Development and testing of a multidimensional parent reported outcome measure for common presenting complaints of infancy: the UK infant questionnaire.

By Amy Sarah Miller BSc MSc,[1] Brechtje Huizinga BSc MSc,[1] Manu Pinkster BSc MSc,[1] Anna Clarissa Jeanne Telford BSc MSc,[1] Jorieke Maria ten Heggeler BSc MSc,[1] Joyce Elaine Miller BS, DC, PhD[2]

1 Anglo-European College of Chiropractic, Bournemouth, United Kingdom

2 Corresponding author: Joyce Elaine Miller BS, DC, PhD, Anglo-European College of Chiropractic, Bournemouth, United Kingdom
Email: jmiller@aecc.ac.uk

Conflict of Interest & Funding: The authors declare no conflict of interest. No funding was obtained for this study.

ABSTRACT

Objectives: Patient reported outcome measures are recognized as important and valuable tools to monitor patient progress in healthcare. It is fundamental to clinical practice to understand whether the treated patient has improved or not. Despite the highest use of outpatient healthcare among all pediatric age groups, no age-appropriate outcome measures are available for the infant. Therefore, the objective of this study was to develop and test a new infant outcomes instrument for the most common presenting complaints of infancy. **Methods:** This was a multi-phase study designed to develop a questionnaire using maternal interviews and to test it for reliability and validity for use in well child clinical practice. After collecting the mother's views, grounded theory and content analysis were used to derive themes and domains for the questionnaire. After achieving face validity, the instrument was evaluated for test-retest reliability, homogeneity and concurrent criterion validity. Subjects comprised a convenience sample of mothers who presented their infants to a university-affiliated chiropractic teaching clinic on the south coast of England. **Results:** Maternal interviews revealed mothers' concerns about feeding, sleeping, crying and other aspects of infant activities of daily living resulting in construction of a 12 question instrument. The questionnaire showed excellent test-retest reliability (ICC = 0.96) and good internal consistency (Cronbach's α = 0.8). In validity testing, ten questions showed positive correlation to a statistically significant degree against their established gold standard references. In all, 294 mother/infant dyads were involved in the research project. **Conclusion:** The UK Infant Questionnaire is the first parent reported outcome measure for use with the most common complaints of the infant patient based on maternal views. As such, this instrument meets the standard set by the UK National Health Service to involve the parent's voice in their child's care, and is therefore innovative in its field. Although further testing is indicated, and we make no claims that this instrument is comprehensive in all aspects of infant well-child care, it may be used by individual clinicians in routine daily practice to gain understanding of clinical progress of individual patients.

Keywords: Outcome measures, Pediatrics, Chiropractic

Introduction

The infant (0-1 year of age) incurs the highest health care costs of any age throughout childhood.[1] Common complaints are crying, feeding and sleeping problems, which are reported in 20-33% of infants,[2,3] and an additional 15% of infants are afflicted with two or more of these issues.[3] Problems in early infancy are associated with short term risks, including early discontinuation of breastfeeding,[4-6] infant abuse,[7,8] maternal depression,[8,9] and long term risks including developmental problems.[3] As birth injury has been largely implicated in these early infant complaints,[10] biomechanical factors may have some degree of influence on the short and long-term prognoses of these conditions.[11]

Chiropractic care is a modality commonly sought by families, and large numbers of infants are seen by chiropractors.[12-14] Despite high usage of chiropractic care for infants, the profession has been criticized for a lack of evidence upon which to support pediatric chiropractic care.[15] However, virtually all branches of health care have been guilty of too little evidence-based-practice for children.[16,17] A crucial example in medicine is the alarming and continued use of off-label prescriptions, in the pediatric population, whilst recognizing the lack of evidence for safety and efficacy required by regulatory standards.[16,17] A key reason for the scarcity of high quality research in pediatric care, across all health care arenas, is the paucity of relevant outcome measures in routine practice for this age group.[18] The development of age-appropriate outcome measurements is a pragmatic and appropriate next step considering the high

usage of health care in the first year of life.

Background

Patient reported outcome measures (PROMS) have been increasingly utilized by healthcare communities to measure patient responses to treatment. Their use in research is well known, as PROMS were initially developed to enable a clinician to measure outcomes in clinical trials and to take into account a patient's subjective health status and quality of life.[18] The routine use of outcome measures is increasingly called for in all types of health care as a valuable measure of change, be it improvement or worsening, in order to document whether or not a specific treatment has value to the individual patient. Collecting outcome data from large numbers of patients in chiropractic practice has been proven feasible.[19]

Including children in the endeavor to use patient choices to improve their care is crucial.[20,21] Pragmatic and efficient outcome instruments for the infant patient would be particularly useful because other methods to study responses to care are onerously time and cost consuming. For example, large scale direct observation of the infant with video-recording or in-house recordings would be invasive and problematic for both the family and the researcher. Diaries have been used as the gold standard (validated against in-house recordings) for the infant's behaviour,[22,23] but these are time consuming for both parents and researchers and are therefore, under-utilized, ignored or abandoned. Questionnaires may be a practical replacement for diaries as an efficient tool to measure outcomes, provided they can be documented as equally credible. Mothers have been shown to be reliable reporters of their infant's behavior,[22-24] and therefore intake and discharge questionnaires could be a quick, accurate and pragmatic way to study this population's behaviors and treatment results.

A small number of questionnaires have been developed for individual aspects of the infant's problematic behavior, including sleeping,[24-26] pain,[27,28] crying,[29] and feeding.[30] Despite these focused attempts, there is no established validated pediatric instrument that covers all key aspects of an infant's behavior.[24-31] Many individual problems faced by the infant population (feeding, crying, sleeping, postural problems/pain) are overlapping and interlinked,[2] and as such should be viewed as a piece of the clinical puzzle, rather than the whole picture.

Because of the requirement to represent the exact needs of the patient,[18] with infant patients the maternal voice must be heard by the researcher. Qualitative approaches used to ascertain what is most important in infant health care from the mother's perspective can be considered the key foundation for any new outcome measure.[32,33] This not only ensures content validity, but focuses on and respects the voices of those for whom the outcome measurement is intended, which in this case is the mother of the infant.

Therefore, the goals of this project were to ask mothers what key concepts were most important in their infant's health, use these domains to develop a questionnaire, investigate the reliability and validity of the questions in the instrument and test it for intake and follow-up to understand parent report of outcomes of infant care.

Methods

The development stages and exploration of instrument testing are shown in Table 1 (next page) which was generated to summarize the methodological procedures to be viewed at a glance.

Subjects and setting: A convenience sample of mothers who presented their infant to a chiropractic teaching clinic was recruited. Inclusion criteria were English fluency and consent to be part of the study. The treating clinician was not aware as to whether the mother had enrolled in the study or not.

Ethical approval was granted by the AECC Research Ethics Subcommittee in July 2014. There was no funding for this project and no financial incentive for either the subjects or the investigators.

Phase 1: Qualitative Study: Mothers attending the clinic with their infant were asked to take part in an interview. Interviews were based on a discussion guide,[32] broadly based on common presenting complaints in the infant age group and outcomes and experiences of healthcare. This document was evolving and was changed as new topics arose in interviews, a principle used in reflexivity.[34,35] Interviews were anonymously audio-recorded and transcribed verbatim.

An exceptionally large sample size in the qualitative phase was determined for the purpose of obtaining as many opinions from the mothers as was reasonably possible. Data saturation is generally considered complete after 12 interviews.[36] Brod et al. suggested that after 12 interviews, between 88% and 92% of analysis codes (themes or domains) could be identified.[36] However in this study it was decided to continue qualitative research until no new themes had emerged for several interviews in order to be assured of saturation. This was important to ensure high content validity of the domains for use in the questionnaire.

Concepts from content analysis and grounded theory were used to extract data from transcripts, based on work by Lasch et al. and Brédart et al.[32,33] Themes were the topics

which featured most heavily throughout each transcript, and were decided after each researcher individually read each transcript, and discussed to reach agreement by at least three researchers.

Phase 2: Development of the questionnaire: Following data analysis, each team member wrote questions around the themes and domains from the interviews. Questions were subsequently reworded to utilize the mothers' language in order to be conscious of the subjects' needs and parlance.[32]

Phase 3: Face validity: Because this was a new type of instrument, never before tested, it was considered that face validity should be established to determine whether the questionnaire has merit or "face value." Experts in the fields of either pediatric chiropractic or research were asked for feedback on all aspects of the questionnaire including topics covered, anchors for the answers to each question, and the wording and layout. An 11-point scale was used with descriptive anchors at zero, five and ten, which varied depending on the question. Feedback was recorded and discussed with the team before implementing changes.

As the central focus of this project, mothers were given the final consideration and asked both to complete and comment on two versions of the questionnaire. A convenience sample of mothers who presented their infant to the clinic for care were observed during the completion for any hesitation and the time taken to complete. They were asked what each question meant to them and if they had any issues with filling out the questionnaire and if any key points were missing. The questionnaires were then re-formatted and the final version of the United Kingdom Infant Questionnaire (UKIQ) was generated ready for reliability and validity testing.

Phase 4: Reliability: Test-retest reliability: First, test-retest reliability was measured by asking mothers to complete the same questionnaire twice on the same day. The items on the second questionnaire were given in a different order to reduce the possibility of the participant memorizing their initial response,[38,39] although it was considered that this risk

Table 1. Summary of Methods to Develop and Test an Infant Outcomes Instrument

Inclusion criteria: Mothers of infants <12 months, fluent in English

Exclusion criteria: Mothers of infants >12 months, not fluent in English

Subjects and setting: Convenience sample of mothers presenting to the AECC outpatient clinic Ethics Research proposal approved by AECC Research Ethics Subcommittee

• Phase 1: Qualitative Study

Interviews and focus groups were conducted using a discussion guide. Topics included experiences of care and aspects of healthcare important to the mother. Data analysis: Grounded theory content analysis was used to create themes.

• Phase 2: Development of questionnaire

Themes from interviews were used to create questions corroborated with research into infant public health issues. Questions were generated to reflect maternal concerns and worded to reflect maternal language.

• Phase 3: Face validity

Experts

Experts in either pediatric care or research were asked to give feedback on the questionnaire during interviews or focus groups.

Mothers

Two questionnaires (versions one and two) with the same layout but different wording were completed by mothers. The time needed to complete the questionnaire, signs of hesitation, erasures or skipped questions were observed. Feedback from mothers was recorded and integrated.

• Phase 4: Reliability

Test-retest reliability

The questionnaire was given to mothers twice on the same day to determine test-retest reliability. The questions on the retest questionnaire were reorganized to avoid memory aids.

Homogeneity (internal consistency)

Internal consistency was calculated over each of the administrations of three questionnaires (including pre- and post-treatment) using Cronbach's α and item corrected total correlations.

• Phase 5: Validation of the questionnaire

Responsiveness

Mothers completed a questionnaire pre- and post-treatment.

Six-day behaviour diary

Mothers completed a 24-hour behaviour diary (gold standard) for six consecutive days. After six days the diaries were returned to the researcher and mothers completed the UKIQ and its correlated validation questionnaire.

Validation

Each question of the UKIQ was correlated to a question out of a reference questionnaire or the six day behaviour diary using Spearman's ρ, Phi Coefficient, Cramer's Phi or Cohen's Kappa depending on data requirements.

• Phase 6: Pilot testing

Questionnaire was implemented in a busy infant practice to test the practicalities of usage from clinician, parent and office staff perspectives.

was low with fatigued and stressed new mothers completing the forms. The test-retest reliability was calculated using intra-class correlation (ICC) coefficient two-way mixed single measures (ICC3.1).[39-42] It assessed the reliability of ratings by comparing the variability of different ratings of the same subject to the total variation across all ratings and all subjects.

Homogeneity: Second, homogeneity (internal consistency) was assessed. Homogeneity measures whether all of the items in the questionnaire are tapping different aspects of the same attribute. If this is the case, the items in the questionnaire can be added to give a total score.[38-42] Cronbach's α statistic[40-44] and item-corrected total correlations (Pearson's correlation coefficient, *r*)[40] were used. Cronbach's α provided a measure of the internal consistency of the questionnaire in its totality. Cronbach's α uses inter-item correlations to determine whether constituent items are measuring the same domain;[44,45] it compares the variance of each question with the variance of the total score. Item corrected-total correlation is calculated on a per-question basis and shows whether individual items are correlated to the total score of the questionnaire.

Phase 5: Validation of the questionnaire: Although many widely used instruments in health services have never been tested for validity,[42] it was decided that reliability alone was insufficient to support the use of this instrument. Validity was tested to determine the degree to which the instrument measured the domains it purported to measure, by testing each question against its own validated measure. Validity is measured by degrees and is not a binary judgement. Validity testing was carried out in stages to test face, construct and criterion concurrent validity.

Validity: Each domain of the questionnaire was matched to an external measure that was the gold standard for that dimension (Table 2, next page). This criterion-related approach seeks the amount of correlation with another test designed to measure the same thing. Because the UKIQ was the first questionnaire developed directly from the current views of mothers, some of the topics were novel and had not previously been investigated, only a reference standard of a six day 24-hour behaviour diary could be used, as diaries have been validated as accurate records of infant behavior.[22-24] The gold standard questions were then compiled to form a reference questionnaire. This was required because it was considered unethical to ask mothers to complete eight questionnaires along with a diary for comparison at one time.

Mothers were given the UKIQ, the six-day diary and the reference standard questionnaire at indicated times during the infant's intake, treatment and follow-up. Data from each survey were entered into SPSS and statistical tests were chosen relative to the type of data tested.

Analysis: Because of the nature of the data, non-parametric statistics were predominantly used. Each question was tested more than once, if more than one appropriate test could be used. Pearson's *r* tested scale based data. The phi coefficient was used for nominal based questions (yes/no). Where questions had more than two categories, Cramer's Phi coefficient was used, with cut-off points of 0.3 (medium) and 0.5 (large) effect as standards. The kappa measure of agreement was used to determine agreement between the two instruments (the UKIQ against the reference standard). A value of 0.5 for kappa represents moderate agreement, 0.7 good and 0.8 excellent.[42-46] All were tested for statistically significant associations.

A patient global impression of change (GIC) question was included in the follow up questionnaire and this was used as the gold standard to assess clinically significant change over time,[47-49] or responsiveness. The questionnaire was completed by mothers pre- and post-treatment and compared to the GIC. The correlation was calculated between the changed score for each question, as well as the corrected changed total score of the questionnaire overall.

Results

In all, 294 mother/infant dyads were recruited into the study with the baby's mean age of 8 weeks and the mother's mean age of 31. The infants were presented for crying (21%), feeding problems (20%), inability to sleep supine (19%), other sleeping problems (16%), head shape (8%) or check-up/difficult birth (16%).

Qualitative phase: The qualitative phase of the project gave domains which were infant behaviors (feeding, crying, sleeping, pain, movement patterns and abilities) and maternal feelings (anxiety, depression and quality of life). These were translated into questions that could be scored relative to the degree of the problem.

Phase 2: Development of the questionnaire: This resulted in a 12 question intake and follow-up questionnaire with the same questions. The follow up questionnaire also included a global impression of change (GIC) question used as the gold standard to assess clinically significant change.[48,49]

Phase 3: Face and content validity: A total of eight participants were included, six experts in pediatric care and two experts in research, who agreed relevance, merit, content and face value of the instrument. Seventeen mothers found it relevant and useful.

The UK Infant Questionnaire was then presented to 20

Table 2. UKIQ Questions and Gold Standard Reference Measures

No.	UKIQ question	Gold standard/ surrogate measure	Why the gold standard/ surrogate measure was used
1.	Over the past few days, on average, have you considered your baby's feeding to be a problem?	Infant behavior diary	- Feeding easily correlated with 24 hour behaviour diary (Barr et al. 1988) - Eating Behavioural Questionnaire only validated for children (CEBQ), not for infants (BEBQ)
2.	Over the past few days, on average, have you considered your baby's sleeping to be a problem?	Brief Infant Sleeping Questionnaire (BISQ): "Do you consider your child's sleep as a problem?"	Research findings provide psychometric, clinical, and ecologic support for the use of the BISQ as a brief infant sleep measure for clinical and research purposes (Sadeh 2004)
3.	Over the past few days, on average, have you considered your baby's crying to be a problem?	Infant behavior diary	Infant behaviour diary is the gold standard and is easily correlated with the 24 hour behaviour diary (Barr et al. 1988)
4.	Over the past few days, on average, how much did your baby cry?	Infant behavior diary	- Infant behaviour diary is the gold standard for crying and an easy tool to measure crying quantities (Barr et al. 1988) - Only one other validated instrument: QUIC, which provides valid information about infant crying, but has not been widely disseminated and utilized (Miller and Green 2011)
5.	Over the past few days, on average, how easy or difficult has it been to console (comfort, calm) your baby when he/she cried?	FLACC (Face, Legs, Activity, Cry, Consolability) behavioral pain assessment scale	- Validated scale to assess consolability amongst four other pain measures - Validated as a parent-assigned pain score, which has been used as proxy measures of pain for children, such as those with cognitive impairment (CI), who cannot self-report (Voepel-Lewis et al. 2005)
6.	Over the past few days, on average, how comfortable (settled, relaxed) has your baby been while lying on his/her back?	None	There are no validated instruments
7.	Over the past few days, on average, how would you rate your baby's discomfort or pain?	FLACC (Face, Legs, Activity, Cry, Consolability) behavioral pain assessment scale	- Validated for children between 2 months and 7 years of age or individuals unable to communicate their pain. - Assesses for pain in children - Validated as a parent-assigned pain score, which has been used as proxy measures of pain for children, such as those with cognitive impairment (CI), who cannot self-report (Voepel-Lewis et al. 2005)
8.	Over the past few days, on average, how anxious (worried) or distressed (upset) have you been feeling about your baby's behavior?	Edinburgh Post Natal Depression Scale (EPDS): "I have been anxious or worried for no good reason"	The EPDS is the most widely used well-validated 10-item questionnaire designed to screen post-natal depression in mothers (Cox et al. 1987) - It has been validated in thirty-seven different research papers up to 2008 (Gibson et al. 2009)
9.	Over the past few days, on average, how depressed (feeling down, sad) have you been feeling?	Edinburgh Post Natal Depression Scale (EPDS): "I have felt sad or miserable"	The EPDS is the most widely used well-validated 10-item questionnaire designed to screen post-natal depression in mothers (Cox et al. 1987) - It has been validated in thirty-seven different research papers up to 2008 (Gibson et al. 2009)
10.	Overall, how would you rate your experience of motherhood and quality of life with this baby?	The short version of the World Health Organization Quality of Life questionnaire (WHOQOL-BREF): "How would you rate your quality of life?"	-Validated for postnatal mothers (Webster et al. 2010) - The WHOQOL-BREF is well-accepted and valid instrument for new mothers and can be utilized in postnatal clinical settings or for research purposes
11.	Over the past few days has your baby turned his/her head freely to both sides?	None	There are no validated instruments
12.	Over the past few days, on average, how much tummy time has your baby had?	Infant behavior diary	There are no validated instruments
13.	Since the beginning of treatment at this clinic, how would you describe the change (if any) in your baby's condition and or behavior?	Patient Global Impression of Change (GIC) scale	The GIC is the gold standard to measure clinically significant change in adults (Hurst and Bolton 2004)

mothers to complete. The average time to complete the questionnaire was five minutes. All participants were content with the length, wording, order, layout and content of the questionnaire and no changes were made at this stage.

Phase 4: Reliability: The results of the Intra-Class Correlation (ICC) coefficient indicated excellent test-retest reliability for each individual question. The overall ICC value of 0.96 (n=29) indicated excellent test-retest reliability of the entirety of the UK Infant Questionnaire.

Homogeneity (internal consistency): Cronbach's α was higher than 0.8, showing good internal consistency and that the instrument taps on different aspects of the same attribute.

Test-retest reliability: After test-retest and homogeneity testing, the UKIQ was considered reliable. The corrected item total correlations test showed that questions 1 — 10 can be added to form a total score, but items 11 and 12 did not contribute to the overall score. It was decided *a priorii* that two questions would not be added into a total score. Question 11 (cervical spine rotation) was an untested question because of the absence of a gold standard comparison. However, it can be considered to have face validity based on maternal and expert opinion. Question 12 asks for specific times of prone play, rather than a rating scale, and therefore, is dissimilar in formatting to the first ten questions. Both are clinically important, but do not contribute to a total score.

Phase 5: Validity: Of the 12 questions tested, 10 were confirmed as valid to a statistically significant degree (Table 3). (Item 1 (feeding) correlated with the diary (Spearman's ρ .729). Question 2 (sleep) correlated with the Brief Infant Sleep Questionnaire (BISQ), showing good agreement. Items 3 and 4 (crying) correlated with the diary (Spearman's ρ .568 -.612). Items 5 (crying), and 6 (pain) measured good agreement with the Face Legs, Activity, Cry, Consolability (FLACC) pain questionnaire. Question 9 (depression) indicated good agreement with the Edinburg Post-Natal Depression Scale (EPDS) at .737 (Cohen's Kappa) and Cramer's Phi (.780). Likewise, question 10 (maternal quality of life) correlated at a good level of agreement with the World Health Organization Quality of Life (WHOQOL) questionnaire (Spearman's ρ .667 and Cramer's Phi .514). Question 12 (prone play time) correlated with the diary (Spearman's ρ .688). Item 13 (UKIQ GIC) on the follow-up questionnaire showed high agreement with the existing validated Patient Global Impression of Change with Spearman's ρ .905.

Two items (question 7 — maternal anxiety and question 11 — cervical spine rotation) cannot be considered validated, as question 7 had only divergent validity (discrimination) with its reference standard, and question 11 has never been previously tested and therefore has no gold standard reference. Question 7 was removed from the UKIQ. Question 11 was considered important by both mothers and experts and therefore was retained, although it cannot be added to a total score. Item 12 cannot be summed either, as it is reverse scored from the rest of the questions (this was done to improve understanding by the mothers and to obtain clinically important information). The first ten items can be used as a sum of scores and each item can be used individually. The scores showed no apparent floor or ceiling effect.

Discussion

The 12-item UKIQ was developed through a systematic process of literature reviews, qualitative research, expert review, and pilot testing using 294 mother-infant dyads. Pa-

Table 3. Results of validity testing

Item	External measure used for correlation	Spearman's R	Cohen's Kappa	Phi (2x2)	Cramer's Phi
1	Infant behavioural diary	.729*			
2	Brief Infant Sleep Questionnaire (BISQ)	.641*	.476*		.470*
3	Infant behavioural diary	.612*			
4	Infant behavioural diary	.568*			
5	FLACC^			.678*	
6	FLACC^	.541*			
7	FLACC^	.688*			
8	EPDS& (anxiety)			.388(0.07)	
9	EPDS& (depression)		.737*		.780*
10	WHOQOL-BREF+	.667*	.410*		.514*
11	Cervical spine (validity not tested)				
12	Infant behavioural diary	.688*			
13	Patient Global Impression of Change	.905*			

** Correlation is significant at the 0.01 level or better; ^Face, Legs, Activity, Cry and Consolability behavioural pain assessment scale; &Edinburgh Postnatal Depression Scale; +World Health Organization Quality of Life Questionnaire (abbreviated)*

tient Reported Outcomes Measures (PROMs) for children are rare, and those that do exist are of low quality and lack reliability, validity or both or only measure a single variable.[25]

Because of maternal concern for their infant's problems, along with the risk for persistence of those problems when inadequately addressed, the purpose of this project was to develop and investigate a questionnaire to follow progress during infant clinical practice. At first, the Bournemouth Questionnaire (BQ) was used as a model. However, this was abandoned to set the questions because of the need to use maternal interviews[50] (qualitative research) as a starting point, whereas the BQ had used a literature search as a basis to determine domains. However, their process for testing the questions was retained. Content validity was addressed by the implementation of current research[50] which advises including the parent and child into health care decisions. Thus qualitative research was used as the foundation for the instrument developed. The goal was to establish content validity that best served the target patient population by implementing their own needs and views. As such, the UKIQ is the only known pediatric questionnaire based predominantly on the views of the subjects it addresses, and this may be considered its strength.

This is balanced by significant weaknesses. Qualitative research has been criticized for lack of rigour.[51] Further, it can lack generalizability. That is a distinct problem with this sample, as all subjects had presented to an outpatient chiropractic teaching clinic, and therefore were not necessarily representative of the general population of infants in the UK. However, the pediatric population in this clinic has previously been found to be representative of the broader infant population.[52] Further, validating each question against its own gold standard reference is unique as was reliance on the infant behavioral diary. Although the behavior diary has been tested and is a gold standard, it has been noted by many researchers[22,53] that the diary is onerous to complete and causes high drop-out rates. This is balanced by the significant rigorous testing of the infant diary and it remains the only gold standard in infant behavior today.[22-24]

Despite its drawbacks and need for further testing, relying on maternal views has led to a broad spectrum questionnaire which can be used clinically to understand the individual infant's health status and response to care in a well-child practice. It can provide a starting point, testing for clinical utility and pragmatic use in daily practice. Chiropractic care has been highly criticized for paucity of research in the infant patient.[15,54] A reliable, valid and easy to use questionnaire as a replacement for the onerous gold standard behavior diary may go some way to encouraging more research due to the ease of measuring outcomes. The aim was to allow both clinicians and researchers to integrate key aspects of the infant health parameters into their studies and practices. As such, the UKIQ is simply a first step to open the wider discussion between clinicians and researchers to continue development toward a widely useful instrument.

Meanwhile, this outcome measure may be useful to further the evidence base for infant care by chiropractors, which is urgently needed. It may also be useful for individual clinicians providing infant care to track progress with patients in their own clinic. The routine use of PROMS in clinical practice is widely advocated as a means of supporting patient-centered care, informing decisions and driving service quality.[19,55]

A unique aspect of the UKIQ is that it can be used to assess the infant's progress relative to current public health issues of supine sleep, breastfeeding, and positional head deformation. As such, it tests the specific and broad concerns of the mothers for their child in well-patient care. Establishing efficacious treatments for infants facing difficulties in these areas is extremely important, not only for the health of the individual, but for their families and the broader community including the already economically stretched health services.

Conclusion

The UKIQ is a new parent-reported outcome measure for use with the infant patient in well-child practice. It was founded in maternal views through qualitative research, and demonstrated face validity, reliability, and validity as a clinically useful tool. As such, it provides a starting point for more investigation and discussion for further development. The UKIQ may be used in clinical practice by individual clinicians to monitor progress of infant patients, and could also be used in future outcomes focused research for this age group, as well as community surveys to sample maternal concerns. The practical and clinical utility with widely varied populations along with clinical significance of scores requires further study.

References

1. Alemayehu B, Warner KE: The lifetime distribution of health care costs. *Health Serv Res* 2004, 39:627-642.

2. 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, 96:622-629.

3. Wolke D, Rizzo P, Woods S: Persistent infant crying and hyperactivity problems in middle childhood. *Pediatrics* 2002, 109:1054-1060.

4. Hall RT, Mercer AM, Teasley SL, McPherson DM, Simon SD, Santos SR, Meyers BM, Hipsh NE: A breast-feeding assessment score to

evaluate the risk for cessation of breast-feeding by 7 to 10 days of age. *J Pediatr* 2002, 141:659-664.

5. Tamagawa K, Weaver J: Analysing adverse effects of epidural analgesia in labour. *British Journal of Midwifery* 2012, 20:704-708.

6. Guyer J, Millward LJ, Berger I: Mothers' breastfeeding experiences and implications for professionals. *British Journal of Midwifery* 2012, 20:724-733.

7. Reijneveld SA, van der Wal MF, Brugman E, Sing RA, Verloove-Vanhorick SP: Infant crying and abuse. *Lancet* 2004, 364:1340-1342.

8. Almond P: Postnatal depression: a global public health perspective. *Perspect Public Health* 2009, 129:221-227.

9. Borra C, Iacovou M, Sevilla A: New Evidence on Breastfeeding and Postpartum Depression: The Importance of Understanding Women's Intentions. In *Matern Child Health J.* Volume 19. Boston; 2015: 897-907

10. Zwart P, Vellema-Goud MG, Brand PL: Characteristics of infants admitted to hospital for persistent colic, and comparison with healthy infants. *Acta Paediatr* 2007, 96:401-405.

11. Schertz M, Zuk L, Zin S, Nadam L, Schwartz D, Bienkowski RS: Motor and cognitive development at one-year follow-up in infants with torticollis. *Early Hum Dev* 2008, 84:9-14.

12. Hestbaek L, Jorgensen A, Hartvigsen J: A description of children and adolescents in Danish chiropractic practice: results from a nationwide survey. *J Manipulative Physiol Ther* 2009, 32:607-615.

13. Miller J: Demographic survey of pediatric patients presenting to a chiropractic teaching clinic. *Chiropr Osteopat* 2010, 18:33.

14. Barnes PM, Bloom B, Nahin RL: Complementary and alternative medicine use among adults and children: United States, 2007. Natl Health Stat Report 2008:1-23.15. Ernst E: Chiropractic spinal manipulation for infant colic: a systematic review of randomised clinical trials. *Int J Clin Pract* 2009, 63:1351-1353.

16. Palmaro A, Bissuel R, Renaud N, Durrieu G, Escourrou B, Oustric S, Montastruc JL, Lapeyre-Mestre M: Off-label prescribing in pediatric outpatients. *Pediatrics* 2015, 135:49-58.

17. Frattarelli DA, Galinkin JL, Green TP, Johnson TD, Neville KA, Paul IM, Van Den Anker JN: Off-label use of drugs in children. *Pediatrics* 2014, 133:563-567.

18. Hunter C, Fitzpatrick R, Jenkinson C, Darlington AS, Coulter A, Forder JE, Peters M: Perspectives from health, social care and policy stakeholders on the value of a single self-report outcome measure across long-term conditions: a qualitative study. *BMJ Open* 2015, 5:e006986.

19. Newell D, Diment E and Bolton JE. An electronic patient-reported outcome measures system in the UK chiropractic practices: A feasibility study of routein collection of outcomes and costs. *Journal of Manipulative and Physiolgical Therapeutis* 2016;39(1):31-41.

20. CentreHSCIC: Hospital Episode Statistics NHS Maternity Statistics — England, 2013-14 Health and Social Care Information Centre; 2015.

21. Children and young people's survey 2014 www.cqc.org.uk/content/children-and-young-peoples-survey-2014

22. Barr RG, Kramer MS, Boisjoly C, McVey-White L, Pless IB: Parental diary of infant cry and fuss behaviour. *Arch Dis Child* 1988, 63:380-387.

23. St James-Roberts I, Hurry J, Bowyer J: Objective confirmation of crying durations in infants referred for excessive crying. *Arch Dis Child* 1993, 68:82-84.

24. Müller S, Hemmi MH, Wilhelm FH, Barr RG, Schneider S: Parental report of infant sleep behavior by electronic versus paper-and-pencil diaries, and their relationship to actigraphic sleep measurement. *Journal of Sleep Research* 2011, 20:598.

25. Sadeh A: A Brief Screening Questionnaire for Infant Sleep Problems: Validation and Findings for an Internet Sample. *Pediatrics* 2004, 113:e570-e577.

26. Spruyt K, Gozal D: Pediatric sleep questionnaires as diagnostic or epidemiological tools: a review of currently available instruments. *Sleep Med Rev* 2011, 15:19-32.

27. RCN: The recognition and assessment of acute pain in children. In *Clinical practice guidelines*. pp. 73. London: Royal College of Nursing; 2009:73.

28. Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S: The FLACC: a behavioral scale for scoring postoperative pain in young children. *Pediatr Nurs* 1997, 23:293-297.

29. Miller JE, Green A: QUIC: Initial Validation of an Instrument to Measure Infant Crying. *Journal of Clinical Chiropractic Pediatrics* 2011, 12:843-848.

30. Llewellyn CH, van Jaarsveld CH, Johnson L, Carnell S, Wardle J: Development and factor structure of the Baby Eating Behaviour Questionnaire in the Gemini birth cohort. *Appetite* 2011, 57:388-396.

31. Morrell JMB: The Infant Sleep Questionnaire: A New Tool to Assess Infant Sleep Problems for Clinical and Research Purposes. *Child Psychology and Psychiatry* 1999, 4:20-26.

32. Bredart A, Marrel A, Abetz-Webb L, Lasch K, Acquadro C: Interviewing to develop Patient-Reported Outcome (PRO) measures for clinical research: eliciting patients' experience. *Health Qual Life Outcomes* 2014, 12:15.

33. Lasch KE, Marquis P, Vigneux M, Abetz L, Arnould B, Bayliss M, Crawford B, Rosa K: PRO development: rigorous qualitative research as the crucial foundation. *Qual Life Res* 2010, 19:1087-1096.

34. Salmond SW: Qualitative Metasynthesis. In *Comprehensive Systematic Review for Advanced Nursing Practice* New York, USA: Springer; 2011: 209-236

35. McNair R, Taft A, Hegarty K: Using reflexivity to enhance in-depth interviewing skills for the clinician researcher. *BMC Med Res Methodol* 2008, 8:73.

36. Rabiee F: Focus-group interview and data analysis. *Proc Nutr Soc* 2004, 63:655-660.

37. Brod M, Tesler LE, Christensen TL: Qualitative research and content validity: developing best practices based on science and experience. *Qual Life Res* 2009, 18:1263-1278.

38. Bolton JE, Breen AC: The Bournemouth Questionnaire: a short-form comprehensive outcome measure. I. Psychometric properties in back pain patients. *J Manipulative Physiol Ther* 1999, 22:503-510.

39. Bolton JE, Humphreys BK: The Bournemouth Questionnaire: a short-form comprehensive outcome measure. II. Psychometric properties in neck pain patients. *J Manipulative Physiol Ther* 2002, 25:141-148.

40. Landers R: Computing Intraclass Correlations (ICC) as Estimates of Interrater Reliability in SPSS. *The Winnower* 2015.

41. Shrout PE, Fleiss JL: Intraclass Correlations: Uses in Assessing Rater Reliability. *Psychological Bulletin* 1979, 86:420-428.

42. Streiner DL, Norman GR, Cairney J: *Health measurement scales: a practical guide to their development and use*. Oxford: Oxford University Press; 2015.

43. Tao X, Chow SK, Wong FK: Determining the validity and reliability of the Chinese version of the Kidney Disease Quality of Life Questionnaire (KDQOL-36). *BMC Nephrol* 2014, 15:115.

44. Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD: What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health* 2007, 10 Suppl 2:S94-s105.

45. Rattray J, Jones MC: Essential elements of questionnaire design and development. *J Clin Nurs* 2007, 16:234-243.

46. Tavakol M, Dennick R: Making sense of Cronbach's alpha. *Int J Med Educ* 2011, 2:53-55.

47. Spruyt K, Gozal D: Development of pediatric sleep questionnaires as diagnostic or epidemiological tools: A brief review of Dos and Don'ts. *Sleep Medicine Reviews* 2011, 15:7.

48. Farrar JT, Young JP, Jr., LaMoreaux L, Werth JL, Poole RM: Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain* 2001, 94:149-158.

49. Hurst H, Bolton J: Assessing the clinical significance of change scores recorded on subjective outcome measures. *J Manipulative Physiol Ther* 2004, 27:26-35.

50. Weil LG, Lemer C, Webb E, Hargreaves DS: The voices of children and young people in health: where are we now? *Arch Dis Child* 2015, 100:915-917.

51. Mays N, Pope C: Rigour and qualitative research. *Bmj* 1995, 311:109-112.

52. Miller J, Newell D, Bolton J. Efficacy of chiropractic manual therapy in infant colic: a pragmatic single-blind, randomised controlled trial. *Journal of Manipulative and Physiological Therapeutics* 2012; 35(8):600-7.

53. Lam J, Barr RG, Catherine N, Tsui H, Hahnhaussen CL, Pauwels J, Brant R: Electronic and paper diary recording of infant and caregiver behaviors. *J Dev Behav Pediatr* 2010, 31:685-693.

54. Bronfort G, Haas M, Evans R, Leininger B, Triano J: Effectiveness of manual therapies: the UK evidence report. *Chiropr Osteopat* 2010, 18.

55. Bodger K, Ormerod C, Shackcloth D, Harrison M, Ahmed S, Bowering K, Brown E, Hurst A, Gledhill T: Development and validation of a rapid, generic measure of disease control from the patient's perspective: the IBD-Control questionnaire. *Gut* 2014, 63:1092-1102.