 ORIGINAL ARTICLES

Systematic Review and Meta-Analysis: Fussing and Crying Durations and Prevalence of Colic in Infants

Dieter Wolke, PhD, Dr rer nat h.c.[1,2], Ayten Bilgin, MSc[1], and Muthanna Samara, PhD[3]

Objective To determine the mean duration of fussing and crying and prevalence of colic using modified Wessel criteria in infants in the first 3 months of life.

Study design A systematic literature search was performed using the databases Medline, PsycINFO, and Embase. The major outcome measure was mean total fuss/cry duration during 24 hours at ages 1-2 weeks (11 samples), 3-4 weeks (6 samples), 5-6 weeks (28 samples), 8-9 weeks (9 samples), and 10-12 weeks (12 samples).

Results Of 5687 articles reviewed, 28 diary studies (33 samples) were suitable for inclusion in meta-analysis; these studies included 8690 infants. No statistical evidence for a universal crying peak at 6 weeks of age across studies was found. Rather, the mean fuss/cry duration across studies was stable at 117-133 minutes (SDs: 66-70) in the first 6 weeks and dropped to a mean of 68 minutes (SD: 46.2) by 10-12 weeks of age. Colic was much more frequent in the first 6 weeks (17%-25%) compared with 11% by 8-9 weeks of age and 0.6% by 10-12 weeks of age, according to modified Wessel criteria and lowest in Denmark and Japan.

Conclusions The duration of fussing/crying drops significantly after 8-9 weeks of age, with colic as defined by modified Wessel criteria being rare in infants older than 9 weeks. Colic or excessive fuss/cry may be more accurately identified by defining fuss/cry above the 90th percentile in the chart provided based on the review. *(J Pediatr 2017;185:55-61).*

Colic is a common source of concern for parents, a frequent reason for seeking help and advice from healthcare professionals.[1,2] It may be a trigger for abusive head trauma.[3] However, definitions for colic vary widely ranging from gastrointestinal symptoms[4,5] to inconsolable crying,[6] which has resulted in variations of the reported prevalence rate from 1.5% to 11.9%.[7] Increasingly colic is defined in terms of total daily duration of fussing and crying.[8-11] The most widely used definition for colic is the "Rule of Three's"[12]: an infant is considered to have colic if the infant fusses or cries for >3 hours, >3 days per week, for >3 weeks. However, it is impracticable for parents to assess and document fuss/cry duration for a 3-week period using detailed diaries.[13] Thus, the "modified Wessel criteria" are most often used, requiring the infant to have fussed/cried for more than 3 hours a day, on at least 3 days in any 1 week.[12,14,15]

Although the modified Wessel criteria are widely used, normative studies in the general infant population are lacking.[15] Considerable changes in infant care have occurred since the Wessel criteria were published,[16,17] thus, the prevalence may be different than that noted in the 1950s. Although modified Wessel criteria have been used in different countries, the impact of cultural variations such as caregiving styles[18-23] and maternal soothing techniques[23,24] on the duration of infant fussing and crying need to be taken into consideration. Furthermore, the cry/fuss duration may depend on the patterns of feeding (breast vs bottle).[25] Finally, several studies have documented a developmental pattern of fuss/cry duration in the first 3 months of life,[1,26-29] indicating a gradual increase that peaks at 5-6 weeks of age with a decrease to one-half the amount by 3-4 months of age.[18,22,27,28,30,31] This "normal crying curve" has been interpreted as universal across cultures,[18] although some have not found evidence.[20,32,33]

We conducted a systematic review and meta-analysis of fuss/cry durations reported in diary studies from around the world. Twenty-four-hour behavior diaries are considered to be the international gold standard for measurement.[14,34-36] We investigated the change in fuss/cry duration over the first 12 weeks of life to determine if there is a universal "crying curve" (5- to 6-week fuss/cry duration peak) and if mean fuss/cry duration varies across studies in different countries, according to feeding type or study quality. We also determined the prevalence of colic according to the modified Wessel criteria at different ages in the first 12 weeks.

Methods

The current meta-analysis was conducted in line with PRISMA guidelines.[37] We searched the databases Medline (1964-December 2015), PsycINFO (1964-December 2015), and Embase (1964-December 2015) using the search headings "infant and crying" OR "crying and amount" OR "crying and duration" OR "fussing and infant." In addition, infant cry researchers who had participated in the International Cry Research Workshops were approached concerning unpublished

From the [1]Department of Psychology, University of Warwick, Coventry; [2]Division of Mental Health and Wellbeing, Warwick Medical School, University of Warwick, Coventry, United Kingdom; and [3]Department of Psychology, Kingston University London, Kingston upon Thames, London

A.B. is supported by a PhD scholarship from the Republic of Turkey Ministry of Education. The authors declare no conflicts of interest.

http://dx.doi.org10.1016/j.jpeds.2017.02.020

data. Finally, we conducted a separate bibliography search and included all new relevant research.

Criteria for inclusion of articles in the analysis were as follows: (1) at least one 24-hour behavior diary to measure fuss/cry duration; (2) unselected sample (ie, no infants had been excluded according to fuss/cry duration [eg, only colic infants or all noncolic infants]); (3) observation study (ie, no intervention trial); (4) infant age between 1 and 13 weeks; and (5) the authors reported (or provided after request) mean fuss/cry duration as well as distribution indices (ie, SD). For the colic prevalence analysis, only the studies that reported at least three 24-hour behavior diaries were included to meet the modified Wessel criteria. Abstracts were screened according to the selection and inclusion criteria explained above by 2 authors, each screening one-half of the abstracts. Study selection and data extraction were performed independently by 2 authors.

The quality of studies was evaluated according to 8 criteria: (1) subject selection (whole vs convenience population), (2) recruitment rate (≥50% vs <50%), (3) participation rate (≥75% vs <75%), (4) sample size (≥ 101 vs <101), (5) whether the following 4-sample characteristics were reported: socioeconomic status, parity, infant sex, and maternal age (3 of 4 reported vs <3 reported), (6) feeding type (reported vs not reported), (7) resolution time for the diary (5 vs 15 minutes), (8) number of days requested for diary (≥4 vs <4 days), and (9) whether modified Wessel criteria were employed. Each sample, at each measurement age, received a score of 0 or 1 for each of the criteria. A score of 0 was also given in cases where the information for the criterion was not reported. The individual scores were summed to give a total quality score that could range from 0 to 8 (**Table I**; available at www.jpeds.com).

The major outcome measure was mean total fuss/cry duration during 24 hours. The studies were grouped according to age at assessment: 1-2 weeks (11 samples), 3-4 weeks (6 samples); 5-6 weeks (28 samples), 8-9 weeks (9 samples), and 10-12 weeks (12 samples). Furthermore, information regarding the sample size and feeding type (bottle-fed, breastfed, mixed) was extracted from the articles.

To test fuss/cry peak duration, we calculated a weighted mean and the pooled weighted SD for each period. To test for mean differences, ANOVA was performed between individual weighted means. To evaluate the prevalence of colic, 3 samples which used the diaries for less than 3 days were removed from the 8- to 9-week analysis. The prevalence of colic for each study at each assessment point was computed according to the modified-Wessel criteria, and overall prevalence rates are reported as weighted mean and pooled weighted SD.

Meta-analysis was conducted with the comprehensive meta-analysis software.[38] Effect sizes are reported as standardized mean difference with 95% CIs for each study. The mean difference (Cohen *d*) compares the individual study's mean with the overall weighted mean across studies at each assessment time. A *d* of .20 is a small, .50 medium, and .80 or more a large effect.[39] Effect sizes were analyzed using the random effects model, in which the error term is composed of variation originating from both within-study variability and between-study differences.[40,41]

The distribution of effect sizes was examined using tests of heterogeneity. Significant heterogeneity indicates that differences across effect sizes are likely due to sources other than sampling error, such as different study characteristics. Categorical moderator tests were applied to test for within groups Q (Q_w) and between groups Q (Q_b). A significant value for Q_w indicates that the effect sizes within a category of the moderator variable are heterogeneous, whereas a significant value for Q_b indicates that the effect sizes are significantly different across different categories of the moderator variable. Meta-regression analyses were performed to test quality of assessment as a continuous moderator.

We examined the potential for publication bias by using 2 methods suggested for observational studies. First, biases according to study size were assessed with use of the Begg and Mazumdar[42] rank correlation test (Kendall tau b). Second, the Duval and Tweedie[43,44] "trim and fill" method was applied.

Results

The online search yielded 5680 articles. An additional seven potential studies were identified through searches of bibliographies and from the Infant Cry Research Workshops (**Figure 1**; available at www.jpeds.com). After removing the duplicates, the overall systematic literature search included 4109 articles. We reviewed the titles and abstracts of all articles found (N = 4109), resulting in 227 abstracts for joint review. After excluding 138 articles based on their abstract, a total of 89 full-text articles were independently reviewed by 2 authors. Based on the inclusion criteria, 43 articles were further excluded. Among the remaining 46 articles, there were 18 studies with missing data, which required their authors to be contacted to obtain further information about the fuss/cry duration or moderator variables. However, some authors were not able to provide missing data (eg, means, SD etc)[8,14,15,22,45-49] or could not be reached[50-54]; and some studies did not meet the inclusion criteria (eg, selected population, no fuss/cry duration data etc.).[55-58] These studies were, therefore, not included in the meta-analysis. Five study reports[20,21,59-61] reported on more than one sample, resulting in a total of 28 articles with 33 samples being included in the meta-analysis (**Table I**). The majority of the studies used at least 3 days diary except 3 samples from 2 study reports.[60,62] The studies included in the analysis with their quality rating scores and descriptions of each study are shown in **Table I**.

The overall agreement in the selection of articles according to the predefined criteria was Cohen k = 0.89 at the full-text retrieval stage. The discrepancies in articles were discussed and mutually resolved by the coders.

Fuss/Cry Duration across 1-12 Weeks of Age

Mean Fuss/Cry Duration. The weighted mean average for each period was computed (**Figure 2**, A). As shown, mean fuss/cry durations were 117-133 minutes (SDs: 66-70) in the first 6 weeks and then dropped to 68 minutes (SD: 46) by 10-12 weeks of age. Post-hoc comparisons showed that fuss/cry duration did not significantly differ from each other across the

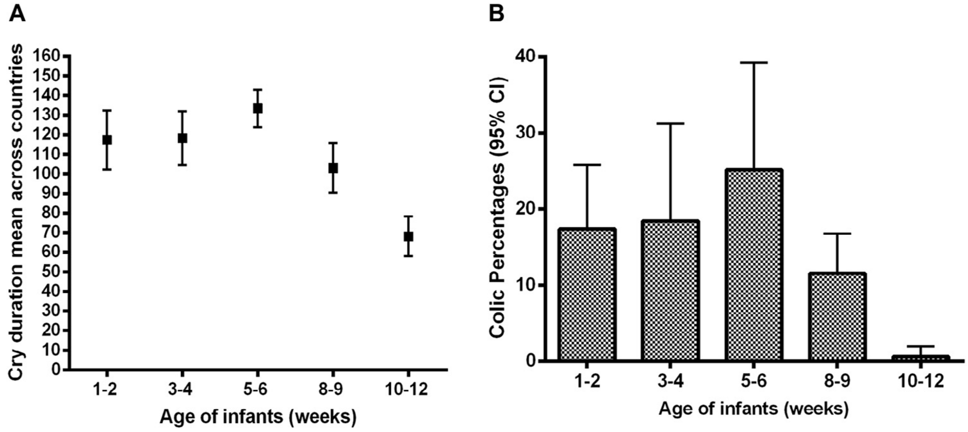


Figure 2. **A,** Weighted mean fuss/cry duration (in minutes) across countries (1-2 vs 10-12 weeks of age [$P < .001$]; 3-4 vs10-12 weeks of age [$P < .01$]; 5-6 vs 10-12 weeks of age [$P < .001$]) and **B,** overall colic percentages (95% CI) for all studies combined in each period (1-2 vs 10-12 weeks of age [$P < .01$]; 3-4 vs 10-12 weeks of age [$P < .001$]; 5-6 vs 10-12 weeks of age [$P < .001$]; 8-9 vs10-12 weeks of age [$P < .05$]; and 5-6 vs 8-9 weeks of age [$P < .05$]).

first 6 weeks. However, the fuss/cry duration at 10-12 weeks (mean: 68.03, SD: 46.2) was significantly lower than at 1-2 weeks of age (mean: 117.3, SD: 66.8; $P < .001$); 3-4 weeks of age (mean: 118.2, SD: 69.3; $P < .01$); and 5-6 weeks of age (mean: 133.3, SD: 70.1; $P < .001$).

Potential Moderator Variables.

Country. Random effects meta-analyses (Figure 3; available at www.jpeds.com) showed that the standardized fuss/cry duration means in Germany at 1-2 (mean: 69, SD: 60) weeks and 3-4 weeks of age (mean: 80.8, SD: 67.4) and in Japan at 5-6 weeks of age (mean: 107, SD: 36) were significantly lower than the overall weighted average mean (Figure 3). Similarly, the standardized fuss/cry duration means in Denmark were significantly lower than the overall weighted average mean across the first 12 weeks except 8-9 weeks of age. On the other hand, the standardized fuss/cry duration mean in Canada at 3-4 weeks of age (mean: 149.8, SD: 73.5) and in The Netherlands at 5-6 weeks of age (mean: 150.4, SD: 66.3) was significantly higher than the overall weighted fuss/cry duration mean (Table II). The Q test for heterogeneity was significant at each age ($P < .001$).

Feeding Type. Feeding type was found to be a significant moderator at ages 1-2 weeks ($Q_b = 22.91$; $P < .001$), 5-6 weeks ($Q_b = 12.28$; $P < .01$), and 10-12 weeks ($Q_b = 21.01$; $P < .001$). Samples which included babies who were bottle-fed ($z = -3.461$; $P < .01$) or mixed-fed (breast and bottle) ($z = -3.656$; $P < .01$) had significantly lower fuss/cry durations than the overall weighted fuss/cry mean at 5-6 weeks of age. In contrast, samples which included babies who were breastfed had significantly higher fuss/cry durations at 3-4 weeks of age ($z = 3.500$; $P < .01$). Furthermore, samples that did not report on the type of feeding (at 1-2 and 10-12 weeks of age) reported significantly higher fuss/cry durations than the overall weighted fuss/cry.

Quality Assessment. Univariate meta-regression analyses indicated a positive significant moderating influence of study quality at 8-9 weeks of age (the slope: point estimate = 0.15; z = 4.09; SE = 0.04; $P < .001$; $Q_b = 16.79$; df: 1; $P < .001$): As study quality increased, fuss/cry duration also increased.

Prevalence of Colic

We calculated the overall mean weighted colic prevalence of all studies at each assessment point (Figure 2, B). Mean colic prevalence at 10-12 weeks age (0.6%) was significantly lower than the mean colic prevalence at 1-2 weeks of age (17.4%, z = 2.95; $P < .01$), 3-4 weeks of age (18.4%, z = 3.40; $P < .001$), 5-6 weeks of age (25.1%, z = 3.64; $P < .001$), and 8-9 weeks of age (10.8%, z = 2.93, $P < .01$). Furthermore, it was found that colic prevalence at 5-6 weeks of age was significantly higher than colic prevalence at 8-9 weeks of age (z = 2.01; $P < .05$).

Potential Moderator Variables. Significant moderating effects were observed for country at 1-2 weeks of age ($Q_b = 16.24$; $P < .01$), 3-4 weeks of age ($Q_b = 22.91$; $P < .001$), and 8-9 weeks of age ($Q_b = 9.44$; $P < .05$). The average standardized difference (*d*) in mean colic prevalence of the United Kingdom studies at 1-2 weeks of age (28%), Canada at 3-4 weeks of age (34.1%), and Italy at 8-9 weeks of age (20.9%) was significantly higher than the overall weighted colic prevalence. In contrast, Denmark (5.5%) and Germany (6.7%) had lower colic rates at 3-4 weeks of age (Table II).

Although country was not a significant moderator at 5-6 weeks of age, fewer infants with colic were reported across all Danish studies (6.7%), the Japanese study (2.1%), and the United Kingdom studies (18.1%) compared with the overall prevalence.

Feeding type was found to be a significant moderator at 5-6 weeks of age ($Q_b = 14.23$; $P < .01$) and 10-12 weeks of age ($Q_b = 4.55$; $P < .05$). At 5-6 weeks of age, studies that

Table II. Countries with statistically significant deviations from mean fuss/cry duration and colic percentages compared with the overall mean across all countries (all studies)

	Fuss/cry duration					
	Number of studies	Overall mean (SD) in min	Mean (SD) in min	Higher or lower than overall	Z score	P
1-2 wk						
Denmark	1	117.3 (66.8)	80 (44)	Lower	–3.86	<.001
Germany	1	117.3 (66.8)	69 (60)	Lower	–4.55	<.001
3-4 wk						
Canada	1	118.2 (69.3)	149.8 (73.5)	Higher	3.50	<.001
Denmark	1	118.2 (69.3)	90 (58)	Lower	–3.24	<.001
Germany	1	118.2 (69.3)	80.8 (67.4)	Lower	–3.78	<.001
5-6 wk						
Denmark	3	133.3 (70.1)	85.6 (64.7)	Lower	–11.27	<.001
Japan	1	133.3 (70.1)	107 (36)	Lower	–2.04	<.05
The Netherlands	1	133.3 (70.1)	150.4 (66.3)	Higher	2.05	<.05
10-12 wk						
Denmark	1	68 (46.2)	48 (44)	Lower	–3.01	<.01

	Colic percentage					
	Number of studies	Overall colic percentage	Colic percentage	Higher or lower than overall	Z score	P
1-2 wk						
United Kingdom	7	17.4%	28%	Higher	2.73	<.01
3-4 wk						
Canada	1	18.4%	34.1%	Higher	2.76	<.01
Denmark	1	18.4%	5.5%	Lower	–2.84	<.001
Germany	1	18.4%	6.7%	Lower	–2.36	<.001
5-6 wk*						
Denmark	3	25.1%	6.7%	Lower	–7.22	<.001
Japan	1	25.1%	2.1%	Lower	–2.16	<.05
United Kingdom	7	25.1%	18.1%	Lower	–2.56	<.05
8-9 wk						
Italy	2	11.5%	20.9%	Higher	1.98	<.05

Please note that this only illustrates the findings for countries that were significantly different from overall.
*Country overall was not a significant moderator.

reported infants who were bottle-fed ($z = -3.87$; $P < .001$) and mixed fed ($z = -3.54$; $P < .001$) had lower prevalence of colic. On the other hand, at 10-12 weeks of age, studies that did not report the feeding type (6 studies) had significantly higher colic prevalence ($z = 2.62$; $P < .05$) compared with overall weighted colic prevalence.

The heterogeneity analysis was significant at the following ages: 1-2 weeks: Q = 29.42; $P < .01$; 3-4 weeks: Q = 24.87; $P < .001$; and 5-6 weeks: Q = 74.57; $P < .001$.

Quality Assessment. Univariate meta-regression analyses showed that study quality had a positive significant moderating influence at 8-9 weeks of age (the slope: point estimate = 0.53; z = 2.57; SE = 0.21; $P < .05$; $Q_b = 6.61$; df: 1; $P < .5$). Increased quality of study was associated with increased prevalence of colic.

Publication Bias

The Begg and Mazumdar Rank Correlation Test (correlation between study size and effect size) suggest that there was little evidence for publication bias. We assessed the possibility of publication bias by using a funnel plot to assess for asymmetry. The Duval and Tweedie's trim and fill method indicates that 2 studies are missing left to the mean at 1-2 weeks of age (combined studies: 0.01; 95% CI: –0.23 to 0.26; using trim and fill, the imputed point estimate: –0.08; 95% CI: –0.32 to 0.16), 3 studies are missing left to the mean at 3-4 weeks of age (combined studies: –0.02; 95% CI: –0.38 to 0.34; using trim and fill, the imputed point estimate: 0.29; 95% CI: –0.67 to 0.09), and 3 studies are missing left to the mean at 10-12 weeks of age (combined studies: 0.14; 95% CI: –0.09 to 0.38; using trim and fill, the imputed point estimate: –0.01; 95% CI: –0.26 to 0.24).

Discussion

This review and meta-analysis found no statistical evidence for a "universal" increase of fuss/cry duration over the first 6 weeks of life culminating in a "crying peak" at 5-6 weeks of age as proposed previously,[28,29] although visual inspection shows a slight increase.[63] Overall, fuss/cry durations were high across the first 6 weeks of life, then reduced significantly over the following 6 weeks. All studies found a "universal" reduction in fuss/cry duration between 6 and 12 weeks of age.

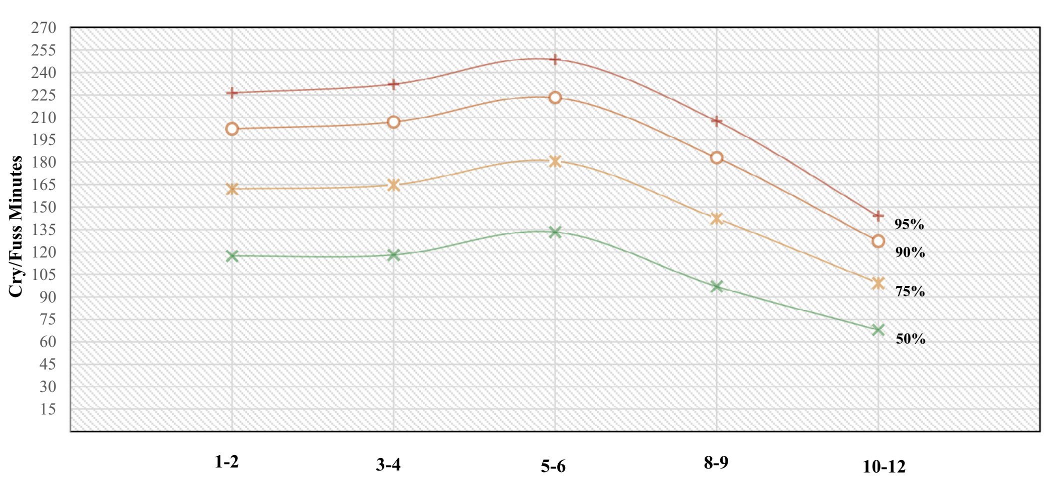
The significant differences in mean fuss/cry durations between studies were moderated by country of origin. The most consistent finding was the lower fuss/cry durations reported in Denmark at several age points.[32] On the other hand, with the exception of Denmark, no consistent pattern for significantly higher fuss/cry duration between other countries was found. At different assessment ages, studies from The Netherlands[64] and Canada[65,66] had significantly higher fuss/cry durations compared with the overall mean weighted fuss/cry duration.

The findings regarding the country differences appear robust according to publication bias results. However, we can only speculate on the reasons why there are country differences, in particular between Denmark and the rest of Europe and North America. These could range from economic conditions, such as less social inequality, to caretaking patterns such as responsiveness, carrying behavior and management in Denmark that have been shown to differ from the United Kingdom.[20] However, there may also be population genetic differences, and the infants both inherit their parents' genes and are reared by them (gene-environment correlation).[67] Nevertheless, further analysis of caretaking patterns may prove to provide clues for effective preventative strategies.

Feeding type was a further moderator of fuss/cry duration. Bottle or mixed feeding was associated with reduced duration of fussing and crying or colic from 3-4 weeks of age onward. Switch in feeding type is one frequently adopted method by parents dealing with a crying baby[68] and has been found to reduce crying regardless of what formula change is instituted, suggesting a placebo effect.[69] Feeding type has also been previously reported to be associated with more night waking in infants.[70,71] Night waking is often signaled by fussing or crying and, thus, may have increased the total fuss/cry duration in diary reports in those breast feeding. Alternatively, cultural differences might have influenced the accuracy of diary keeping. Furthermore, mothers' perception of the frequency of their infants' crying might be enhanced by cultural variations in support for shouldering the burden in caring for their infant.

The prevalence of colic according to Wessel modified criteria ranged from 17% to 25% in the first 6 weeks, then decreased to 11% by 8-9 weeks of age and finally, to only 0.6% by 10-12 weeks of age. Notably, the lowest colic prevalence rate was found for Danish infants (6%) and Japanese (2%) infants during the first 6 weeks. In contrast, the highest mean prevalence rate was found for the United Kingdom studies at 1-2 weeks of age (ranged from 17% to 47%). If colic is considered the extreme of a normal distribution of fuss/cry duration, then it is not surprising that fewer infants with colic were found in Denmark where the mean fuss/cry duration was lower than in other countries. However, if alternatively, colic is considered to be qualitatively different from normal fussing and crying[8] then a similar prevalence should have been found across countries. Our findings are consistent with the first interpretation that colic reflects the extreme of normal fuss/cry distribution. Further, but less consistent, moderation of colic prevalence was found by feeding type. There was a weak trend for infants who were bottle or mixed-fed to have lower prevalence of colic at 5-6 weeks of age.

There are strengths and limitations that require comment. The study only included those with diary measures: all but 3[60,72] used 5-minute resolution, and most samples had 3 or more days of diaries. On the other hand, there were unequal numbers of studies from different countries. Although we identified empirical studies in Australia, Iran, and Korea, despite contacting the authors, the required distribution measures were not available. Thus, this is a review of studies in North America and parts of Europe with only 1 study from Japan. No studies from threshold or developing countries were available, but these would be needed to provide adequate feedback to parents on other continents. Feeding type information was also not available for some studies. We consider it unlikely that relevant studies with diary data were missed, but it cannot be excluded. Moreover, there might be a loss of studies in the title and abstract screening procedure, which was conducted by 1 author. Furthermore, multiple statistical comparisons were not



Please note that this is an approximation of average to high fuss/cry amount percentiles based on the samples included in the meta-analysis. It should only be used as a rough guide to identify excessively crying infants according to age.

Figure 4. Percentile chart of above average fuss/cry at the ages 1-2, 3-4, 5-6, 8-9, and 10-12 weeks in infants.

Bonferroni adjusted and need to be interpreted cautiously. Finally, the lack of a significant peak at 5-6 weeks of age should be interpreted cautiously because our study might be underpowered to detect a small peak amounting to 15 minutes.

There are several implications for research on colic and clinical practice. Firstly, colic is best defined as the extreme of the distribution of fuss/cry duration. Secondly, the cut-off points need to take into account the rapid developmental changes occurring in fuss/cry durations during the first 3 months. The modified Wessel criteria may have served researchers or clinicians well for more than sixty years but may be inaccurate when applied at any time across the first 3 months. As shown here, fuss/cry duration is highest in the first 6 weeks and reduces rapidly during the next 6 weeks. Thus cut-off points need to be determined for the first 6 weeks of age, at 8-9 weeks of age, and 10-12 weeks of age. **Figure 4** provides a percentile chart for "average" and "excessive" fussing and crying on basis of this meta-analysis across countries, allowing clinicians an approximation of whether the infant is excessively fussing or crying according to age or within the normal range. This feedback to parents is a first step of education on fussing or crying and whether their infant's fuss/cry is within the normal range. Those above the 90th percentile may be identified as excessive criers or infants with colic. As the mean fuss/cry duration was found to vary between countries, future normative country specific studies may be required on representative samples. The rapid developmental change in fuss/cry duration has implications for treatment and interpretation of treatment studies. Colic is the extreme of normal fuss/cry behavior, self-limiting, and, thus, the vast majority will spontaneously remit.[29] Adequate management of fussing and crying in the first 3 months rather than treatment may be required.[73] However, if excessive fuss/cry persists beyond the first 3 months, there is increasing evidence that this may indicate regulatory problems with adverse consequences for future development and may require treatment.[74] ■

We would like to thank Marissa Alvarez, PhD, for her contributions to the study.

Submitted for publication Aug 26, 2016; last revision received Jan 6, 2017; accepted Feb 7, 2017

Reprint requests: Dieter Wolke, Ph.D, Dr rer nat h.c, Department of Psychology, University of Warwick, Coventry CV4 7AL, United Kingdom. E-mail: D.Wolke@warwick.ac.uk

References

1. St James-Roberts I, Halil T. Infant crying patterns in the first year: normal community and clinical findings. J Child Psychol Psychiatry 1991;32:951-68.
2. Forsyth BW, Leventhal JM, McCarthy PL. Mothers' perceptions of problems of feeding and crying behaviors. A prospective study. Am J Dis Child 1985;139:269-72.
3. Talvik I, Alexander RC, Talvik T. Shaken baby syndrome and a baby's cry. Acta Paediatr 2008;97:782-5.
4. Lehtonen LA, Rautava PT. Infantile colic: natural history and treatment. Curr Probl Pediatr 1996;26:79-85.
5. Crowcroft N. Effectiveness of treatment for infatile colic. Findings apply only to the most severely affected infants. BMJ 1998;317:1451-2.
6. Clifford TJ, Campbell MK, Speechley KN, Gorodzinsky F. Sequelae of infant colic: evidence of transient infant distress and absence of lasting effects on maternal mental health. Arch Pediatr Adolesc Med 2002;156:1183-8.
7. Reijneveld SA, Brugman E, Hirasing RA. Excessive infant crying: the impact of varying definitions. Pediatrics 2001;108:893-7.
8. Barr RG, Paterson JA, MacMartin LM, Lehtonen L, Young SN. Prolonged and unsoothable crying bouts in infants with and without colic. J Dev Behav Pediatr 2005;26:14-23.
9. Iacono G, Carroccio A, Montalto G, Cavataio F, Bragion E, Lorello D, et al. Severe infantile colic and food intolerance: a long-term prospective study. J Pediatr Gastroenterol Nutr 1991;12:332-5.
10. Lothe L, Lindberg T, Jakobsson I. Cow's milk formula as a cause of infantile colic: a double-blind study. Pediatrics 1982;70:7-10.
11. Taubman B. Parental counseling compared with elimination of cow's milk or soy milk protein for the treatment of infant colic syndrome: a randomized trial. Pediatrics 1988;81:756-61.
12. Wessel MA, Cobb JC, Jackson EB, Harris GS, Detwiler AC. Paroxysmal fussing in infancy: sometimes called "colic". Pediatrics 1954;14:421-34.
13. Lucassen PL, Assendelft WJ, van Eijk JT, Gubbels JW, Douwes AC, van Geldrop WJ. Systematic review of the occurrence of infantile colic in the community. Arch Dis Child 2001;84:398-403.
14. St. James-Roberts I, Hurry J, Bowyer J. Objective confirmation of crying durations in infants referred for excessive crying. Arch Dis Child 1993;68:82-4.
15. Barr RG, Rotman A, Yaremko J, Leduc D, Francoeur TE. The crying of infants with colic: a controlled empirical description. Pediatrics 1992;90:14-21.
16. Lynn R. What has caused the Flynn effect? Secular increases in the Development Quotients of infants [Review]. Intelligence 2009;37:16-24.
17. Collishaw S, Maughan B, Natarajan L, Pickles A. Trends in adolescent emotional problems in England: a comparison of two national cohorts twenty years apart. J Child Psychol Psychiatry 2010;51:885-94.
18. Barr RG, Konner M, Bakeman R, Adamson L. Crying in !Kung San infants: a test of the cultural specificity hypothesis. Dev Med Child Neurol 1991;33:601-10.
19. Lee K. The crying pattern of Korean infants and related factors. Dev Med Child Neurol 1994;36:601-7.
20. St. James-Roberts I, Alvarez M, Csipke E, Abramsky T, Goodwin J, Sorgenfrei E. Infant crying and sleeping in London, Copenhagen and when parents adopt a "proximal" form of care. Pediatrics 2006;117:1146-55.
21. Lucas A, St. James-Roberts I. Crying, fussing, and colic behaviour in breast- and bottle-fed infants. Early Hum Dev 1998;53:9-18.
22. Hunziker UA, Barr RG. Increased carrying reduces infant crying: a randomized controlled trial. Pediatrics 1986;77:641-8.
23. van der Wal MF, van den Boom DC, Pauw-Plomp H, de Jonge GA. Mothers' reports of infant crying and soothing in a multicultural population. Arch Dis Child 1998;79:312-7.
24. St James-Roberts I, Bowyer J, Varghese S, Sawdon J. Infant crying patterns in Manali and London. Child Care Health Dev 1994;20:323-37.
25. Alison J. Increasing breastfeeding rates: do changing demographics explain them? Womens Health Issues 2007;17:84-92.
26. Alvarez M, St James-Roberts I. Infant fussing and crying patterns in the first year in an urban community in Denmark. Acta Paediatr 1996;85:463-6.
27. Brazelton TB. Crying in infancy. Pediatrics 1962;29:579-88.
28. Barr RG. The early crying paradox: a modest proposal. Hum Nat 1990;1:355-89.
29. Barr RG, Rivara FP, Barr M, Cummings P, Taylor J, Lengua LJ, et al. Effectiveness of educational materials designed to change knowledge and behaviors regarding crying and shaken-baby syndrome in mothers of newborns: a randomized, controlled trial. Pediatrics 2009;123:972-80.
30. Rebelsky F, Black R. Crying in infancy. J Gen Psychol 1972;121:49-57.
31. Barr RG. Crying in the first year of life: good news in the midst of distress. Child Care Health Dev 1998;24:425-39.
32. Alvarez M. Caregiving and early infant crying in a Danish community. J Dev Behav Pediatr 2004;25:91-8.

33. St. James-Roberts I, Plewis I. Individual differences, daily fluctuations, and developmental changes in amounts of infant waking, fussing, crying, feeding, and sleeping. Child Dev 1996;67:2527-40.
34. Salisbury A, Minard K, Hunsley M, Thoman EB. Audio recording of infant crying: comparison with maternal cry logs. Int J Behav Dev 2001;25:458-65.
35. Barr RG, Kramer MS, Boisjoly C, McVey-White L, Pless IB. Parental diary of infant cry and fuss behaviour. Arch Dis Child 1988;63:380-7.
36. St James-Roberts I, Conroy S, Wilsher K. Bases for maternal perceptions of infant crying and colic behaviour. Arch Dis Child 1996;75:375-84.
37. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009;6:e1000097.
38. Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive meta-analysis (Version 2) [Computer Software]. Englewood (NJ): Biostat; 2005.
39. Cohen J. Statistical power analyses for the behavioral sciences. 2nd ed. Hillsdale (NY): Erlbaum; 1988.
40. Cooper H, Hedges L. The handbook of research synthesis. New York (NY): Russel Sage Foundation; 1994.
41. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Inroduction to meta-analysis. Chichester (UK): John Wiley & Sons; 2009.
42. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994;50:1088-101.
43. Duval S, Tweedie R. A nonparametric trim and fill method of accounting for publication bias in meta-analysis. J Am Stat Assoc 2000a;95:89-98.
44. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000b;56:455-63.
45. Keller H, Chasiotis A, Risau-Peters J, Volker S, Zach U, Restemeier R. Psychobiological aspects of infant crying. Early Dev Parent 1996;5:1-13.
46. Milgrom J, Westley DT, McCloud PI. Do infants of depressed mothers cry more than other infants? J Paediatr Child Health 1995;31:218-21.
47. St. James-Roberts I, Conroy S, Hurry J. Links between infant crying and sleep-waking at six weeks of age. Early Hum Dev 1997;48:143-52.
48. Hyodynmaa E, Tammela O. Cradling in the prevention of excessive crying and colic symptoms in infants. Inf Mental Health J 2005;26:217-30.
49. Kaley F, Reid V, Flynn E. Investigating the biographic, social and temperamental correlates of young infants' sleeping, crying and feeding routines. Infant Behav Dev 2012;35:596-605.
50. Harrison Y. The relationship between daytime exposure to light and night-time sleep in 6-12-week-old infants. J Sleep Res 2004;13:345-52.
51. van Sleuwen BE, L'Hoir MP, Engelberts AC, Busschers WB, Westers P, Blom MA, et al. Comparison of behavior modification with and without swaddling as interventions for excessive crying. J Pediatr 2006;149:512-7, e512.
52. Walker AM, Menahem S. Normal early infant behaviour patterns. J Paediatr Child Health 1994;30:260-2.
53. Walker AM, Menahem S. Intervention of supplementary carrying on normal baby crying patterns: a randomized study. J Dev Behav Pediatr 1994;15:174-8.
54. Saeidi R, Abadi MZL, Saeidi A, Robatsangi MG. The effectiveness of mother infant interaction on infantile colic. Iran J Neurol 2013;4:34-8.
55. St James-Roberts I, Conroy S, Wilsher K. Links between maternal care and persistent infant crying in the early months. Child Care Health Dev 1998;24:353-76.
56. Killerby E. A comparison of cry-behaviours of clinic-referred and non-referred infants 4-12 weeks old. Early Child Dev Care 1992;81:77-100.
57. Baildam EM, Hillier VF, Ward BS, Bannister RP, Bamford FN, Moore WMO. Duration and pattern of crying in the first year of life. Dev Med Child Neurol 1995;37:345-53.
58. Worobey J, Pena J, Ramos I, Espinosa C. Infant difficulty and early weight gain: does fussing promote overfeeding? Matern Child Nutr 2014;10:295-303.
59. Barr RG, Kramer MS, Pless IB, Boisjoly C, Leduc D. Feeding and temperament as determinants of early infant crying/fussing behavior. Pediatrics 1989;84:514-21.
60. Darlington ASE, Wright CM. The influence of temperament on weight gain in early infancy. J Dev Behav Pediatr 2006;27:329-35.
61. Fujiwara T, Barr RG, Brant R, Barr M. Infant distress at five weeks of age and caregiver frustration. J Pediatr 2011;159:425-30, e422.
62. Wake M, Morton-Allen E, Poulakis Z, Hiscock H, Gallagher S, Oberklaid F. Prevalence, stability, and outcomes of cry-fuss and sleep problems in the first 2 years of life: prospective community-based study. Pediatrics 2006;117:836-42.
63. Barr RG, Trent RB, Cross J. Age-related incidence curve of hospitalized Shaken Baby Syndrome cases: convergent evidence for crying as a trigger to shaking. Child Abuse Negl 2006;30:7-16.
64. de Weerth C, Buitelaar JK. Childbirth complications affect young infants' behavior. Eur Child Adolesc Psychiatry 2007;16:379-88.
65. Kramer MS, Barr RG, Dagenais S, Yang H, Jones P, Ciofani L, et al. Pacifier use, early weaning, and cry/fuss behavior: a randomized control trial. JAMA 2001;286:322-6.
66. Miller AR, Barr RG, Eaton WO. Crying and motor behavior of six-week-old infants and postpartum maternal mood. Pediatrics 1993;92:551-8.
67. Rutter M. Genes and behavior: nature-nurture interplay explained. Oxford: Blackwell; 2006.
68. Forsyth BW, McCarthy PL, Leventhal JM. Problems of early infancy, formula changes, and mothers' beliefs about their infants. J Pediatr 1985;106:1012-7.
69. Berseth CL, Johnston WH, Stolz SI, Harris CL, Mitmesser SH. Clinical response to 2 commonly used switch formulas occurs within 1 day. Clin Pediatr (Phila) 2009;48:58-65.
70. Schmid G, Schreier A, Meyer R, Wolke D. Predictors of crying, feeding, and sleeping problems: a prospective study. Child Care Health Dev 2011;37:493-502.
71. Thunström M. Severe sleep problems among infants in a normal population in Sweden: prevalence, severity, and correlates. Acta Paediatr 1999;88:1356-63.
72. Bensel J. Was sagt mir mein Baby, wenn es schreit? Ratingen, Germany: Oberstebring Verlag; 2003.
73. Wolke D, Gray P, Meyer R. Excessive infant crying: a controlled study of mothers helping mothers. Pediatrics 1994;94:322-32.
74. Hemmi MH, Wolke D, Schneider S. Associations between problems with crying, sleeping and/or feeding in infancy and long-term behavioural outcomes in childhood: a meta-analysis. Arch Dis Child 2011;96:622-9.

References

75. St. James-Roberts I, Sleep J, Morris S, Owen C, Gillham P. Use of a behavioural programme in the first 3 months to prevent infant crying and sleeping problems. J Paediatr Child Health 2001;37:289-97.
76. St. James-Roberts I, Menon-Johansson P. Predicting infant crying from fetal movement data: an exploratory study. Early Hum Dev 1999;54:55-62.
77. St. James-Roberts I, Goodwin J, Peter B, Adams D, Hunt S. Individual differences in responsivity to a neurobehavioural examination predict crying patterns of 1-week-old infants at home. Dev Med Child Neurol 2003;45:400-7.
78. St James-Roberts I, Peachey E. Distinguishing infant prolonged crying from sleep-waking problems. Arch Dis Child 2011;96:340-4.
79. Blum NJ, Taubman B, Tretina L, Heyward RY. Maternal ratings of infant intensity and distractibility: relationship with crying duration in the second month of life. Arch Pediatr Adolesc Med 2002;156:286-90.
80. Stifter CA, Spinrad TL. The effect of excessive crying on the development of emotion regulation. Infancy 2002;3:133-52.
81. Stifter CA, Bono M, Spinrad T. Parent characteristics and conceptualizations associated with the emergence of infant colic. J Reprod Infant Psychol 2003;21:309-22.
82. McRury JM, Zolotor AJ. A randomized, controlled trial of a behavioral intervention to reduce crying among infants. J Am Board Fam Med 2010;23:315-22.
83. Bonichini S, Axia G, St James-Roberts I, DeCian S. Infant crying and maternal holding in the first 2 months of age: An Italian diary study. Infant Child Dev 2008;17:581-92.
84. Mazzotti S, Bonichini S, Axia G. Patterns of infant cry from 0 to 3 months in an Italian sample. [Italian]. Imago 2003;10:331-43.
85. Keller H, Lohaus A, Volker S, Cappenberg M, Chasiotis A. Relationship between infant crying, birth complications, and maternal variables. Child Care Health Dev 1998;24:377-94.
86. Lohaus A, Keller H, Voelker S. Relationships between eye contact, maternal sensitivity, and infant crying. Int J Behav Dev 2001;25:542-8.
87. Wurmser H, Rieger M, Domogalla C, Kahnt A, Buchwald J, Kowatsch M, et al. Association between life stress during pregnancy and infant crying in the first six months postpartum: a prospective longitudinal study. Early Hum Dev 2006;82:341-9.
88. Sondergaard C. Infant crying and fussing pattern in a follow-up study of 432 infants. Infant Crying Workshop: Taby, Sweden;2000.
89. Shinohara H, Kodama H. Relationship between duration of crying/fussy behavior and actigraphic sleep measures in early infancy. Early Hum Dev 2012;88:847-52.

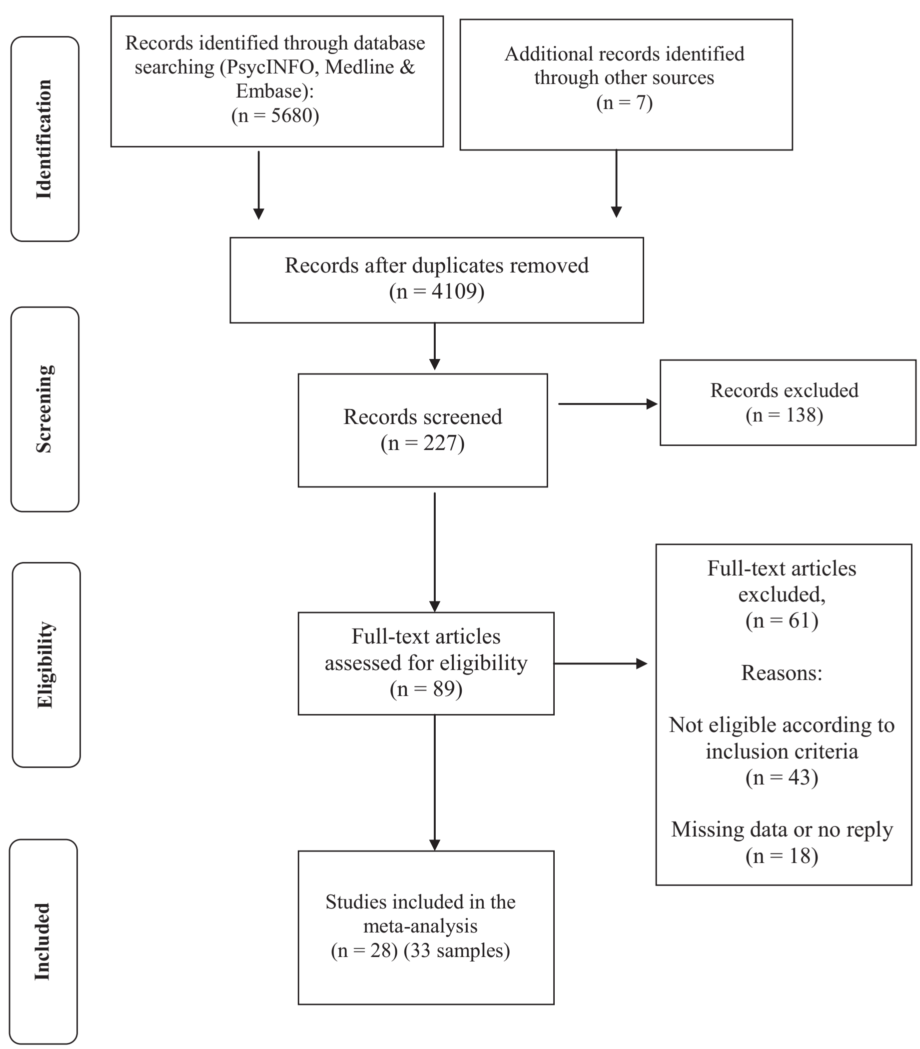
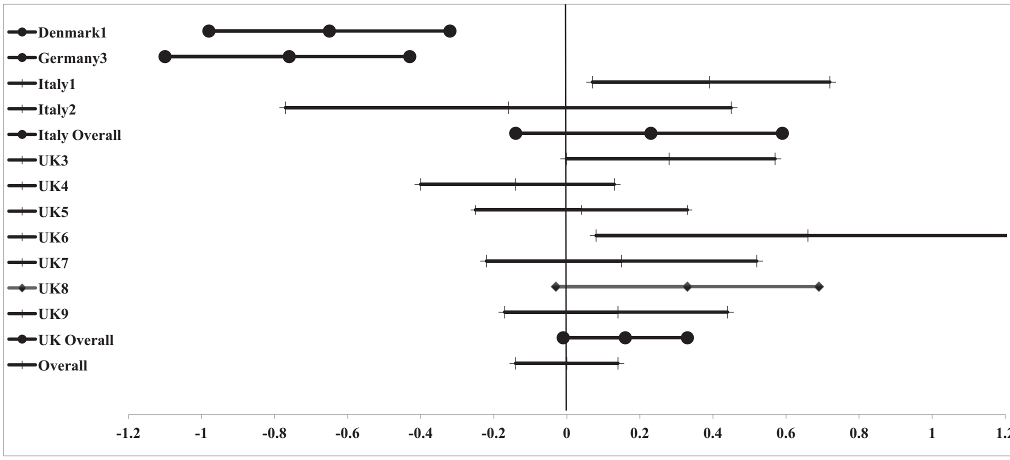
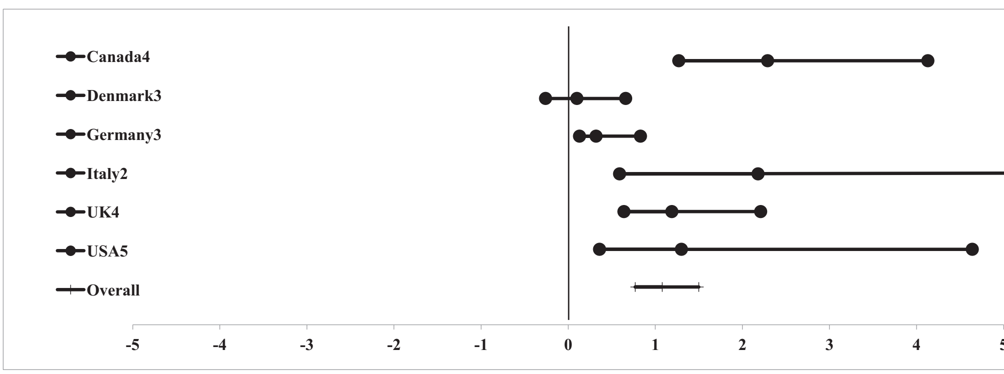


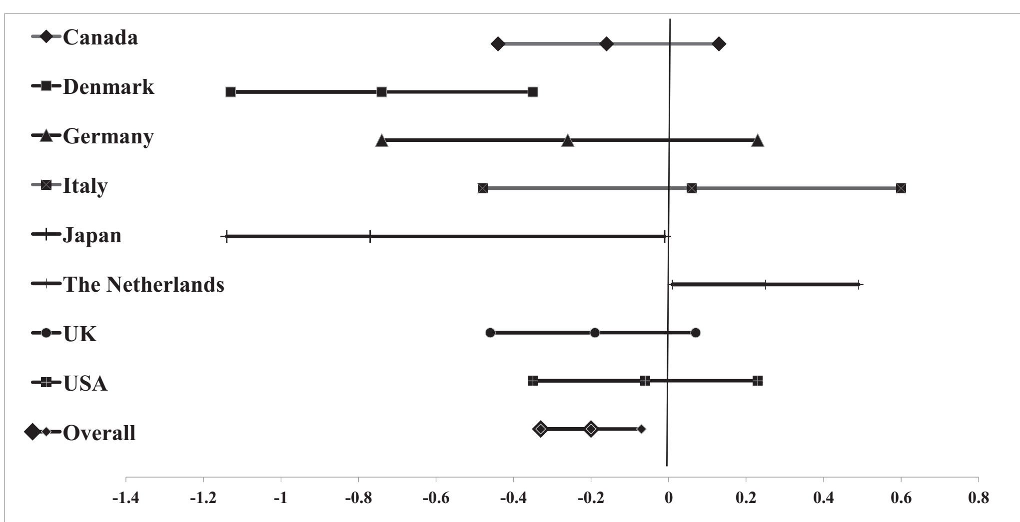
Figure 1. Search strategy for the systematic review.



(a)



(b)



Please note that only the overall country effect sizes are shown for each country. The overall effect size at the bottom indicates the finding of the moderator analysis, revealing that the effect size was related to the country where the studies were conducted. Therefore, it does not cross 0.

(c)

Figure 3. Random effect meta-analysis comparing crying amounts across countries at 1-2 weeks of age, 3-4 weeks of age, 5-6 weeks of age, 8-9 weeks of age, and 10-12 weeks of age. **A,** Random effect meta-analysis comparing crying amounts across countries at 1-2 weeks of age. **B,** Random effect meta-analysis comparing crying amounts across countries at 3-4 weeks of age. **C,** Random effect meta-analysis comparing crying amounts across countries at 5-6 weeks of age. **D,** Random effect meta-analysis comparing crying amounts across countries at 8-9 weeks of age. **E,** Random effect meta-analysis comparing crying amounts across countries at 10-12 weeks of age. (*Continues*)

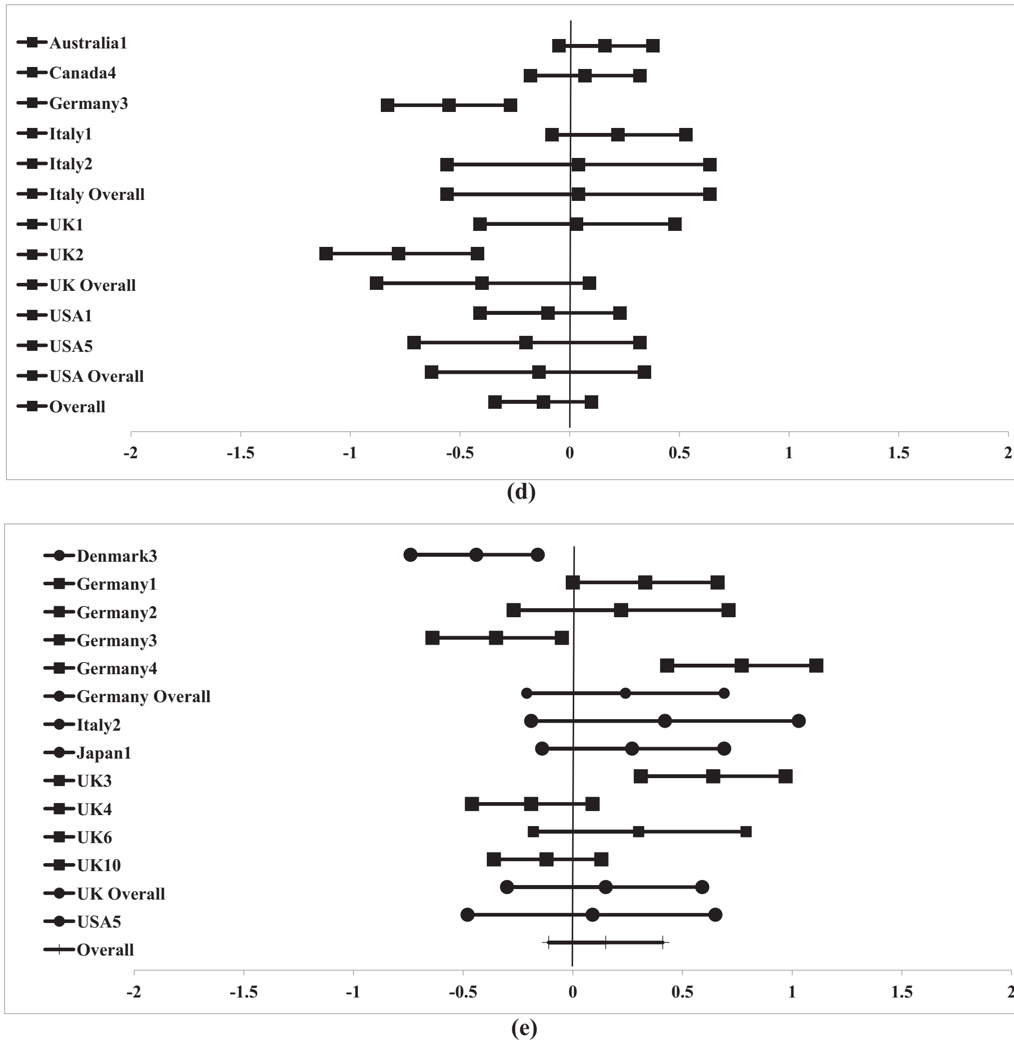


Figure 3. Continued.

Table I. Summary table of all samples included in the meta-analysis with quality ratings

	Sample codes	study		Wk	Sample size	Mean (SD) of cry/fuss duration in min	Recruitment rate	Participation rate at each age	Characteristic of sample	Subject selection (whole vs defined population)	Feeding type	Diary duration	Resolution time for the diary (5 vs 15 min)	Modified- Wessel definition	Quality rating scores
		Authors	Year												
1	UK1*	Darlington and Wright [60]	2006	8	24	105 (47.7)	Not reported	20%	Yes	Whole	Breastfed	2 d	15 min	No	3
2	UK2*	Darlington and Wright[60]	2006	8	51	58 (45.8)	Not reported	42%	Yes	Whole	Bottlefed	2 d	15 min	No	3
3	UK3	St. James-Roberts and Plewis[33]	1996	2	128	133 (77)	74%	64%	No	Whole	Not reported	3 d	5 min	No	4
				6	94	128 (70)		47%							3
				12	69	97(44)		35%							3
4	UK4	St. James-Roberts et al[75]	2001	1	191	107 (77)	35%	94%	Yes	Whole	Mixed	3 d	5 min	No	5
				3	181	122 (72)		89%							6
				6	173	102 (66)		85%							6
				12	152	60 (42)		75%							6
5	UK5	St. James-Roberts et al[20]	2006	1	111	120 (63)	81%	64%	Yes	Whole	Breastfed	4 d	5 min	Yes	8
				5	81	126 (58)		47%							7
6	UK6	St. James-Roberts and Menon-Johansson[76]	1999	1	14	170 (133.9)	Not reported	56%	No	Whole	Not reported	3 d	5 min	No	2
				6	20	129 (65.1)		80%							3
				12	20	82 (45.4)		80%							3
7	UK7	Lucas and St. James-Roberts[21]	1998	2	43	128 (79)	Not reported	96%	Yes	Whole	Breastfed	3 d	5 min	Yes	5
				6	36	149 (66)		80%							5
8	UK8	Lucas and St. James-Roberts[21]	1998	2	49	144 (98)	Not reported	94%	Yes	Whole	Bottlefed	3 d	5 min	Yes	5
				6	41	110 (51)		79%							5
9	UK9	St. James-Roberts et al[77]	2003	1	93	126.8 (75.3)	Not reported	68%	Yes	Selected	Mixed	3 d	Not reported	Yes	4
10	UK10	St. James-Roberts and Peachey[78]	2011	5-6	352	104.7 (63.2)	Not reported	Not reported	Yes	Whole	Mixed	3 d	5 min	No	3
				12	316	63 (42.1)									
11	Canada1	Barr et al[35]	1988	6	10	125 (48.1)	Not reported	Not reported	Yes	Defined	Mixed	7 d	5 min	No	4
12	Canada2	Barr et al[59]	1989	6	283	93 (61.1)	84%	69%	No	Whole	Breastfed	8 d	5 min	No	6
13	Canada3	Barr et al[59]	1989	6	91	88.7 (65.7)	84%	22%	No	Whole	Bottlefed	8 d	5 min	No	5
14	Canada4	Kramer et al[65]	2001	4	183	149.8 (73.5)	Not reported	71%	Yes	Defined	Breastfed	3 d	5 min	No	3
				6	156	131.6 (72.5)		61%							3
				9	148	107.6 (64.6)		57%							3
15	Canada5	Miller et al[66]	1993	5	88	136.8 (66)	Not reported	78%	No	Defined	Breastfed	7 d	5 min	Yes	6
16	Canada 6	Fujiwara et al[61]	2011	5	1065	163.4 (75)	78.6%	58.3%	Yes	Whole	Not reported	4 d	5 min	No	6
17	USA1	Blum et al[79]	2002	5	60	140.9 (75.1)	Not reported	53%	Yes	Defined	Not reported	4 d	5 min	No	3
				6	59	127.1 (69.9)		52%							3
				8	58	97.9 (46.9)		51%							3
18	USA2	Stifter and Spinrad[80]	2002	6	116	116.2 (58.2)	Not reported	74%	Yes	Whole	Not reported	4 d	5 min	No	6
19	USA3	Stifter et al[81]	2003	6	128	120.6 (64.1)	Not reported	89%	Yes	Whole	Not reported	4 d	5 min	Yes	7
20	USA4	Fujiwara et al[61]	2011	5	1857	152 (71.8)	54.2%	68.4%	Yes	Whole	Not reported	4 d	5 min	No	6
21	USA5	McRury and Zolotor[82]	2010	4	16	126 (72)	1408 fliers distributed, 51 responses to fliers	69%	Yes	Whole	Not reported	3 d	5 min	No	2
				6	16	114 (66)									
				8	17	90 (72)									
				12	14	72 (48)									
22	Italy1	Bonichini et al[83]	2008	2	70	147.6 (90.2)	Not reported	77%	Yes	Whole	Not reported	3 d	5 min	Yes	5
				5	70	150.6 (100.1)		77%							5
				8	70	118.5 (78.5)		77%							5
23	Italy2†	Mazzotti et al[84]	2003	2	12	106.8 (46)	Not reported	79%	No (2/4)	Whole	Not reported	3 d	5 min	Yes	4
				4	12	145.4 (77.1)									
				6	12	119.9 (53.9)									
				8	12	105.7 (88.8)									
				10	12	86.7 (34.5)									
24	The Netherlands1	de Weerth and Buitelaar[64]	2007	6	103	150.4 (66.3)	Not reported	89%	Yes	Whole	Mixed	4 d	5 min	No	7
25	Germany1	Keller et al[85]	1998	12	62	84.6 (54)	70%	82%	Yes	Whole	Not reported	3 d	5 min	No	5
26	Germany2	Lohaus et al[86]	2001	12	20	78.8 (59.9)	Not reported	100%	Yes	Defined	Not reported	3 d	5 min	No	3
27	Germany3	Bensel[72]	2003	2	96	69 (60)	Not reported	72%	Yes	Defined	Mixed	2-3 times per wk	15 min	No	2
				3	97	80.8 (67.4)		72%							2
				6	99	85 (78)		74%							2
				9	101	66.3 (69)		75%							2
				12	91	51 (51)		68%							2
28	Germany4	Wurmser et al[87]	2006	6	64	145.6 (84.4)	Not reported	68	Yes	Whole	Not reported	5 d	5 min	No	4
				12	63	107.9 (58.2)		67							
29	Australia1*	Wake et al[62]	2006	8	446	113.5 (64.1)	Not reported	92%	Yes	Whole	Mixed	1 d	5 min	No	6
30	Denmark1	St. James-Roberts et al[20]	2006	1	70	80 (44)	88%	80%	Yes	Whole	Breastfed	4 d	5 min	Yes	8
				5	64	81 (60)		79%							8
31	Denmark2	Sondergaard[88]	2000	6	432	88 (67)	Not reported	Not reported	Yes	Defined	Mixed	4 d	5 min	Yes	6
32	Denmark3	Alvarez[32]	2004	3	118	90 (58)	55%	79%	Yes	Defined	Mixed	3 d	5 min	Yes	7
				6	111	79 (67)		74%							6
				12	110	48 (44)		73%							6
33	Japan1	Shinohara and Kodama[89]	2012	4-6	31	107 (36)	50%	Not reported	Yes	Whole	Mixed	3 d	5 min	No	4
				8-10		80 (36)									

*Excluded from the colic prevalence analysis.
†Cross-sectional study. Please note that the following samples were reported in the same study: (1) UK1 and UK2; (2) UK5 and Denmark1; (3) UK7 and UK8; (4) Canada2 and Canada3; and (5) Canada6 and USA4.
References 75-89 available as **Appendix** (available at www.jpeds.com).