ages and Stages Questionnaire; CDI, Child Development Inventories; BINS, Bayley Infant Neurodevelopmental Screener; PDD, pervasive developmental disorder.*