

New Validated Tool to Diagnose Breastfeeding Dysfunction

Journal of Human Lactation
2024, Vol. 40(1) 120–131
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/08903344231209306
journals.sagepub.com/home/jhl



Rajeev Agarwal, MD¹ , Mars Eddis-Finbow, MSc¹, Jodie Tam, BS²,
Jennifer Broatch, PhD³, and Kimberly J. Bussey, PhD^{2,4} 

Abstract

Background: Breastfeeding behaviors and experiences exist on a continuum. What differentiates normal from dysfunctional is defined by frequency and severity. No current validated tool addresses the subjective experience of dyads with a predictive score that can be followed over time.

Research Aim: To create and validate a self-report tool to assess breastfeeding and evaluate its ability to predict risk of breastfeeding dysfunction.

Methods: This study used a cross-sectional design to determine the validity of a novel instrument to assess breastfeeding dysfunction. We gave the initial questionnaire to 2085 breastfeeding dyads. We assessed content validity by comparison with other tools. We used exploratory factor analysis with varimax rotation for concept identification and Cronbach's alpha for internal consistency. We employed logistic regression to assess the tool's ability to differentiate between normal breastfeeding and breastfeeding dysfunction.

Results: Factor analysis mapped 17 questions to four concepts to create a score (FLIP; flow, latch, injury [to the nipple], and post-feed behavior). Internal consistency and reliability of the scores in these concepts were acceptable (Cronbach's alpha ≥ 0.087 for all measures). A logistic regression model that controlled for infant age, with a breastfeeding dysfunction risk classification threshold of 60%, yielded a correct classification of 88.7%, with 93.1% sensitivity, 64.6% specificity, and a 6.5% false positive rate.

Conclusions: The FLIP score was determined to be a valid and reliable instrument for quantifying the severity of breastfeeding dysfunction in children under 1 year old. Further studies will assess its usefulness in the management of breastfeeding dysfunction.

Keywords

Academy of Breastfeeding Medicine, American Academy of Pediatrics, ankyloglossia, breastfeeding, breastfeeding assessment, breast pain, breastfeeding support, International Board Certified Lactation Consultant, lactation, tongue-tie

Background

Breastfeeding, defined as the child receiving breastmilk either directly from the breast or expressed (World Health Organization [WHO], 2008), is associated with benefits to both infant and breastfeeding parent (Couto et al., 2020; Horta, 2019; North et al., 2022). Both parental and infant experiences determine its continuation and longevity. These experiences can include both factors external to the breastfeeding dyad and physiological factors in both parent and infant.

External factors contribute to early cessation of breastfeeding. Socioeconomic status, racial disparities in the initiation of breastfeeding, social stigma around breastfeeding in public, lack of access to support systems for continuing breastfeeding upon return to work, and limited access to culturally competent lactation resources, have all been found to affect both whether and how long parents breastfeed (Apanga et al., 2022; Beauregard

et al., 2019; Gianni et al., 2019; Hamner et al., 2021; Li et al., 2019; Whipps et al., 2019). The perception of external recommendations and social support from partners and parental grandparents also influence breastfeeding patterns (Whipps et al., 2019; Zhang et al., 2023).

Among physiological factors, a constellation of symptoms indicative of breastfeeding problems, including nipple pain

¹Agave Pediatrics, Glendale, AZ, USA

²University of Arizona, Tucson, AZ, USA

³Arizona State University, Glendale, AZ, USA

⁴Midwestern University, Glendale, AZ, USA

Date submitted: November 28, 2022; Date accepted: October 6, 2023.

Corresponding Author:

Rajeev Agarwal, MD., Agave Pediatrics, 3575 West Deer Valley Road, Glendale, AZ 85308, USA.

Email: ragarwal@agavepediatrics.com

and poor latch, have been implicated as causes for earlier than desired cessation of breastfeeding (Gasparin et al., 2019; Gianni et al., 2019; Morrison et al., 2019). These symptoms may be associated with functional impairment in sucking or milk transfer that is characterized as “breastfeeding dysfunction.” Breastfeeding dysfunction can have several causes and includes both parental and infant factors that stem from anatomical or physiological issues; partial ankyloglossia is an example. Partial ankyloglossia, commonly referred to as ankyloglossia or tongue-tie, refers to a lingual frenulum that is short or thick. It is associated with restricted lingual mobility and impaired breastfeeding (Cordray et al., 2023; Fraga et al., 2020; LeFort et al., 2021). Thus, partial ankyloglossia severe enough to warrant treatment can serve as marker for known breastfeeding dysfunction.

Because breastfeeding behaviors span a spectrum from normal to dysfunctional, various tools including the Mother–Baby Assessment (MBA; Mulford, 1992), the Bristol Breastfeeding Assessment Tool (BBAT; Ingram et al., 2015), The Infant Breastfeeding Assessment Tool (IBFAT; Matthews, 1988), and LATCH (Latch, Audible swallowing, nipple Type, Comfort, Holding; Jensen et al., 1994) are used to evaluate full-term infants for evidence of breastfeeding difficulties. These tools are filled out by the healthcare provider, who must elicit the information from the parent, and depend on the provider’s examination, which requires extra training. While these tools have been applied in the context of identifying dyads at risk for breastfeeding difficulties, they were not designed nor validated for the identification of breastfeeding dysfunction. LATCH was originally designed as a charting aid (Jensen et al., 1994). MBA was designed to assess breastfeeding sessions in much the same way APGAR (Appearance, Pulse, Grimace, Activity, Respiration) scores are used to assess newborn health (Mulford, 1992). BBAT was designed to assist midwives, lactation consultants, and other healthcare professionals in assessing breastfeeding proficiency (Ingram et al., 2015). IBFAT was constructed to allow parents, midwives, and nurses to assess infant performance during feeding (Matthews, 1988).

In addition to tools that assess behavioral indicators of breastfeeding difficulties, there are tools that focus on evaluating infant anatomy and function including The Frenectomy Decision Rule for Breastfeeding Infants (FDRBI; Srinivasan et al., 2006) and Lingual Frenulum Protocol with Scores for Infants (Lopes De Castro Martinelli et al., 2012). The FDRBI is an assessment performed by a clinician to guide intervention. The clinician evaluates whether there is nipple pain and/or trauma and difficulty latching along with the range of motion of the tongue. Clinicians using the Lingual Frenulum Protocol with Scores for Infants assess breastfeeding difficulty by asking about maternal pain, length of feeding and infant behaviors, for example, fatigue or short duration in a historical context. It then requires an evaluation of a feed as well as infant oral anatomy.

No single tool assesses the complete symptom spectrum of breastfeeding dyads. The FDRBI and the Lingual

Key Messages

- No validated self-report tool exists for the identification of breastfeeding dysfunction that addresses the symptomology of both partners in the breastfeeding dyad.
- FLIP exhibits high content validity and predictive performance to identify the risk of breastfeeding dysfunction in dyads with infants under 1 year of age.
- FLIP can be used to monitor breastfeeding dysfunction over time and can serve as a tool to assess the effectiveness of a broad range of interventions to improve the breastfeeding experience.

Frenulum Protocol with Scores for Infants focus primarily on infant behaviors with little consideration given to maternal factors that influence breastfeeding success. LATCH and IBFAT are centered on maternal pain and infant latch, but fail to elicit information about other behaviors associated with breastfeeding dysfunction. BBAT focuses primarily on the infant assessing positioning, attachment, sucking, and swallowing. MBA documents steps in the breastfeeding process and both parent and infant are scored based on whether they successfully complete a step. MBA fails to document specific behaviors or other indicators of breastfeeding dysfunction (e.g., nipple pain). In addition, only BBAT and LATCH have been assessed for both reliability and validity. IBFAT and MBA have only been evaluated for reliability (Bickell et al., 2018), and the FDRBI and Lingual Frenulum Protocol with Scores for Infants have not been validated.

Symptom severity is a subjective observation that depends largely on parental perception, and often can be misleading, particularly if assessed longitudinally. Symptom frequency, on the other hand, has better stability over time (Krabbe & Forkmann, 2012). There is a need for a tool that does not require extensive training on the part of the provider to implement and can quantify symptom severity in terms of symptom frequency to differentiate between normal breastfeeding behaviors and breastfeeding dysfunction. The purpose of this study was to create and validate a self-report tool to assess breastfeeding and evaluate its ability to predict risk of breastfeeding dysfunction.

Methods

Research Design

This study used a cross-sectional design to evaluate the validity of a novel instrument to assess breastfeeding dysfunction. Because the instrument is new, responses were evaluated for content validity, internal consistency, and

known groups validity, in order to facilitate its use both clinically and in further research studies assessing intervention effectiveness. The study was approved by the Western Institutional Review Board (IRB number 20181213).

Setting and Relevant Context

Paper surveys were administered in English as part of routine medical office paperwork in a large suburban pediatric practice in the Phoenix, Arizona metropolitan area. Socioeconomic factors (e.g., family income, education, race, and ethnicity), a lack of access to lactation support, and social stigma, all exacerbate the difficulties associated with breastfeeding dysfunction and contribute to early cessation of breastfeeding (Beauregard et al., 2019; Li et al., 2019). The communities surrounding the practice locations had median annual incomes ranging from \$60,499 to \$97,409 ($N=2,228,368$), poverty rates ranging from 6.7% to 17.3% ($N=2,228,368$), and college education rates ranging from 22.1% to 60.2% ($N=2,228,368$; U.S. Census Bureau, 2023). Minority populations (Black, Asian, Native American, Hawaiian or Pacific Islander, and Hispanic) in these communities ranged from 21.3% to 58.6% ($N=2,228,368$; U.S. Census Bureau, 2023). The 2018 to 2019 breastfeeding initiation rate for Maricopa County, where the Phoenix metropolitan area is located, was 89.0% (Centers for Disease Control and Prevention [CDC] 2021). The Arizona Department of Health hosts a website with information about breastfeeding resources in the state for parents and employers. The state also provides a 24-hour breastfeeding hotline (Arizona Department of Health Services [AZDHS], 2022). Access to an International Board Certified Lactation Consultant (IBCLC) was part of the routine care offered in the practice. The CDC periodically assessed public opinions about breastfeeding through the SummerStyles Survey. In 2021, the majority of respondents (69.02%) agreed with the statement “I believe women should have the right to breast feed in public spaces” (CDC, 2023) and Arizona state law protects the right of mothers to breastfeed in public (Arizona Revised Statutes—Title 41, 2006)

Sample

The target population was healthy breastfeeding parents over the age of 18 and healthy full-term infants under a year old with or without symptoms of breastfeeding dysfunction seen during routine clinical care. Participants were either primary patients of the practice or had been referred to the practice from outside for lactation consultation. Participants from a variety of socioeconomic backgrounds were included (e.g., those who qualified for the Special Supplemental Nutrition Program for Women, Infants, and Children [WIC]). However, specific information about socioeconomic background, race, ethnicity, and parental age were not collected. Medical record review was used to verify age and breastfeeding status, determine whether the infant had a frenectomy, and when, relative

to the survey date, and to identify confounding conditions. We excluded infants with prematurity or physical (e.g., clefting), neurological, cardiac, or genetic problems that could contribute to breastfeeding dysfunction. Since symptoms of breastfeeding dysfunction overlap with normal breastfeeding behaviors and can be caused by suboptimal latch, most dyads reporting breastfeeding difficulties had already been evaluated by an IBCLC before their appointment to rule out suboptimal latch, and all dyads were offered IBCLC support after the evaluation. Participants were not compensated.

The cross-sectional sampling scheme is illustrated in Figure 1. Following survey administration, completed surveys were identified. Note that a completed survey only means that some portion of the survey questions were answered and does not imply that all survey questions were answered. Medical record review was carried out and participants were divided into two groups. Participants who had evidence of partial ankyloglossia severe enough to have progressed to frenectomy were classified as the known breastfeeding dysfunction group ($n=1737$). Those who were seen in the office for well-child checks and no note was made in the medical record about breastfeeding difficulties were classified as the normal breastfeeding group ($n=348$). Dyads who were referred for breastfeeding difficulties and had frenectomy prior to taking the survey were excluded. We also excluded those dyads who were recommended for frenectomy but chose not to proceed because in these situations it is unclear where on the spectrum of breastfeeding behaviors these dyads would fall. For example, the decision to not proceed to frenectomy may indicate a level of symptomatology within the realm of normal. The final sample was 2,085 dyads.

The 2,085 dyads proved to be a more than sufficient sample size for the subsequent analyses. The Kaiser's Measure of Sampling Adequacy: Overall MSA=0.893 for the exploratory factor analysis (EFA) indicates an adequate sample size (Shrestha, 2021). The post-hoc power analysis for the logistic regression model for classification revealed a power of > 0.99 for FLIP sub-scores Flow, Latch, Injury, and 0.91 for the predictor Post-feed Behavior. The post-hoc power analysis for the logistic regression on the training data set only revealed similar power results with a power of > 0.99 for FLIP sub-scores Flow, Latch, Injury, and 0.80 for the predictor Post-feed Behavior. The statistically acceptable lower value of power for post-feed behavior was a consequence of that sub-score consisting of fewer questions than other sub-scores.

Measurement

The demographic variables evaluated were assigned sex and infant age at survey. Ages were computed in days based on the date of the survey and the birthdate of the infant obtained from the medical record. Ages were grouped as follows: 0–14 days were assigned to the 0–2 weeks category; 15–30 days were assigned to the 2–4 weeks category; 31–61 days were assigned to the 1–2 month category; 62–91 days were assigned to the 2–3 month category; 92–183 days were assigned to the

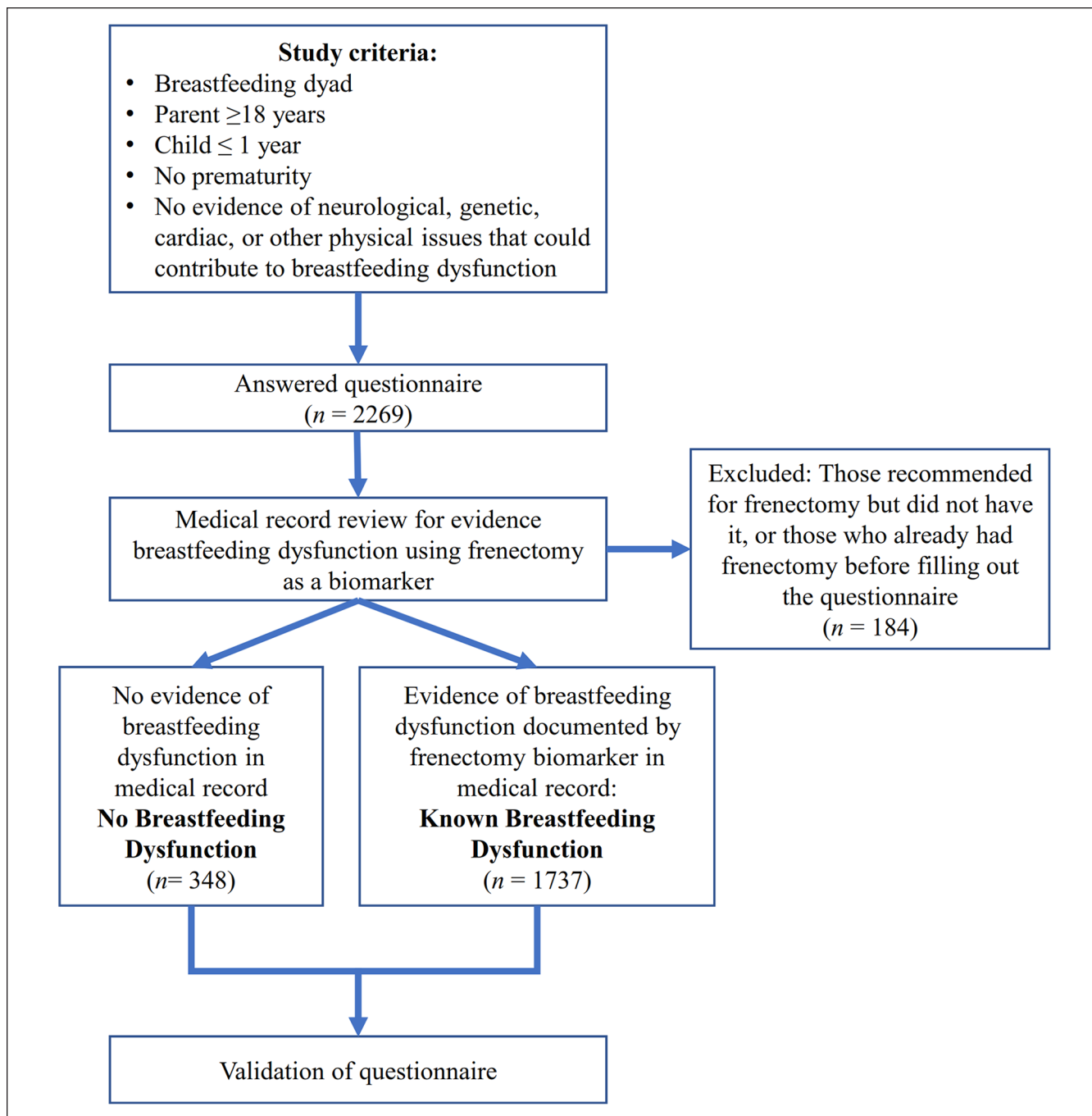


Figure 1. Study Design.

3–6-month category; 184–273 days were assigned to the 6–9-month category, and 274–365 days were assigned to the 9+ months category. We chose these categories as they align with the recommended timing of well-child checks.

The most common parental and infant symptoms were identified through a literature survey and formed the concepts guiding the Results (IQ) construction. The IQ can be found in Supplemental Table 1 (see the online supplemental material). The frequency of symptoms (e.g., how often they

occur) rather than the severity of symptoms was chosen for assessment to mitigate bias, as reported severity was more variable and subjective than frequency, depending on the immediate psychological state of the parent (Krabbe & Forkmann, 2012). Except for the standard pain question which rated pain between 0 and 10 (Hawker et al., 2011), and two questions that asked for specific amounts of time, questions asked the respondent to indicate how often a symptom or behavior occurred: never, less than 25% of the time, 50%

Table 1. Demographics of Infants Included in the Study by Group (N = 2085).

	No BF dysfunction (n = 348)	BF dysfunction (n = 1737)
	n (%)	n (%)
Assigned sex at birth		
Female	154 (44.25)	779 (44.82)
Male	194 (55.75)	958 (55.12)
Age group		
0-2 weeks	154 (44.25)	416 (23.96)
2-4 weeks	9 (2.59)	454 (26.15)
1-2 months	94 (27.01)	443 (25.52)
2-3 months	38 (10.92)	195 (11.23)
3-6 months	36 (10.34)	170 (9.79)
6-9 months	13 (3.74)	40 (2.3)
9+ months	4 (1.15)	18 (1.04)

Note. BF = breastfeeding.

of the time, greater than 75% of the time, or 100% of the time. This was re-coded to a 5-point Likert-like scale (Jebb et al., 2021), which precludes the existence of outliers. Numeric responses reported as ranges were recoded as the arithmetic average of the end points of the supplied range (e.g., if patient stated 1–3, a score of 2 was recorded) and used in the subsequent analyses.

For the purposes of future practical use and rapid clinical implementation, data for four questions were reformatted to maintain a consistent scale across all questions. The standard pain question (originally on a 0–10 scale) was divided by 2 to remain consistent with the other questions on a 0–5 scale. The question “Do you feel the baby empties your breasts?” was reverse scaled as higher scores were associated with improving rather than worsening symptoms. Responses from the questions “Average time baby is on breast/bottle per feed?” and “How many times did you breastfeed, last 24 hours?” were normalized so that responses in the middle 20% would warrant a score of 1. More extreme observations received higher scores, resulting in the responses within the top 10% and bottom 10% receiving a 5.

Parental Symptoms. Questions on parental symptoms focused on nipple pain and blanching. As nipple pain is a top reason for early breastfeeding cessation (Morrison et al., 2019), we included a pain scale as well as a frequency scale to best quantify it. All other questions related to parental symptoms assessed frequency of occurrence.

Infant Symptoms and Feeding Behavior. Symptoms resulting from infant anatomy included shallow latch, clicking during feeding, curling the upper lip under while feeding, and leaking from the corners of the mouth. We also asked about infant gassiness and spit-up. All questions asked about the frequency of symptoms.

Infant feeding behavior referred to symptoms present during or directly after feeding. These included latching on and off, choking or gagging, chewing, or clamping down on the nipple, fussiness immediately after feeding, and the presence of milk coming out of the nose during feeding. The frequency of symptoms was assessed.

Milk Production. Milk production was addressed through several questions. We asked about the frequency with which the parent felt that the infant emptied the breasts. We also asked how many ounces milk or formula were given via bottle. Finally, we elicited information about the frequency and duration of feeds.

Data Collection

Paper surveys were administered from January 2016 to September 2019 by healthcare providers at the point of care. The breastfeeding parent was asked to fill out the survey as part of the in-office paperwork for the visit. Informed consent was obtained verbally by the individual administering the questionnaire. Surveys were not used by providers to inform patient care. Data were stored separately from patient medical records and securely on-site following all HIPAA regulations. Surveys were de-identified prior to data analysis to maintain confidentiality.

Data Analysis

Descriptive statistics (counts and percentage of total) were computed for assigned sex and age variables for the infants in both groups, as well as for items and four key concept scales (mean and standard deviation). Missing data were omitted for summative statistics and only complete cases (n = 1529) were used for subsequent analyses. All data analysis was conducted using SAS software (Version 9.4). The Kruskal-Wallis test for the association of sub-score with misclassification was calculated in R (Version 4.1.1), using the stats package.

Tool Validation. Content validity was performed to determine how the content of the tool compared with other tools. Validation then proceeded by an exploratory factor analysis to identify consistent underlying concepts addressed by the questions and associate questions with concepts. Questions that did not map to concepts were dropped from the IQ as they were not measuring the content they were intended to measure. Internal consistency was evaluated using Cronbach’s alpha (American Educational Research Association et al., 2014). Known-group analysis was used to evaluate the predictive power of the tool.

Content Validity Analysis. The IQ was compared to other similar tools: LATCH: A breastfeeding charting system and documentation tool (Jensen et al., 1994), Frenectomy Decision Rule

Table 2. Correspondence of Questions to Existing Tools.

Name of survey	Scored by:	Correspondence between tools
LATCH: A Breastfeeding charting system and documentation tool	Healthcare Provider	The scale of "L" of LATCH corresponds to: Does the infant have a shallow latch? Does the infant click during breastfeeding? Does the infant latch on and off the nipple during breastfeeding? The scale for "C" of LATCH corresponds to: What is your pain level when breastfeeding from 1-10? Does the child chew/gum/clench/clamp down/bite the nipple during breastfeeding?
Frenectomy Decision Rule for Breastfeeding Infants (FDRBI)	Healthcare Provider	The first part of the decision tool corresponds to the following: What is your pain level when breastfeeding from 1-10? Does the infant have a shallow latch? How often do you have nipple pain during breastfeeding?
Bristol Breastfeeding Assessment Tool (BBAT)	Healthcare Provider	The scale for swallowing and attachment corresponds to the question: Does the infant click during breastfeeding?
Lingual Frenulum Protocol with Scores for Infants	Healthcare Provider	Historical assessment and Evaluation of Orofacial Functioning section 2.2-2.4 correspond to the following: Does your child feed less than every 2 hours? Does your child appear tired, fall asleep or get frustrated during feeding? Does the infant click during breastfeeding? Does the infant latch on and off the nipple during breastfeeding? Does the infant cough during breastfeeding? Does the child chew/gum/clench/clamp down/bit the nipple during breastfeeding?

for Breastfeeding Infants (FDRBI; Srinivasan et al., 2006), Bristol Breastfeeding Assessment Tool (BBAT; Ingram et al., 2015), and Lingual Frenulum Protocol with Scores for Infants (Lopes De Castro Martinelli et al., 2012) to identify areas of overlap with other validated tools.

Factor Analysis. The IQ contained 28 items (questions) to be evaluated (see Supplemental Table 1 in the online supplemental material). We performed an exploratory factor analysis (EFA) with a varimax rotation on all patient responses, regardless of group. We verified the appropriateness of the sample size (Lawley & Maxwell, 1971) and method (Bartlett’s test of sphericity: 10346.5268, $p < .001$). Questions were deemed to be related to a construct when their factor loading exceeded 0.35. Items unrelated to any construct were removed from the analysis.

Internal Consistency Analysis. We used Cronbach’s alpha (Cronbach, 1951) to demonstrate the level of consistency in the responses across the entire question set. The sub-scores identified by the factor analysis were also analyzed for internal consistency. A value ≥ 0.6 was considered sufficient.

Known-Groups Validity Analysis. The IQ’s intended use was to identify the presence and severity of breastfeeding dysfunction. The diagnosis of partial ankyloglossia severe enough to warrant frenectomy identifies a known group with breastfeeding dysfunction. The data were split into a test group and validation group. The survey responses for 70% ($n=1234$) of patients with breastfeeding dysfunction

and 70% ($n=224$) of the patients with no breastfeeding dysfunction were used to develop a logistic regression model to estimate the likelihood of breastfeeding dysfunction based solely on the constructs, controlling for the age of the infant by modeling it as a categorical variable. The 70% were selected from the overall sample and not selected proportional to the group sizes. The accuracy of the model was then tested on the remaining 30% of the patients in each group (known breastfeeding dysfunction $n=503$ and no breastfeeding dysfunction $n=124$) with the threshold of breastfeeding risk set at 60%. Only cases without missing data were used in the analysis. We assessed overall classification rate, area under the curve, sensitivity, specificity, false positive rate, and false negative rate.

Results

Characteristics of the Sample

The infant participants ranged in age from 0 to 264 days. Out of the 2085 infant participants, 993 (47.63%) were assigned female and 1152 (55.25%) were assigned male. The details of age and assigned sex by group classification of the infant participants are given in Table 1.

Content Validity

We identified the questions in our questionnaire that overlapped with concepts covered in similar tools. As seen in Table 2, the IQ covered similar questions in the areas of

Table 3. Mean and Standard Deviation of Responses on the FLIP Tool by Group.

Subsection	Question	No BF dysfunction (n = 348)		BF dysfunction (n = 1737)		Combined (N = 2085)		Response Scaling
		n	Mean (SD)	n	Mean (SD)	N	Mean (SD)	
Flow	Does infant choke/gag during feeding? [1-5]	340	1.54 (0.78)	1692	2.45 (1.27)	2032	2.29 (1.24)	1 = infant rarely/never chokes, 5 = infant always/often chokes
	Does infant hiccup immediately after feeding? [1-5]	339	1.99 (0.92)	1710	2.86 (1.15)	2049	2.72 (1.16)	1 = infant rarely/never hiccups, 5 = infant always/often hiccups
	Does infant cough during feeding? [1-5]	338	1.39 (0.67)	1700	2.08 (1.13)	2038	1.96 (1.10)	1 = infant rarely/never coughs, 5 = infant always/often coughs
	Does infant spit up after feeding? [1-5]	339	1.97 (0.94)	1705	2.54 (1.27)	2044	2.44 (1.24)	1 = child never/rarely spits up, 5 = child often/always spits up
	Does child leak from corners of mouth after feeding? [1-5]	340	1.73 (0.84)	1711	2.77 (1.36)	2051	2.59 (1.35)	1 = child rarely/never leaks, 5 = child always/often leaks
Latch	Does infant get tired/frustrated during feeding? [1-5]	336	2.17 (1.07)	1696	3.45 (1.21)	2032	3.24 (1.28)	1 = infant rarely/never gets tired, 5 = infant often/always gets tired
	Does infant have shallow latch? [1-5]	323	1.84 (1.10)	1622	3.80 (1.23)	1945	3.47 (1.41)	1 = latch not shallow, 5 = latch very shallow
	Does infant click during feeding? [1-5]	327	1.77 (1.13)	1669	2.96 (1.44)	1996	2.77 (1.46)	1 = child never/rarely clicks, 5 = child clicks often/always clicks
	Does infant latch on/off during feeding? [1-5]	334	2.45 (1.20)	1704	3.65 (1.25)	2038	3.45 (1.32)	1 = infant rarely/never latches on/off, 5 = infant always/often latches on/off
	Does child tuck upper lip under? [1-5]	317	1.82 (1.09)	1603	3.35 (1.42)	1920	3.10 (1.48)	1 = child never/rarely tucks, 5 = child clicks often/always tucks
Injury	What is your pain level when nursing? [0-5]	326	1.11 (1.19)	1668	2.07 (1.39)	1994	1.91 (1.40)	0 = no pain, 5 = maximum pain possible
	Do you have nipple pain when feeding? [1-5]	340	2.08 (1.17)	1692	3.30 (1.40)	2032	3.10 (1.44)	1 = never have pain, 5 = always have pain
	Does infant chew/clamp down/bite nipple during feeding? [1-5]	341	1.75 (1.00)	1680	3.00 (1.43)	2021	2.79 (1.44)	1 = infant rarely/never latches, 5 = infant always/often latches
	Do you notice blanching (fading/whitening/discoloration) of the nipples with breastfeeding? [1-5]	339	1.47 (0.93)	1659	2.70 (1.54)	1998	2.49 (1.52)	1 = never blanches, 5 = always blanches
Post Feed Behavior	Do you feel baby is fussy right after feeding? [1-5]	342	1.52 (0.79)	1713	2.61 (1.20)	2055	2.43 (1.21)	1 = baby rarely/never fussy, 5 = baby always/often fussy
	Do you feel baby has excessive gas [1-5]	339	2.02 (1.00)	1703	3.03 (1.39)	2042	2.86 (1.38)	1 = never excessive gas, 5 = often/always excessive gas
	Does child arch back during/after feeding [1-5]	335	1.45 (0.79)	1684	2.23 (1.26)	2019	2.10 (1.23)	1 = child rarely/never arches back, 5 = child always/often arches

Note. SD = standard deviation.

parental pain and infant feeding behavior to currently validated tools. The correspondence in content ranged from one question (BBAT) to six questions (Lingual Frenulum Protocol with Scores for Infants).

Construct Identification

The IQ consisted of 28 initial questions. Factor analysis identified four key constructs that were related to one another and significantly contributed to the presence of breastfeeding dysfunction. The constructs were defined by question-factor loading of at least 0.35 onto one unique factor. The unique factors were able to be clinically classified in the following constructs: Flow, Latch, Injury, and Post-Feed Behavior. Questions that did not load onto any of the four factors were excluded from the final validated tool. Based on this analysis and general data inconsistencies, 11 questions were removed from the FLIP Tool (see Supplemental Table 2). These included questions related to milk production. The final tool (see Supplemental Table 3) comprised 17 questions grouped

into four sub-scores corresponding to the identified constructs. The FLIP tool was scored by summing the scaled score of all questions in a sub-group to generate the sub-score. Sub-scores were then added to generate a total score. The sub-scores had the following potential ranges: Flow: 5–25, Latch: 5–25, Injury: 4–20, and Post-Feeding Behavior: 3–15. The total FLIP score could range from 17 to 85. The questions that contributed to each construct along with the mean and standard deviations of the responses of the total sample and broken down by normal and known breastfeeding dysfunction groups are presented in Table 3 and Supplemental Table 4.

Internal Consistency and Reliability

We evaluated the internal consistency and reliability of the remaining 17 questions as a whole and within constructs by computing the correlation of each question with the total score as well as Cronbach's alpha. The overall Cronbach's alpha for raw variables was 0.884. The standardization of variables did

Table 4. Internal Consistency and Reliability of Individual Questions.

	Deleted variable	Raw variables
		Cronbach's alpha
Flow	Does infant choke/gag during feeding?	0.88
	Does infant hiccup immediately after feeding?	0.88
	Does infant cough during feeding?	0.88
	Does infant spit up after feeding?	0.88
	Does child leak from corners of mouth after feeding?	0.88
Latch	Does infant get tired/frustrated during feeding?	0.88
	Does infant have shallow latch?	0.87
	Does infant click during feeding?	0.88
	Does infant latch on/off during feeding?	0.87
	Does child tuck upper lip under?	0.88
Injury	What is your pain level when nursing?	0.88
	Do you have nipple pain during breastfeeding?	0.88
	Does infant chew/clamp down/bite nipple during feeding?	0.88
	Do you notice blanching (fading/whitening/discoloration) of the nipples with breastfeeding?	0.88
	Do you feel baby is fussy right after feeding?	0.88
Post-Feed Behavior	Do you feel baby has excessive gas?	0.88
	Does child arch back during/after feeding?	0.88

not change this. All four constructs had high internal consistency reliability: Flow (Cronbach's alpha=0.880), Latch (Cronbach's alpha=0.876), Injury (Cronbach's alpha=0.880), and Post-Feed Behavior (Cronbach's alpha=0.880). Detailed information about the questions in each construct can be found in Table 4 and the details for the sub-scores can be found in Supplemental Table 5.

Known-Groups Comparison

We assessed whether the questionnaire could distinguish between participants with known breastfeeding dysfunction and those who had no evidence of breastfeeding dysfunction. Descriptive statistics for the four sub-scores for the total sample and broken by groups are given in Supplemental Table 4. We built a logistic regression model to predict the probability of breastfeeding dysfunction, accounting for infant age, using 70% (n=1458) of the data and then tested the classification on the remaining 30% (n=627). We assessed the model's performance to classify dyads as having breastfeeding dysfunction with a probability of 60% or higher. The model correctly classified 88.7% of the participants. Analysis of area under the curve (AUC) for the prediction is shown in Figure 2. Sensitivity was 93.1% and specificity was 64.6%. The false positive rate was 6.5%, and false negative rate was 36.8%.

Misclassification was more pronounced among the no breastfeeding participants compared to participants with known breastfeeding dysfunction (no breastfeeding dysfunction: n=70, N=121; known breastfeeding dysfunction: n=18, N=505). While all sub-scores were significantly

associated with misclassification (Kruskal-Wallis Chi-Square, Supplemental Table 6), participant misclassification was driven by the Latch sub-score (Figure 3). For the "No breastfeeding dysfunction" dyads, 68.6% (n=48) of the misclassified participants had Latch sub-scores >10, whereas 76.4% (n=39) of the correctly classified participants had scores ≤ 9. For the "Known breastfeeding dysfunction" participants, 100% (n=18) of misclassified participants had Latch sub-scores ≤ 10, while 74.9% (n=365) of correctly classified participants had scores ≥ 15.

The model predicted the risk of breastfeeding dysfunction. However, quick clinical implementation required translating the risk cut-off of 60% into a score on the tool. We evaluated the relationship between predicted risk, FLIP score, and the Latch sub-score since the Latch sub-score was most strongly associated with misclassification (Figure 4). Dyads with a FLIP score > 47 had a predicted risk of greater than 90% of having breastfeeding dysfunction with Latch sub-score above 15. A score < 28 resulted in most participants having a risk of breastfeeding dysfunction under 60%. Scores between 28 and 47 constituted a group of participants where the risk of breastfeeding dysfunction was predicted to be in the 50%–80% range, with Latch sub-scores suggestive of breastfeeding dysfunction.

Discussion

Our study described the development, validation, and discriminative performance of a 17-item self-report questionnaire to evaluate the risk of breastfeeding dysfunction in breastfeeding dyads with infants under 1 year old. The risk

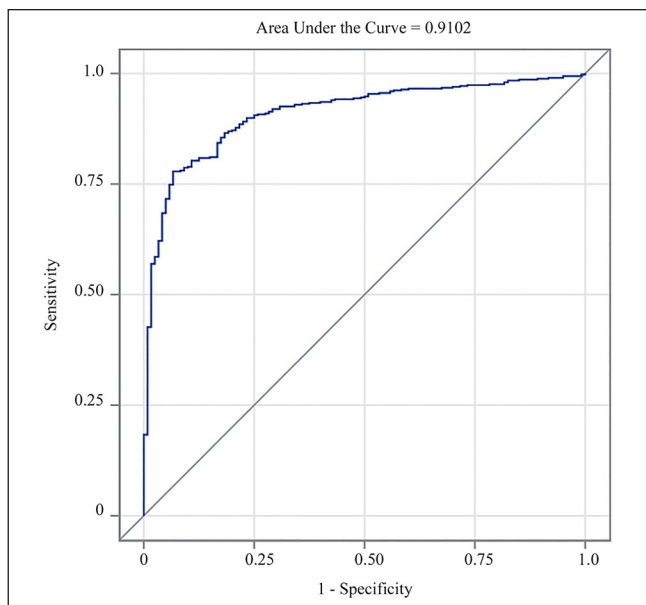


Figure 2. Classification Performance of the FLIP (Flow, Latch, Injury and Post-Feed behavior) Tool With a Breastfeeding Dysfunction Predicted Probability of 60% or Greater.

Note. The curve indicates the performance of the tool, while the diagonal line demonstrates how the tool would perform if the classification was made by chance.

model accounted for infant age and allowed us to translate risk into a FLIP score without needing separate scales for different age groups. The FLIP tool could risk-stratify and triage breastfeeding dyads into those with low risk of breastfeeding dysfunction, those who should be followed closely to determine if symptoms resolve, and those with sufficiently severe breastfeeding dysfunction to warrant an evaluation of underlying causes. The bias of the misclassification to the “No breastfeeding dysfunction” group and the false positive rate of the FLIP score reflected the likely inclusion in the no breastfeeding dysfunction group of individuals who had breastfeeding dysfunction but did not have the biomarker we used to identify known breastfeeding dysfunction. These individuals generally had Latch sub-scores greater than 10. The false negative rate was driven by individuals who had Latch sub-scores of less than 10. This suggested that in addition to paying attention to the total score, particular attention should be given to the Latch sub-score to further refine risk.

FLIP scores translated a dyad’s breastfeeding experiences into objective data that can facilitate longitudinal communication between families, IBCLCs, and other healthcare providers. In contrast to many other tools used in the context of identifying breastfeeding difficulties, the FLIP score can be filled out by parents prior to being seen by the provider, improving efficiency in-office, especially in high-volume practices where providers have short windows of time for providing care. This also provides opportunities for follow-up that do not require an in-person visit to the clinic. Other tools

including the MBA (Mulford, 1992), BBAT (Ingram et al., 2015), IBFAT (Matthews, 1988), and LATCH (Jensen et al., 1994) require the provider to observe one or more feeding sessions which limits their usefulness in high-volume practices and requires in-person follow-ups to assess the usefulness of any suggested interventions.

The FLIP tool differentiates between normal symptoms associated with breastfeeding that reflect the learning curve of both parent and infant and those resulting from treatable causes of breastfeeding dysfunction that could affect the longevity of breastfeeding. The total FLIP score and the sub-scores can determine the severity of breastfeeding dysfunction, and the inclusion of both parental and infant symptomology make it a comprehensive evaluation of the breastfeeding dyad, which, outside of the score, provides information to the clinician to pinpoint and respond to difficulties. For example, a parent could indicate that they always have pain when feeding, but no other symptoms. This information could guide the care given to that parent to help resolve the underlying issue. The quantitative nature and psychometric profile of the FLIP score and sub-scores makes the tool suited to the evaluation of interventions to improve breastfeeding dysfunction, including the timing of intervention. While one candidate for such studies would be the evaluation of the effectiveness of surgical intervention for anatomical issues like partial ankyloglossia, FLIP could be used to evaluate pedagogy for patient education in breastfeeding or the effectiveness of methods to correct positioning to improve latch.

Certain symptoms, while clinically important, were removed from the present study due to issues with variability in the data or too few responses. Milk production is important to consider in a dyad with breastfeeding difficulties, as a feedback loop exists between low milk production and difficulty in milk extraction. However, milk production is highly individualized, difficult to quantify without 24-hr weight testing, and perception versus reality of milk supply does not always match (Westerfield et al., 2018). Thus, questions on milk production were removed from the present study.

Another important set of questions that were removed regarded the length of feedings and time between feeds. In our clinical experience, when taken into consideration with other feeding symptoms, excessively short or long feeds may indicate breastfeeding dysfunction. Excessively long feeds or short time between feeds may be due to poor milk transfer, whereas excessively short feeds and long time between feeds may be due to infant fatigue and breast-aversion due to feeding difficulties. Parental milk production and age of the infant also influence the average length of feed, with older infants taking less time to feed than younger infants (Kent et al., 2013). Supplementation of human-milk substitutes and use of nipple shield are also confounding factors. This variation led to these questions being removed from the final questionnaire.

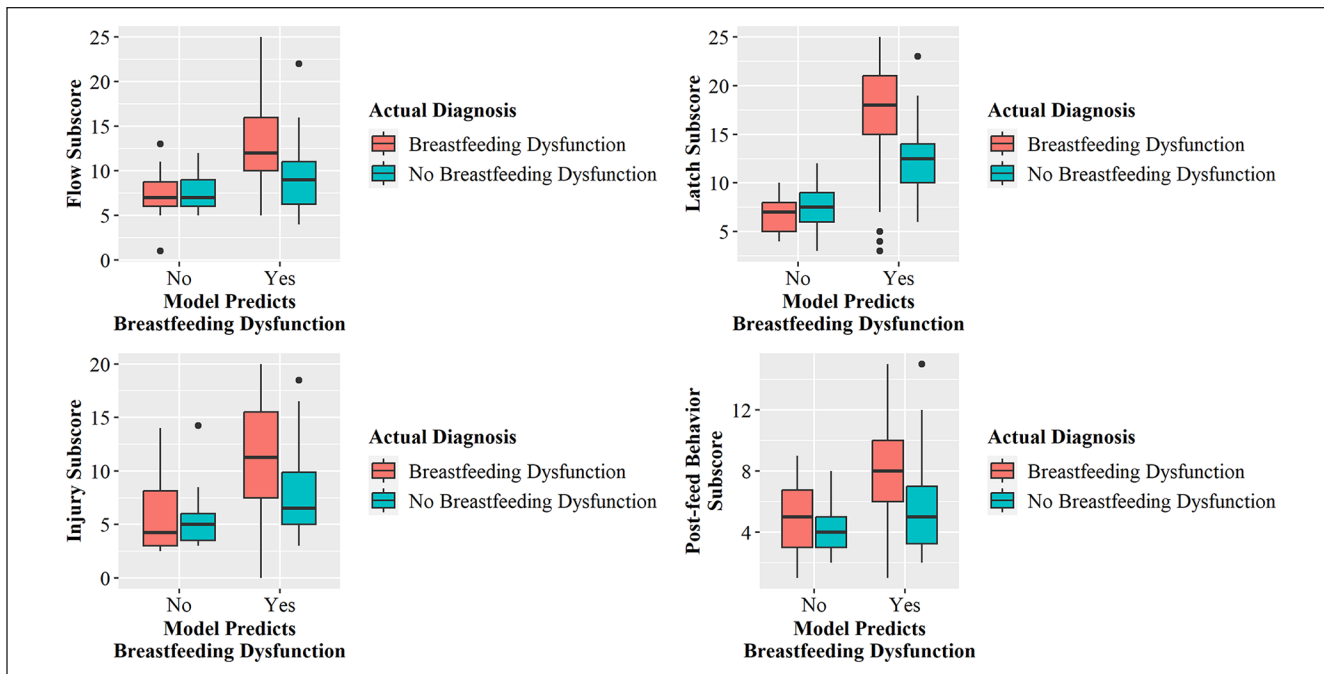


Figure 3. Sub-Score Distribution by Known Groups and Classification Outcome for Flow, Latch, Injury, and Post-Feed Behavior (FLIP). Note. Boxes correspond to the 25th and 75th percentile with median indicated by the horizontal line. Vertical lines define 1.5 times the interquartile range above and below the boxes. Individual points represent data beyond 1.5 times interquartile range.

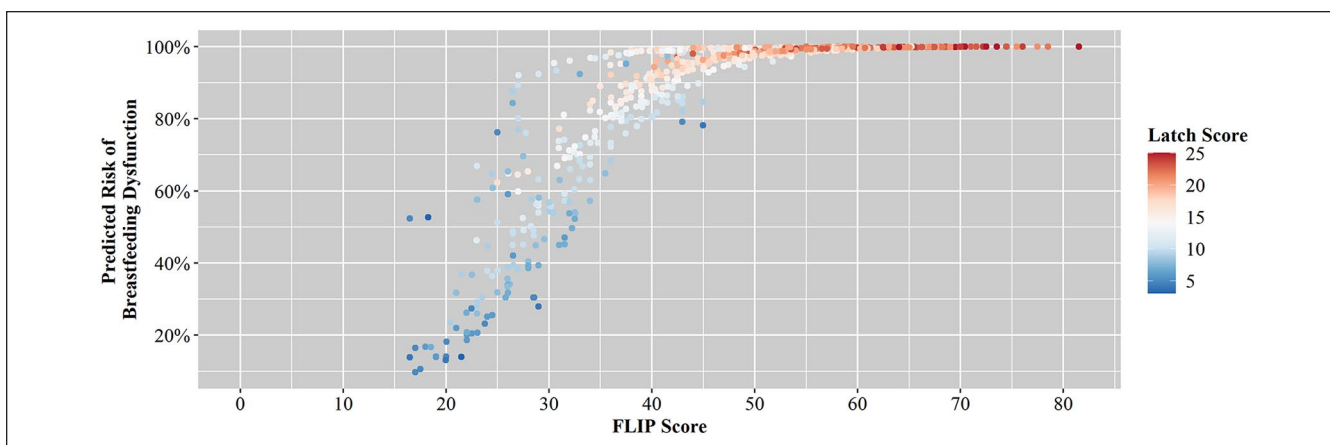


Figure 4. The Relationship Between Predicted Risk of Breastfeeding Dysfunction, FLIP Score, and Latch Sub-Score. Note. The color of the point indicates the Latch sub-score with blue being low scores and red being high scores.

Limitations

The limitations of this study include the wording of the scaled values. As evaluated in this study, the wording could introduce confusion on the part of the parent when rating something they feel happens at a frequency in between two choices. This could explain the ranges we encountered that we then needed to account for by taking the arithmetic mean of the range endpoints. This wording might also introduce self-report and social desirability bias. Future iterations of the tool should use wording that indicates distinct categories:

never, rarely (no more 25% of the time), sometimes (25–75% of the time), often (more than 75% of the time), and always. Additionally, while the risk model accounts for patient age, the number of dyads with children older than 6 months was limited. Further studies using FLIP should assess the validity for use in various other populations with a focus on older infants. Next, the tool was only administered in English in the context of a large pediatrics practice with significant lactation support services. Future studies should evaluate the tool in a wider variety of practice settings. Finally, while all

dyads seen during the survey period were offered the survey, only completed surveys were included in the analysis. This may have introduced unintended bias into the participant population, as we do not know whether the demographics of those dyads who chose not to complete the survey differ from those who did complete the survey.

Conclusions

We have developed and validated a 17-question self-reported questionnaire, FLIP, that can diagnose breastfeeding dysfunction in breastfeeding dyads. This tool grouped participants into three categories: no breastfeeding dysfunction, wait and watch, and breastfeeding dysfunction, based on the total score. The four sub-scores, Flow, Latch, Injury, and Post-Feed Behavior, can be used to further refine clinical management. This is particularly true for the Latch sub-score, as it had the strongest association with misclassification in the underlying risk model. Future studies will assess its usefulness in managing breastfeeding dysfunction.

Acknowledgments

We thank Drs Michelle Mancenido, Abhay Vats, and Laurie B. Jones for their discussions of the data. We also acknowledge Alexandra Lopez's help with data collection, Jilma Joy's help with data entry, Alex Warthen's help with content validation, Samuel Hanson's help with data clean-up and validation, and Qidi Xu's help with data collection, clean-up, and analysis.

Author Contributions

Rajeev Agarwal: Conceptualization; Data curation; Investigation; Supervision; Writing – review & editing.

Mars Eddis-Finbow: Data curation; Investigation; Project administration; Writing – review & editing.

Jodie Tam: Data curation; Writing – original draft; Writing – review & editing.

Jennifer Broatch: Data curation; Formal analysis; Methodology; Visualization; Writing – review & editing.

Kimberly J. Bussey: Data curation; Formal analysis; Visualization; Writing – original draft; Writing – review & editing.

Disclosures and Conflicts of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Rajeev Agarwal  <https://orcid.org/0000-0002-5253-1802>

Kimberly J. Bussey  <https://orcid.org/0000-0003-1001-8207>

Supplemental Material

Supplementary Material may be found in the "Supplemental material" tab in the online version of this article.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (U.S.). (2014). *Standards for educational and psychological testing*. <https://www.apa.org/science/programs/testing/standards>
- Apanga, P. A., Christiansen, E. J., Weber, A. M., Darrow, L. A., Riddle, M. S., Tung, W.-C., Liu, Y., Kohnen, T., & Garn, J. V. (2022). The role of state breastfeeding laws and programs on exclusive breastfeeding practice among mothers in the special supplemental nutrition program for Women, Infants, and Children (WIC). *International Breastfeeding Journal*, 17(1), 46. <https://doi.org/10.1186/s13006-022-00490-9>
- Arizona Revised Statutes—Title 41. (2006). 41–1443. *Arizona State Legislature*. <https://www.azleg.gov/viewdocument/?docName=https://www.azleg.gov/ars/41/01443.htm>
- Arizona Department of Health Services. (2022). *Breastfeeding program*. <http://www.azdhs.gov/prevention/nutrition-physical-activity/breastfeeding/index.php>
- Beauregard, J. L., Hammer, H. C., Chen, J., Avila-Rodriguez, W., Elam-Evans, L. D., & Perrine, C. G. (2019). Racial disparities in breastfeeding initiation and duration among U.S. infants born in 2015. *Morbidity and Mortality Weekly Report*, 68(34), 745. <https://doi.org/10.15585/mmwr.mm6834a3>
- Bickell, M., Barton, C., Dow, K., & Fucile, S. (2018). A systematic review of clinical and psychometric properties of infant oral motor feeding assessments. *Developmental Neurorehabilitation*, 21(6), 351–361. <https://doi.org/10.1080/17518423.2017.1289272>
- Centers for Disease Control and Prevention. (2021). *Breastfeeding Initiation Rates by County or County Equivalent in Arizona*. <https://www.cdc.gov/breastfeeding/data/county/2018-2019/arizona.html>
- Centers for Disease Control and Prevention. (2023). *Public opinions about breastfeeding*. https://www.cdc.gov/breastfeeding/data/healthstyles_survey/index.htm
- Cordray, H., Mahendran, G. N., Tey, C. S., Nemeth, J., Sutcliffe, A., Ingram, J., & Raol, N. (2023). Severity and prevalence of ankyloglossia-associated breastfeeding symptoms: A systematic review and meta-analysis. *Acta Paediatrica*, 112(3), 347–357. <https://doi.org/10.1111/apa.16609>
- Couto, G. R., Dias, V., & de Jesus Oliveira, I. (2020). Benefits of exclusive breastfeeding: An integrative review. *Nursing Practice Today*, 7(4), 245–254. <https://doi.org/10.18502/npt.v7i4.4034>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- Fraga, M. R. B.A., Barreto, K. A., Lira, T. C. B., Celerino, P. R. R.P., Tavares, I. T. S., & Menezes, V. A. (2020). Ankyloglossia and breastfeeding: What is the evidence of association between them? *Revista CEFAC*, 22, e12219. <https://doi.org/10.1590/1982-0216/202022312219>

- Gasparin, V. A., Strada, J. K. R., Moraes, B. A., Betti, T., Gonçalves, A. C., & Santo, L. C. E. (2019). Pairs seen by lactation consultants and cessation of exclusive breastfeeding in the first month. *Revista Da Escola De Enfermagem Da USP*, 53, Article e03422. <https://doi.org/10.1590/S1980-220X2018010003422>
- Gianni, M. L., Bettinelli, M. E., Manfra, P., Sorrentino, G., Bezze, E., Plevani, L., Cavallaro, G., Raffaelli, G., Crippa, B. L., Colombo, L., Morniroli, D., Liotto, N., Roggero, P., Villamor, E., Marchisio, P., & Mosca, F. (2019). Breastfeeding difficulties and risk for early breastfeeding cessation. *Nutrients*, 11(10), Article E2266. <https://doi.org/10.3390/nu11102266>
- Hamner, H. C., Beaugard, J. L., Li, R., Nelson, J. M., & Perrine, C. G. (2021). Meeting breastfeeding intentions differ by race/ethnicity, Infant and Toddler Feeding Practices Study-2. *Maternal & Child Nutrition*, 17(2), e13093. <https://doi.org/10.1111/mcn.13093>
- Hawker, G. A., Mian, S., Kendzerska, T., & French, M. (2011). Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care & Research*, 63(Suppl. 11), S240–S252. <https://doi.org/10.1002/acr.20543>
- Horta, B. L. (2019). Breastfeeding: Investing in the future. *Breastfeeding Medicine*, 14(S1), S-11. <https://doi.org/10.1089/bfm.2019.0032>
- Ingram, J., Johnson, D., Copeland, M., Churchill, C., & Taylor, H. (2015). The development of a new breast feeding assessment tool and the relationship with breast feeding self-efficacy. *Midwifery*, 31(1), 132–137. <https://doi.org/10.1016/j.midw.2014.07.001>
- Jebb, A. T., Ng, V., & Tay, L. (2021). A review of key Likert scale development advances: 1995–2019. *Frontiers in Psychology*, 12, 637547. <https://www.frontiersin.org/articles/10.3389/fpsyg.2021.637547>
- Jensen, D., Wallace, S., & Kelsay, P. (1994). LATCH: A breastfeeding charting system and documentation tool. *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, 23(1), 27–32. <https://doi.org/10.1111/j.1552-6909.1994.tb01847.x>
- Kent, J. C., Hepworth, A. R., Sherriff, J. L., Cox, D. B., Mitoulas, L. R., & Hartmann, P. E. (2013). Longitudinal changes in breastfeeding patterns from 1 to 6 months of lactation. *Breastfeeding Medicine*, 8(4), 401–407. <https://doi.org/10.1089/bfm.2012.0141>
- Krabbe, J., & Forkmann, T. (2012). Frequency vs. intensity: Which should be used as anchors for self-report instruments? *Health and Quality of Life Outcomes*, 10(1), 107. <https://doi.org/10.1186/1477-7525-10-107>
- Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method*. Butterworths.
- LeFort, Y., Evans, A., Livingstone, V., Douglas, P., Dahlquist, N., Donnelly, B., Leeper, K., Harley, E., & Lappin, S. (2021). Academy of breastfeeding medicine position statement on ankyloglossia in breastfeeding dyads. *Breastfeeding Medicine*, 16(4), 278–281. <https://doi.org/10.1089/bfm.2021.29179.ylf>
- Li, R., Perrine, C. G., Anstey, E. H., Chen, J., MacGowan, C. A., & Elam-Evans, L. D. (2019). Breastfeeding trends by race/ethnicity among US children born from 2009 to 2015. *JAMA Pediatrics*, 173(12), Article e193319. <https://doi.org/10.1001/jamapediatrics.2019.3319>
- Lopes De Castro Martinelli, R., Queiroz Marchesan, I., & Berretin-Felix, G. (2012). Lingual frenulum protocol with scores for infants. *International Journal of Orofacial Myology*, 38, 104–112.
- Matthews, M. K. (1988). Developing an instrument to assess infant breastfeeding behaviour in the early neonatal period. *Midwifery*, 4(4), 154–165. [https://doi.org/10.1016/S0266-6138\(88\)80071-8](https://doi.org/10.1016/S0266-6138(88)80071-8)
- Morrison, A. H., Gentry, R., & Anderson, J. (2019). Mothers' reasons for early breastfeeding cessation. *MCN: The American Journal of Maternal Child Nursing*, 44(6), 325–330. <https://doi.org/10.1097/NMC.0000000000000566>
- Mulford, C. (1992). The Mother-Baby Assessment (MBA): An "Apgar Score" for breastfeeding. *Journal of Human Lactation*, 8(2), 79–82. <https://doi.org/10.1177/089033449200800216>
- North, K., Gao, M., Allen, G., & Lee, A. C. (2022). Breastfeeding in a global context: Epidemiology, impact, and future directions. *Clinical Therapeutics*, 44(2), 228–244. <https://doi.org/10.1016/j.clinthera.2021.11.017>
- Shrestha, N. (2021). Factor analysis as a tool for survey analysis. *American Journal of Applied Mathematics and Statistics*, 9(1), Article 1. <https://doi.org/10.12691/ajams-9-1-2>
- Srinivasan, A., Dobrich, C., Mitnick, H., & Feldman, P. (2006). Ankyloglossia in breastfeeding infants: The effect of frenotomy on maternal nipple pain and latch. *Breastfeeding Medicine*, 1(4), 216–224. <https://doi.org/10.1089/bfm.2006.1.216>
- U.S. Census Bureau. (2023). *QuickFacts: Chandler city, Arizona; Scottsdale city, Arizona; Phoenix city, Arizona; Avondale city, Arizona; Glendale city, Arizona; Arizona*. <https://www.census.gov/quickfacts/fact/table/chandlercityarizona,scottsdalecityarizona,phoenixcityarizona,glendalecityarizona,AZ/INC110221>
- Westerfield, K. L., Koenig, K., & Oh, R. (2018). Breastfeeding: Common questions and answers. *American Family Physician*, 98(6), 368–373. <https://www.aafp.org/pubs/afp/issues/2018/0915/p368.html>
- Whipps, M. D. M., Yoshikawa, H., & Demirci, J. R. (2019). Latent trajectories of infant breast milk consumption in the United States. *Maternal & Child Nutrition*, 15(1), Article e12655. <https://doi.org/10.1111/mcn.12655>
- World Health Organization. (2008). *Indicators for assessing infant and young child feeding practices: Part 1—Definitions. Conclusions of a Consensus Meeting Held 6–8 November 2007, in Washington, DC, USA*. WHO Global Consensus Meeting on Indicators of Infant and Young Child Feeding, Washington, DC. <https://www.who.int/publications-detail-redirect/9789241596664>
- Zhang, Q., Li, K., Wouk, K., Lamichhane, R., & Guthrie, J. (2023). Prenatal perception of WIC breastfeeding recommendations predicts breastfeeding exclusivity and duration in the infants' first year. *The Journal of Nutrition*, 152(12), 2931–2940. <https://doi.org/10.1093/jn/nxac221>