Feeding difficulty is common in young infants in the first 6 months of life. As advancements in healthcare have allowed for increasing survival of critically ill newborns, the number of infants with oral feeding difficulty has been rising. Infants who are born premature and with congenital heart disease are at particularly high risk for problematic feeding; however, some infants who are otherwise apparently healthy also have difficulty with bottle- or breastfeeding.

Difficulty with feeding during infancy is problematic because it often results in suboptimal nutrition during a critical time for brain growth. In infants with chronic disease, suboptimal nutrition contributes to morbidity, such as risk for infection and prolonged hospitalization, as well as mortality. Feeding difficulties during infancy also impact the developing social dynamics of mealtime and the parent-child relationship.

Early identification of infants with feeding difficulty is critical to implement appropriate therapies and optimize the infant's nutrition and oral feeding skill development. However, because infants have limited ability to communicate distress during feeding, problematic feeding behaviors may be subtle and present in heterogeneous ways. All of these factors make differentiation between typical feeding behaviors and problematic feeding difficult.

Clinicians need assessment tools to guide objective evaluation of infant oral feeding that is consistent regardless of who is completing the assessment (eg, nurse or speech-language pathologist). Psychometric development and testing is critical to ensure that (1) the tool is valid (ie, it measures what it is intended to measure) and (2) it is reliable (ie, the score represents something true about the infant).
infant’s feeding and it performs in a consistent way). If a tool is not valid, it may, for example, measure parent-infant interaction during feeding rather than the infant’s feeding behaviors. If a tool is not reliable, the score may indicate a good feeding when the feeding was problematic, or it may give a different score depending on who completes the assessment.

PSYCHOMETRIC DEVELOPMENT AND TESTING

An extensive review of all types of validity and reliability testing is beyond the scope of this article. Common methods for evaluating assessment tools are presented here.

Validity

Content validity examines the extent to which a tool comprehensively measures a construct (eg, feeding). Content validity is determined by the manner in which the tool is developed and is tested by obtaining feedback from intended users of the tool. Ideally, a tool would be developed using a combination of sources (eg, input from experts, clinical experience, and a review of the literature) and then formally tested by experts using a content validity index. Other types of validity testing may include concurrent or predictive validity, which examines how well a tool predicts a score on another measure, either a tool that is considered the “gold standard” or something that is hypothesized to be a related construct (eg, nutritional intake).

Reliability

Reliability testing requires a sample of respondents to complete the tool. Internal consistency reliability evaluates how related items on the tool are to each other, indicating the items measure the same construct. Interrater reliability evaluates how consistent 2 different respondents (eg, nurses) are to each other when assessing the same event. Intrarater reliability evaluates how consistent the same respondent (eg, the same nurse) is when assessing the same event at 2 times (eg, a videotape of the same feeding 25 weeks apart). A reliability of 0.7 or more is considered acceptable and less than 0.7 is considered unacceptable. Ideally, all of these types of reliability would be tested.

PURPOSE

The purpose of this review is to identify and evaluate assessment tools available to guide clinical assessment of bottle- and breastfeeding in infants younger than 6 months. The intended purpose of the tool, user (eg, nurse, parent, researcher), feeding method (eg, breast- or bottle-feeding), and the psychometric testing conducted to support the tools’ reliability and validity will be presented.

SEARCH STRATEGY

CINAHL, HaPI (Health and Psychosocial Instruments Database), PubMed, and Web of Science were searched to identify feeding assessment tools for infants younger than 6 months. The terms used for the search were “infant feeding” and “assessment tool.” The search was limited to English language, human, and full text. Both articles and textbooks were included and no limits were placed on publication date. The literature search was conducted in June 2015.

The literature was reviewed by the research team for presentation of new assessment tools, use of existing tools, or reference to existing tools. Assessment tools were excluded if they were intended for infants older than 6 months, were intended for assessment of solid food feeding (eg, pureed baby food), or were intended to assess a construct other than the infant’s feeding (eg, parent-infant interaction, breastfeeding self-efficacy, feeding readiness). Once tools were identified, a secondary specific search of tools by name via PubMed and CINAHL was conducted to identify additional literature on that specific tool. Since the intent of this article is to provide evidence for clinical practice, assessment tools were further excluded if they lacked sufficient published literature to evaluate the tool, if the target population was limited to a specific diagnosis, or if the tool was intended for research only (ie, not intended for clinical use).

The initial search of databases resulted in 237 articles and texts for review (Figure 1). From this literature, 18 relevant tools were identified that met inclusion criteria. Three of these were excluded because they did not have adequate published literature for evaluation of the tool: B-R-E-A-S-T-Feed Observation Form, Infant Nipple Feeding Assessment and Communication Tool, and Oral Eating Readiness Assessment List. Three additional tools were excluded because their use is limited to specific diagnoses: the Feeding Checklist (infants with non-organic failure to thrive), the Infant Malnutrition and Feeding Checklist for Congenital Heart Disease (infants with congenital heart disease), and the Nutrition and Feeding Risk Identification Tool (infants in early intervention). Finally, the Mother–Infant Feeding Tool was excluded because it assesses mother–infant interaction in the context of feeding but is not a feeding assessment; it is also an observational coding system intended for research purposes only. The final sample included 11 assessment tools (Table 1).

SUMMARY OF EVIDENCE

Bottle-feeding Assessment

Two tools are available to assess bottle- or breastfeeding in infants younger than 6 months: the Early Feeding Skills (EFS) Assessment and the Neonatal...
Oral Motor Assessment Scale (NOMAS). However, both of these tools have been tested only in a sample of bottle-feeding infants. The EFS Assessment is intended for preterm infants through at least 52 weeks postmenstrual age and takes a holistic approach to the assessment of feeding (S. M. Thoyre, C. S. Shaker, K. F. Pridham, unpublished data, 2012). It is a 36-item assessment of behavioral state, readiness, muscle tone, energy level, behavioral stress signs, swallowing, physiologic stability, and oral-motor functioning (S. M. Thoyre, C. S. Shaker, K. F. Pridham, unpublished data, 2012). Subscales of the assessment (eg, Ability to Maintain Physiologic Stability) are scored and indicate areas of strength, areas of some clinical concern, and areas of major clinical concern. The EFS Assessment was developed with neonatal nurses and feeding experts, which supports content validity (S. M. Thoyre, C. S. Shaker, K. F. Pridham, unpublished data, 2012) but has not had formal content validity testing. It has acceptable intra- and interrater reliability (value not reported).

The NOMAS is a 28-item assessment intended for preterm and term infants and focuses primarily on oral-motor skills (ie, movements of the jaw and the tongue) for sucking, with only 2 items about fatigue or incoordination of the suck, swallow, breathe sequence. Subscales of the assessment (eg, Nutritive Suck: Jaw) are scored and characterize the oral-motor pattern as normal, disorganized, or dysfunctional. The NOMAS was developed by a speech-language pathologist and was one of the first assessment tools available for evaluation of objective physiologic cues during feeding, but the development process is not described in the literature and there is no evidence of content validation. There is evidence of concurrent validity between the NOMAS and neurobehavioral scores and cerebral diameter, but the NOMAS was not found to be predictive of feeding outcomes in preterm infants. Results for interrater (0.50-1.00, 0.43-0.62, and 0.33-0.95) and intrarater reliability (0.33-1.00 and 0.41-0.65) have been inconsistent and/or unacceptable. Further development of the tool and testing has been recommended.

Both the EFS Assessment and the NOMAS are intended for clinicians and require specialized training. The EFS Assessment requires a 2-day training course (www.shaker4swallowingandfeeding.com/course-brochure/) or is available by request to Dr Suzanne Thoyre (thoyre@email.unc.edu). The NOMAS requires a 3-day training course (http://www.nomasinternational.org).

**Breastfeeding Assessment**

Nine tools identified for assessment of infant feeding indicate that they could be used only for the assessment of breastfeeding. Two of these tools were found to be assessment guides that did not have a scoring system and, therefore, no psychometric testing: the Breastfeeding Evaluation and Education Tool and Systematic Assessment of the Infant at Breast. The remaining 7 tools are the Bristol Breastfeeding Assessment Tool (BBAT), Infant Breastfeeding Assessment Tool (IBFAT), the LATCH, Mother–Baby Assessment (MBA), Mother–Infant Breastfeeding Progress Tool (MIBPT), Potential Early Breastfeeding Problem Tool (PEBPT), and the Premature Infant Breastfeeding Behavior Scale (PIBBS).

The BBAT is a 4-item assessment intended for healthy, full-term infants and has been tested with 218 breastfeeding sessions of infants up to 10 weeks of age. The items assess positioning, attachment, sucking, and swallowing and are scored on a scale of poor, moderate, or good. The BBAT was developed in a manner that supports content validity, but no formal content validity testing was done. The internal consistency reliability of the BBAT was 0.67, which is slightly below acceptable, but interrater reliability was acceptable (0.78). The BBAT was tested alongside the LATCH and IBFAT and found to be more responsive to changes over time than the latter 2.

The IBFAT is a 6-item assessment intended for healthy, full-term infants in the early postpartum
<table>
<thead>
<tr>
<th>Tool (Author)</th>
<th>Purpose</th>
<th>Target Population and Age</th>
<th>Feeding Method</th>
<th>Intended User</th>
<th>Psychometric Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breastfeeding Evaluation and Education Tool (BEET) [27]</td>
<td>To guide assessment of breastfeeding, identify problems, and facilitate seeking assistance.</td>
<td>Healthy, full-term young infants (age not specified)</td>
<td>Breast</td>
<td>Mother or clinician</td>
<td>None. This is a list of questions/interview and assessment guide. No scoring system.</td>
</tr>
<tr>
<td>Bristol Breastfeeding Assessment Tool (BBAT) [29]</td>
<td>To facilitate accurate, rapid breast feeding assessment and targeted advice to mothers acquiring skills or experiencing problems.</td>
<td>Tested with healthy, full-term infants up to 10-wk old.</td>
<td>Breast</td>
<td>Clinician</td>
<td>Evidence of content validity but no testing done. Internal reliability slightly below acceptable, but interrater reliability acceptable.</td>
</tr>
<tr>
<td>Early Feeding Skills Assessment (EFS) [22]</td>
<td>To standardize assessment of feeding skills and individualize interventions.</td>
<td>Preterm infants through 52 wk postmenstrual age.</td>
<td>Breast or bottle</td>
<td>Clinician (*Specialized training required)</td>
<td>Evidence of content validity but no testing. Acceptable inter- and intrarater reliability.</td>
</tr>
<tr>
<td>LATCH [31]</td>
<td>To standardize assessment and teaching of breastfeeding.</td>
<td>Not specified</td>
<td>Breast</td>
<td>Mother or clinician</td>
<td>Evidence of content validity but no testing. Inconsistent interrater reliability.</td>
</tr>
<tr>
<td>Mother–Baby Assessment (MBA) [32]</td>
<td>To assist clinicians in documenting their observations of breastfeeding.</td>
<td>Not specified</td>
<td>Breast</td>
<td>Clinician</td>
<td>No content validation. Inconsistent interrater reliability.</td>
</tr>
<tr>
<td>Neonatal Oral Motor Assessment Scale (NOMAS) [33]</td>
<td>To qualify and describe patterns of disorganized and dysfunctional sucking.</td>
<td>Preterm, medically fragile, and chronically ill infants.</td>
<td>Breast or bottle</td>
<td>Clinician (*Specialized training required)</td>
<td>No evidence of content validation. Results of reliability testing inconsistent and/or unacceptable.</td>
</tr>
<tr>
<td>Potential Early Breastfeeding Problem Tool (PEBPT) [34]</td>
<td>To elicit early breastfeeding events that may be indicative of breastfeeding problems.</td>
<td>Healthy, full-term infants during postpartum hospitalization.</td>
<td>Breast</td>
<td>Clinician</td>
<td>Evidence of content validity but no testing. Acceptable internal consistency reliability.</td>
</tr>
<tr>
<td>Premature Infant Breastfeeding Behavior Scale (PIBBS) [35]</td>
<td>To facilitate clinical observation of preterm infant breastfeeding behavior.</td>
<td>Hospitalized preterm infants.</td>
<td>Breast</td>
<td>Mother or clinician</td>
<td>Evidence of content validity but no testing. Interrater reliability of revised tool unacceptable.</td>
</tr>
<tr>
<td>Systematic Assessment of the Infant at Breast (SAIB) [28]</td>
<td>To evaluate the effectiveness of the infant's contribution to breastfeeding.</td>
<td>Healthy, full-term infants (age not specified).</td>
<td>Breast</td>
<td>Clinician</td>
<td>None. This is an assessment/teaching guide. No scoring system.</td>
</tr>
</tbody>
</table>
period and can be completed by either the mother or a clinician.\textsuperscript{30} The IBFAT assesses infant state, readiness, rooting, latching, sucking behaviors, and maternal satisfaction with the breastfeeding experience.\textsuperscript{30} Items 2 to 5 comprise the infant feeding assessment and are scored on a 4 response option scale with a sum score of 10 to 12 indicating an effective, vigorous feeder; 7 to 9 indicating a moderately effective feeder; and 0 to 6 indicating that the infant was not able to be roused or did not root and sucked weakly during feeding.\textsuperscript{30} The IBFAT was developed using a literature review and observing infant feedings, which supports content validity, but no formal testing was done.\textsuperscript{30} Interrater reliability between clinicians has been found to be acceptable (0.7-0.78).\textsuperscript{31} Interrater reliability between clinicians and mothers was acceptable (0.91) in a moderate sample of feeding observations (n = 77)\textsuperscript{30} but unacceptable (0.27-0.69) in a study on a small sample (n = 23).\textsuperscript{37} The latter study may not have been of a sufficient sample size for this type of testing. In the pilot study of 15 mother–infant dyads by Schlomer et al,\textsuperscript{38} the IBFAT was found to increase with higher maternal satisfaction and lower breastfeeding problem scores, but the findings were not statistically significant. The IBFAT has also been tested with very low-birth-weight infants (<1.5 kg), and better IBFAT scores were found to be significantly correlated with milk intake volume, intake rate, and percent of time sucking.\textsuperscript{39} No process was indicated to ensure validity of this instrument for this different population and no reliability testing was done.

The LATCH (target population not specified) is a 5-item assessment of latch, audible swallowing, type of nipple, comfort of breast/nipple, and hold (positioning) scored on a scale from 0 to 2, with higher scores indicating more ideal breastfeeding (eg, successful latch and frequent audible swallowing).\textsuperscript{1} The LATCH was developed with experts, but no formal content validity testing is reported.\textsuperscript{31} The LATCH can be completed by clinicians or mothers. Interrater reliability between clinicians has been inconsistent. In a study of 35 feedings, reliability was acceptable (0.85-1.0)\textsuperscript{40} and in another study of 46 feedings, reliability was mostly acceptable (0.65-0.91).\textsuperscript{36} In a third small sample of only 23 feedings, reliability was unacceptable (0.11-0.48).\textsuperscript{37} The only study of interrater reliability between clinicians and mothers (n = 33) found unacceptable reliability (0.18-0.67).\textsuperscript{40}

The MBA (target population not specified) is a 5-item assessment completed by a clinician that evaluates signaling, positioning, fixing, milk transfer, and ending, with each item scored as either 1 (behavior present) or 0 (behavior absent).\textsuperscript{32} The mother and baby each receives a score on each of the 5 items, with a higher score indicating a more effective breastfeeding encounter.\textsuperscript{33} The process of development of the MBA is not discussed and there is no evidence of content validity. Interrater reliability results of the MBA have also been variable, with 1 small study of 23 feedings finding unacceptable reliability (0.33-0.66),\textsuperscript{37} but another study of 46 feedings finding acceptable reliability (0.81-0.88).\textsuperscript{36}

The MIBPT is an 8-item checklist intended for healthy, late preterm and term infants and completed by clinicians.\textsuperscript{34} The 8 items evaluate maternal response to feeding cues, length between feedings, latch, nutritive sucking bursts, independence in positioning and latching infant, nipple trauma, and negative comments made about breastfeeding.\textsuperscript{33} Items are checked as either yes or no and the purpose of the MIBPT is to assess maternal and infant behaviors in order to facilitate teaching and support by the clinician.\textsuperscript{31} The MIBPT was tested with 81 breastfeeding sessions of infants 35 to 42 weeks postmenstrual age during the postpartum hospitalization (2-5 days old).\textsuperscript{33} The tool development supports content validity and content validity was assessed informally by 2 experts. Interrater reliability for each item was acceptable (0.79-0.95).\textsuperscript{33} No other testing is reported.

The PEBPT is a 23-item list of possible breastfeeding events (eg, baby falls asleep, sore nipples, breast infection), rated as either 0 (not experienced), 1 (manageable, no real problem), 2 (usually manageable, sometimes a problem), or 3 (very difficult to manage, makes me consider weaning).\textsuperscript{34} The PEBPT was developed in a manner that supports content validity but has no formal content validity testing. Lower breastfeeding satisfaction and bottle use were predictive of worse PEBPT score (R² = 0.15; P < .01).\textsuperscript{34} Schlomer et al\textsuperscript{38} found the internal consistency reliability for the PEBPT to be acceptable (0.81), but no further development or testing has been done.

The PIBBS is a 12-item checklist intended for hospitalized, preterm infants and can be completed by a clinician or a mother.\textsuperscript{35} Items evaluate rooting, latching, sucking, swallowing, general behavior, letdown, problems with the breast and the nipple, and influence of the environment. Each item is scored differently.\textsuperscript{35} The PIBBS was developed in a manner that supports content validity but has not had formal content validity testing. Interrater reliability of the current tool has been tested only with 10 feedings and has found unacceptable reliability between clinicians (0.44) and between mothers and clinicians (0.5-0.75).\textsuperscript{35}

**RECOMMENDATIONS FOR PRACTICE**

None of the assessment tools currently available have formal content validity testing and none have comprehensive reliability testing. Further development and testing is needed before any of the tools...
can be recommended as valid and reliable measures of infant feeding. Given these limitations, recommendations are made for the assessment tools with the most supportive psychometric development and testing. Figure 2 provides a flowchart to guide the selection of an assessment tool for evaluating bottle- or breastfeeding in the infant younger than 6 months.

For the bottle-feeding preterm or full-term infant up to 52 weeks of postmenstrual age, the EFS Assessment is a more comprehensive assessment of feeding and it has more supportive psychometric development and testing than the NOMAS. Specifically, the development of the EFS Assessment supports content validity of this tool while the development process is not described for the NOMAS. In addition, the EFS Assessment has acceptable interrater and intrarater reliability, while the NOMAS has inconsistent and/or unacceptable interrater and intrarater reliability. For the assessment of the breastfeeding preterm infant younger than 35 weeks of postmenstrual age, the EFS Assessment has more supportive psychometric development and testing than either the NOMAS or the PIBBS. While the PIBBS also has evidence of content validity, the very limited testing of the current tool with 10 feedings has found unacceptable interrater reliability.

**FIGURE 2**

Flowchart to guide selection of a feeding assessment tool. Asterisk indicates recommended. BBAT indicates Bristol Breastfeeding Assessment Tool; BEET, Breastfeeding Evaluation and Education Tool; EFS, Early Feeding Skills Assessment; IBFAT, Infant Breastfeeding Assessment Tool; LATCH; MBA, Mother-Baby Assessment; MIBPT, Mother-Infant Breastfeeding Progress Tool; NOMAS, Neonatal Oral Motor Assessment Scale; PEBPT, Potential Early Breastfeeding Problem Tool; PIBBS, Premature Infant Breastfeeding Behavior Scale; SAIB, Systematic Assessment of the Infant at Breast.

www.advancesinneonatalcare.org

Copyright © 2016 National Association of Neonatal Nurses. Unauthorized reproduction of this article is prohibited.
Summary of Recommendations for Practice and Research

What we know:
- Feeding difficulty in infants is common and challenging to differentiate from typical feeding.
- Feeding assessment tools allow for objective measurement of bottle- and breastfeeding.
- None of the currently available tools have adequate psychometric testing to ensure validity and reliability.

What needs to be studied:
- Feeding assessment tools need to be developed and tested for validity and reliability.
- An ideal tool would measure both bottle- and breastfeeding to allow consistent assessment of both feeding methods.
- A clinician-report tool needs to be developed to guide assessment, tailoring of interventions, and documentation of feeding.
- A parent-report tool needs to be developed for use after hospital discharge to identify infants in need of specialized assessment and treatment.

What we can do today:
- The Early Feeding Skills (EFS) Assessment has the most supportive psychometric development and testing for assessment of bottle- and breastfeeding in the preterm infant through 52 weeks postmenstrual age and the full-term infant experiencing complex feeding difficulties.
- The Bristol Breastfeeding Assessment Tool (BBAT) has the most psychometric support for assessment of breastfeeding in healthy, full-term infants experiencing minimal feeding difficulties.

For the breastfeeding full-term infant, Figure 2 indicates the assessment tools available that are intended for use in the early postpartum period (<5 days old) and those intended for use with infants beyond the first 5 days of life. The EFS Assessment and BBAT are the 2 assessment tools with the most supportive psychometric development and testing of the tools available for the breastfeeding full-term infant, regardless of age. The BBAT has evidence of content validity and acceptable interrater reliability, and internal consistency reliability was nearing acceptable (0.67). The EFS Assessment not only has reported testing of internal consistency reliability but also has evidence of content validity and has acceptable interrater and intrarater reliability. The significant difference between these 2 tools is the comprehensiveness of the assessment. The BBAT has only 4 items while the EFS Assessment has 36 items. The BBAT does not require specialized training and is a faster, more concise assessment, which may be appropriate for assessing the healthy, full-term infant experiencing minimal difficulties. For the infant experiencing more complex feeding difficulties in need of a more comprehensive assessment, the EFS Assessment may be more appropriate. There are currently no tools that have been tested with medically fragile infants other than those born preterm, and there are no tools available for assessment of the bottle-feeding infant older than 52 weeks postmenstrual age.

RECOMMENDATIONS FOR RESEARCH

Valid and reliable assessment tools for the evaluation of bottle- and breastfeeding are needed for infants from birth through 6 months of age. These tools need to be developed and tested for use with infants who are at highest risk for feeding difficulty, particularly those who are medically fragile, as well as with infants who are otherwise healthy. Assessment tools that require specialized training provide an opportunity for continuing education about feeding; however, the requirement for training limits the use of the tool. The development of a tool that is valid and reliable without specialized training would allow for more widespread use. While breastfeeding at breast may be ideal, many medically fragile infants require some bottle-feedings and clinicians need tools to objectively assess the infant’s ability to safely and effectively feed. Many of the tools reviewed in this article are promising, but they have not yet been developed and tested comprehensively. Even the EFS Assessment, which was identified as the most comprehensive assessment with the most supportive development and testing, needs formal content validity testing, testing of internal consistency reliability, and reported values for interrater and intrarater reliability.

Future research should focus on the psychometric development and testing of tools that can be used for both bottle- and breastfeeding. A tool that can be used for assessment of both feeding methods will allow clinicians to learn 1 assessment and consistently evaluate feeding regardless of feeding method. For infants who remain hospitalized or are continuing to develop feeding skills, an assessment tool that is intended for clinicians would guide clinical assessment, documentation of feeding, and tailoring of feeding interventions.

For infants who have been discharged from the hospital but are experiencing feeding difficulty, a
parent-report tool would be best. Parents are in the best position to report on typical feeding behavior, which may be different from behaviors seen by clinicians observing a feeding in an environment that is unusual for the infant. A tool that could be used from the time of discharge up through the introduction of solid foods would allow clinicians to follow the progress of feeding development and/or response to feeding treatment over time using the same tool. In some circumstances, the combination of a clinician-report and parent-report tool may provide a comprehensive assessment and help guide discussions between clinicians and parents about feeding concerns. Development and testing of assessment tools is a lengthy but necessary process to provide clinicians with the tools they need to support best clinical practice and provide optimal care to patients.

References

8. Hehir DA, Cooper DS, Walters EM, Ghaynayem NS. Feeding, growth, nutrition, and optimal interstage surveillance for infants with hypoplastic left heart syndrome. Cardiol Young. 2011;21(suppl 2):59-64.

Copyright © 2016 National Association of Neonatal Nurses. Unauthorized reproduction of this article is prohibited.