Patient Related Outcome Measures

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Clinically Meaningful Difference for the Infant Gastroesophageal Questionnaire Revised version (I-GERQ-R): A Quantitative Synthesis

This article was published in the following Dove Press journal:
Patient Related Outcome Measures

Adam B Smith[1]
Neil Fawkes[2]
Helen Kotze[3]
Victoria Hodgkinson[4]
Cathal Coyle[5]

[1]Health Outcomes, Medical Affairs & Evidence Generation, Reckitt Benckiser, Hull, UK; [2]Medical Excellence, Medical Affairs & Evidence Generation, Reckitt Benckiser, Hull, UK; [3]Global Medical Affairs, Medical Affairs & Evidence Generation, Reckitt Benckiser, Hull, UK; [4]Clinical Operations, Medical Affairs & Evidence Generation, Reckitt Benckiser, Hull, UK; [5]Global Medical Affairs, Medical Affairs & Evidence Generation, Reckitt Benckiser, Slough, UK

Background: Gastroesophageal reflux disease (GORD) is a common condition affecting 30% of infants aged 0–23 months. The Infant Gastroesophageal Questionnaire Revised version (I-GERQ-R) is an observer-reported outcome measures (ObsRO) developed to evaluate the impact of GORD on young infants. However, evidence regarding the clinically important difference (CID) for the I-GERQ-R is limited. The aim of this study was to determine a CID for the I-GERQ-R.

Methods: A literature review was undertaken (PsycInfo, Embase, MedLine and EconLit databases) for longitudinal studies involving the I-GERQ-R. Articles were not limited by language or publication date. A random effects model was applied to calculate an overall CID, along with I^2 and Q statistics. Publication bias was also assessed.

Results: The search identified 42 articles; 11 were selected for full-text review and 7 articles were identified for full data extraction. The studies included a total of 661 infants (range: 30 to 313); 424 infants had been diagnosed with GORD (64%). The age range of the infants across the studies was from birth to 7 months. The overall CID was −6.54 (95% confidence interval: −4.35 to −8.74), Q = 17.96, p=0.08 and I^2=22.04.

Conclusion: This study derived a CID for the I-GERQ-R and indicated a threshold around 6 could signify a clinically important difference for this instrument. The lower limit of the 95% confidence interval suggested a threshold of 3 to 4 could represent a minimally important difference. These results may help inform clinical decisions in evaluating meaningful change in symptom severity in children affected by GORD.

Keywords: I-GERQ-R, gastro-esophageal reflux disease, minimally important difference, clinically important difference

Background

Paediatric gastro-oesophageal reflux (GOR) is common in infants, particularly in pre-term babies, with an estimated incidence of 22% in babies born before 34 weeks of gestation.[1] GOR refers to the passage of gastric contents into the oesophagus or oropharynx with or without vomiting and can be a daily, normal physiological occurrence in infants with few or no troublesome symptoms. Regurgitation is the most common and visible sign/symptom that parent/caregivers can observe and can occur daily in about 50% of infants <3 months of age[1] and 30% in infants aged <6 months.[2] Gastro-oesophageal reflux disease (GORD) refers to the presence of troublesome symptoms arising from the passage of gastric contents into the oesophagus.[3,4]

Correspondence: Adam B Smith
Reckitt Benckiser, Dansom Lane, Hull HU8 7DS, UK
Tel +44 1482 583528
Fax +44 1482 582172
Email adam.smith@rb.com

submit your manuscript | www.dovepress.com
DovePress
http://doi.org/10.2147/PROM.S238673

There is a growing acknowledgement within clinical outcomes assessments of the importance of determining the impact of the disease from a patient perspective to help inform treatment decisions and measure the effectiveness of interventions.[5,6] Patient-reported outcomes (PROs) are validated instruments that may be used to capture patient self-report of symptoms and functioning, however, given the young age of the infants affected by GORD, their symptom burden needs to be assessed by a parent or caregiver by proxy through observer-reported outcome measures (ObsRO).

The revised version of the Infant Gastroesophageal Reflux Questionnaire (I-GERQ-R) is an ObsRO designed specifically to determine symptom severity in paediatric GORD as rated by caregivers. The I-GERQ-R was derived from the lengthier, 138-item I-GERQ.[7] The I-GERQ-R was designed for use in clinical trials to determine the effectiveness of interventions.[8]

The early development paper[8] reported recommended thresholds for minimally and clinically important differences (MID/CID) for the I-GERQ-R based on clinician and parental perceptions of change. However, there has been no subsequent research to further investigate the MID and CID despite a growing body of evidence of its use in trials. Furthermore, there remains in the literature, a degree of uncertainty and inconsistency in the use and interpretation of the terms MID and CID, which have often been employed interchangeably.[9] For the sake of clarity, this study primarily focused on CID, defined here as a change score or difference score sufficiently "large enough to establish the therapeutic importance of the results".[10]

There is a growing awareness, not least of which amongst regulatory agencies,[11] that statistical significance of trial results needs to be supported by clinical significance, ie, a clinically important difference. Therefore, the aim of this study was to synthesise the data available from the literature on the I-GERQ-R when used to assess a change in the symptoms of GORD in paediatric populations in order to support the published MID/CID for the instrument. The study applied a quantitative synthesis to determine an overall CID for the I-GERQ-R to help evaluating and interpreting the effects of interventions in clinical studies.

Methods

I-GERQ-R

The I-GERQ-R is a 12-item observer-reported outcome (ObsRO) measure[8] designed to capture the symptom burden associated with paediatric GORD. The items capture regurgitation, crying, refusal to feed, hiccoughs and arching, as well as apnoea/cyanosis. The threshold for determining a diagnosis of GORD is ≥16. The recall period is 1 week, and response options vary from 2 to 5 categories. A higher score indicates greater symptom burden. The instrument developers have determined a change score of 5 to 6 to be of clinical importance[7,8] based on parental/caregiver, and clinician rating of symptom change. A change of 3 on the I-GERQ-R has been deemed by the developers to represent a minimally important difference (MID), although the basis for this is not known. Therefore, in this study, an MID was defined as the smallest perceptible change perceived as beneficial (or deleterious) to the patient.[12]

Search and Quantitative Synthesis

A structured literature search was undertaken in four databases (PsycInfo, EconLit, Embase and MedLine) using the following search terms:

"infant gastro*esophageal reflux questionnaire" OR GERQ OR "I-GERQ*".

The search was not limited to English language papers or to publication date. The inclusion criteria were: Studies reporting at least 2 measurements on the I-GERQ-R over time.

Exclusion criteria were:

- Healthy Infants;
- Opinion Piece;
- Abstract with no data;
- Different outcomes, ie, not I-GERQ-R;
- No long-term data.

Studies were selected on the basis of a review of the abstract for a full-text review. This and the data extraction was undertaken by a single reviewer (ABS). The following data were extracted from the manuscripts and entered into an MS Excel spreadsheet: sample size, mean age of the infants (or median if this was not available), mean and standard deviation (baseline, post-baseline), as well as study type and intervention (where applicable). In addition to the unweighted (raw) mean change score, the weighted overall mean change was calculated using published algorithms[13] for random effects meta-analyses. A random effects model was applied as given the potential variety in study design and interventions utilising the I-GERQ-R there was an assumption that the treatment effects could vary across studies.[14,15] The weighted overall mean change was used as the CID. A standardised difference was calculated by dividing the change into I-GERQ-R score

from baseline by the baseline standard deviation. This difference and the lower 95% confidence for the weighted mean were used to interpret the MID. Consistency in findings was evaluated using Cochrane's Q, and I^2. Cochrane's Q was used to evaluate heterogeneity across the studies. Cochrane's Q is the weighted sum of squared differences between individual study effects and the pooled effect across studies, and has a Chi-squared distribution with k-1 degrees

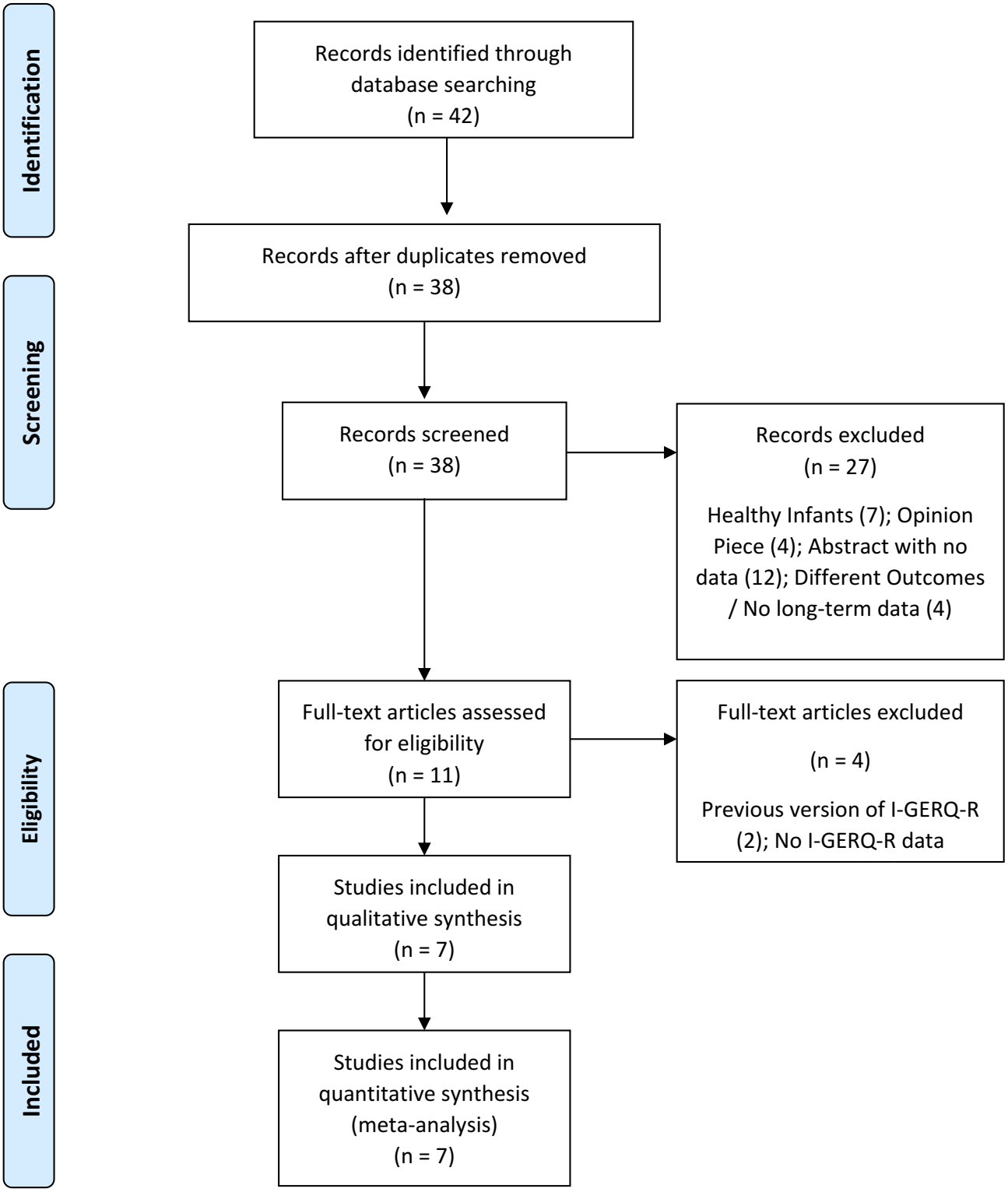

Figure 1 PRISMA flow diagram record selection for the I-GERQ-R.

Table 1 Study Details and Sample Demographics

Authors	Population	Study Design	Outcomes	Intervention	Time Period	Average Age	N	Other Comments
Vandenplas et al (2010)	Infants presenting with frequent regurgitation and reflux-associated symptoms; Inclusion criteria: aged 3 weeks to 3 months, regurgitating >4x per day for at least 2 weeks and presenting with other reflux-associated symptoms	Clinical study	Number infants with decrease in I-GERQ-R score >5; number of infants with normalisation of score (<16)	Multicare AR Bed	1 week intervention	1.5 months (range 0.75–2.5) (median)	30	≥16 for GORD diagnosis (all patients had IGERQ-R≥16 at baseline)
Campanozzi et al (2009)	Infants receiving the diagnosis of infant regurgitation according to the Rome II criteria	Observational, prospective survey: infants were enrolled in the study and re-assessed at 2 month intervals until 2 years old to determine if GOR symptoms had resolved/worsened or a diagnosis of GORD had been made	I-GERQ-R completed at enrolment and follow-up visits	None	24 months (max)	5.6 months (Standard deviation, SD 3.6)	313	A total of 2642 infants were enrolled of whom 313 were diagnosed with infant regurgitation; It is not clear whether the clinicians or parents/caregivers completed the instrument
Van Howe and Storms (2010)	Prospective cohort study, consecutive maternal-infant pairs followed-up for 6 months	I-GERQ-R was completed at 1, 2, 4 and 6 months	I-GERQ-R	None	6 months	0 to 6 months	128	For the 4 scores SE is reported in table but suspect this is SD
Khoshoo and Dhume (2008)	Infants aged ≥3 months referred to specialist clinic for evaluation and treatment of GORD; I-GERQ-R ≥16 at screening.	Patients randomised to 1 of 2 treatment groups; 15 mg lansoprazole once daily vs 7.5 mg lansoprazole twice daily. Also controls with extensively hydrolysed formula (N=15 in each group)	I-GERQ-R	Lansoprazole	2 weeks	4.8 months (1.18) Group A; 4.3 (1.01) Group B; 4.6 (0.99) Group C	45	
Neu et al (2014)	Infants aged 4–10 weeks at enrolment scoring ≥16 on the I-GERQ-R	Patients randomised to receive massage (M) or non-massage therapy (NM) 2x week for 6 weeks (pilot RCT)	I-GERQ-R	Massage therapy	6 weeks	7.3 months (SD 1.6) (M) 7.6 (SD 2.3) (NM)	36	
Orenstein and McGowan (2008)	Infants scoring ≥16 on the I-GERQ-R at enrolment	Non-pharmacological interventions (feeding modification; positioning)	I-GERQ-R	H2RA	2 weeks	13 weeks (median) (4–43 range)	37	The range was used to estimate the standard deviation (range/4). Medians were assumed to approximate the means
Baldassare et al (2020)	Infants scoring ≥ 16 on the I-GERQ-R at enrolment	Randomised multicentre cross-over study	I-GERQ-R	Thickened formula; Magnesium alginate	5 weeks	65.5 days (SD 46.3)	72	The cross-over groups received magnesium alginate and thickened formula; a third group exclusively breastfed infants received magnesium alginate

of freedom (k is the number of studies). A statistically significant Cochrane's Q indicates potential heterogeneity between the studies. I^2 is the percentage variation across studies due to heterogeneity and was calculated as follows: $I^2 = 100\%$ x (Q-df)/Q. The higher the percentage, the greater the degree of heterogeneity. I^2 values <0 are censored at zero. Potential publication bias was assessed by plotting study treatment effect against sample size using a funnel plot.

Results

Studies and Patients

An initial 42 studies were retrieved from the search. Figure 1 includes the PRISMA flow diagram for the selection process. A final total of 7 studies were selected for full data extraction.

Table 1 provides a description of the studies included. A total of 661 (study sample size range: 19–313) infants participated in these studies with ages ranging from 0 up to 6 months; 424 infants had been diagnosed with GORD (64%). Five of the seven studies used the ≥16 threshold as an inclusion criterion. The types of studies included: two randomised control trials (RCTs), one of which included a pharmacological intervention[16] (proton-pump inhibitor (PPI, lansoprazole) versus hydrolysed formula), the other non-pharmacological intervention (massage therapy).[17] One randomised cross-over trial.[23] The other studies comprised a study investigating a positional bed,[18] as well as three observational studies.[19–21] The study periods varied from 1 week up to 24 months. The I-GERQ-R had been completed on at least 2 occasions for each study. Minimum length of time between completion was 1 week and the maximum was 6 weeks.[17]

Quantitative Synthesis

Table 2 includes the effect sizes and standardised mean differences. Figure 2 is the Forest plot of the studies included. One study[21] reported the median change, as well as the associated range. The median was assumed to approximate the mean change for this study, and the standard deviation for the change score was estimated using a commonly used metric (range/4).[22] The unweighted overall mean I-GERQ-R score was −6.46 (95% Confidence Intervals (CI): −4.69 to −8.23). The weighted overall mean was −6.54 (95% CI: −4.35 to −8.74); Cochrane's Q: 17.96, p=0.08; and $I^2$22.04. This suggests an average treatment effect of 6.54 with little heterogeneity across findings.[14,15] The standardised mean difference was 2.9. Figure 3 shows the study effect plotted against the sample size. It may be seen that all points bar one fell within the 95% CIs, suggesting virtually no publication bias.

Table 2 Effects Sizes and Standardised Mean Differences

Authors	Effect Size	95% CIL	95% CIH	SMD
Vandenplas et al (2010)	−6.09	−7.75	−4.43	−1.43
Campanozzi et al (2009)	−4.80	−5.24	−4.36	−1.30
Campanozzi et al (2009)	−3.80	−4.24	−3.36	−0.84
Van Howe & Storms (2010)	−1.52	−2.34	−0.70	−0.32
Van Howe & Storms (2010)	−3.63	−4.60	−2.66	−0.65
Van Howe & Storms (2010)	−5.43	−6.55	−4.31	−0.84
Khoshoo & Dhume (2008)	−6.00	−7.25	−4.75	−2.14
Khoshoo & Dhume (2008)	−6.90	−8.15	−5.65	−1.86
Neu et al (2014)	−7.60	−9.07	−6.13	−1.90
Neu et al (2014)	−9.80	−11.43	−8.17	−2.45
Orenstein & McGowan (2008)	−5.25	−7.14	−3.36	−1.05
Baldassare et al (2020)	−8.96	−10.83	−7.09	−1.29
Baldassare et al (2020)	−9.74	−11.80	−7.68	−1.27
Baldassare et al (2020)	−10.95	−12.47	−9.43	−3.25

Note: The separate rows for the same studies represent the outcomes for different study arms and/or interventions within those studies.
Abbreviations: CIL, lower limit confidence interval; CIH, higher limit confidence interval; SMD, standardised mean difference.

Discussion

This study synthesised change scores in the I-GERQ-R from the published literature. It is the first study to the

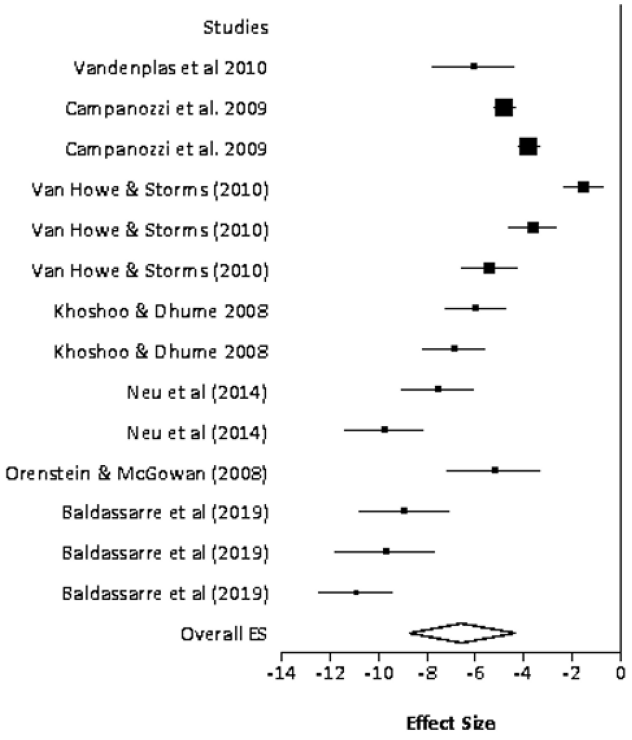

Figure 2 Forest plot of I-GERQ-R studies.
Notes: Cochrane's Q=17.96, p=0.08; I^2=22.04.

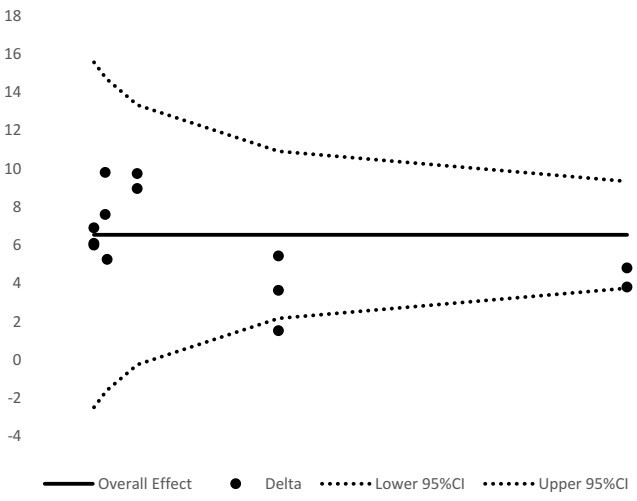

Figure 3 Funnel plot to evaluate potential bias.

authors' knowledge to calculate an overall effect size for the I-GERQ-R based on empirical data. The results suggested an effect size of around 6. This is in line with that previously determined using an anchor-based approach (parent/clinician rating of change)[8] and indicates that a change score of 6 on the I-GERQ-R may be interpreted as clinically significant. In addition to this, the standardised mean difference was around 3 and the lower limit of the 95% CI around 4, suggests a change score or difference of 3 to 4 could be considered as a minimally important difference (MID). This outcome differs slightly from previous studies which have suggested an MID of 3.[8]

The potential limitation to this study is that it was a structured review, rather than a systematic review, ie, there was no grey literature search and the studies included were not evaluated for any bias. Furthermore, the selection was only undertaken by a single reviewer. The small number of studies also meant that further analysis of whether the timing of the intervention, and sample age impacted on the effect size could not be undertaken. This may, therefore, impact on the generalisability of the results. In one study[19] it was not entirely clear whether clinician or parents completed the I-GERQ-R. However, it was decided to retain this paper in the analysis given previous work has suggested a high degree of association between parent/caregiver responses to the I-GERQ-R and those by clinicians.[8] In this study, the CID and MID were calculated using distribution-based methods which are potentially subject to the variability inherent in the samples. Nevertheless, the CID, in particular, corresponded well with those previously derived with anchor-based methods, ie, the clinician and parent global impression of change.[8] Finally, it has been noted that the I-GERQ-R may not able to separate GORD from colic in children aged <3 months.[20] Specifically, that study observed that whereas I-GERQ-R levels decreased over time (up to 6 months), regurgitation frequency remained relatively constant. The presence in studies of higher number of infants with colic could therefore unduly influence the study effect. A re-analysis of the data including only those studies with infants aged >3 months resulted in a treatment effect/CID of 5.41 (95% CI 2.24 to 3.60). It may therefore be concluded that the presence of infants <3 months (with or without colic) had negligible impact on the derived CID.

Conclusions

In summary, the results indicate a degree of consistency in change scores on the I-GERQ-R despite the use of the instrument in different study designs. These may ultimately help shape clinical trial designs for infants with GORD, as well as inform clinical decisions in evaluating meaningful change in symptom severity in children affected by GORD. In particular, the results of this study should enable sponsors to determine adequate sample size to power studies investigating interventions in paediatric GORD. In addition to this, it should also help clinicians in monitoring the symptoms of GORD in infants, and in determining the impact of treatment on the condition.

Acknowledgment

A previous version of the abstract was presented at the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Europe Conference 2018: Smith AB, Hodgkinson V, Kotze H, Coyle C. The infant gastroesophageal reflux disease questionnaire (I-GERQ-R): A meta-analysis. *Value in Health* 2018: 21(S3): S389-S390.

Disclosure

All authors are employees of Reckitt Benckiser. The authors report no other conflicts of interest in this work.

References

1. Dhillon A, Ewer A. Diagnosis and management of gastro-oesophageal reflux in preterm infants in neonatal intensive care units. *Acta Paediatr*. 2004;93(1):88–93. doi:10.1111/j.1651-2227.2004.tb00680.x
2. Salvatore S, Vandenplas Y. Epidemiology. In: Vandenplas Y, editor. *Gastroesophageal Reflux in Children*. London: Springer International Publishing; 2017:1–14.

3. Brown P. Medical management of gastroesophageal reflux. *Curr Opin Pediatr*. 2000;12:247–250. doi:10.1097/00008480-200006000-00012
4. Lightdale JR, Gremse DA, Heitlinger LA, et al. Gastroesophageal reflux: management guidance for the pediatrician. *Pediatrics*. 2013;131:e1684–e1695. doi:10.1542/peds.2013-0421
5. Deshpande PR, Rajan S, Sudeepthi BL, Nazir CPA. Patient-reported outcomes: a new era in clinical research. *Perspect Clin Res*. 2011;2(4):137–144. doi:10.4103/2229-3485.86879
6. Anker SD, Agewall S, Borggrefe M, et al. The importance of patient-reported outcomes: a call for their comprehensive integration in cardiovascular clinical trials. *Eur Heart J*. 2014;35(30):2001–2009. doi:10.1093/eurheartj/ehu205
7. Orenstein SR, Shalaby TM, Cohn JF. Reflux symptoms in 100 normal infants: diagnostic validity of the infant gastroesophageal reflux questionnaire. *Clin Pediatr (Phila)*. 1996;35:607–614. doi:10.1177/000992289603501201
8. Kleinman L, Rothman M, Strauss R, et al. The infant gastroesophageal reflux questionnaire revised: development and validation as an evaluative instrument. *Clin Gastroenterol Hepatol*. 2006;4:588–596. doi:10.1016/j.cgh.2006.02.016
9. King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res*. 2011;11(2):171–184. doi:10.1586/erp.11.9
10. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. *J Pain*. 2008;9(2):105–121. doi:10.1016/j.jpain.2007.09.005
11. European Medicines Agency. Committee for Medicinal Products for Human Use (CHMP) (2011). Guideline on the evaluation of drugs for the treatment of gastro-oesophageal reflux disease. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-evaluation-drugs-treatment-gastro-oesophageal-reflux-disease_en.pdf. Accessed March 6, 2018.
12. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*. 1998;16(1):139–144. doi:10.1200/JCO.1998.16.1.139
13. Lipsey MW, Wilson DB. *Practical Meta-Analysis*. Thousand Oaks (CA): Sage Publications; 2001.
14. Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *Br Med J*. 2011;342:d549. doi:10.1136/bmj.d549
15. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Br Med J*. 2003;327:557–560. doi:10.1136/bmj.327.7414.557
16. Khoshoo V, Dhume P. Clinical response to 2 dosing regimens of lansoprazole in infants with gastroesophageal reflux. *J Pediatr Gastroenterol Nutr*. 2008;46(3):352–354. doi:10.1097/MPG.0b013e31815667d7
17. Neu M, Pan Z, Workman R, Marcheggiani-Howard C, Furuta G, Laudenslager ML. Benefits of massage therapy for infants with symptoms of gastroesophageal reflux disease. *Biol Res Nurs*. 2014;16(4):387–397. doi:10.1177/1099800413516187
18. Vandenplas Y, de Schepper J, Verheyden S, et al. A preliminary report on the efficacy of the Multicare AR-bed in 3-week-3-month-old infants on regurgitation, associated symptoms and acid reflux. *Arch Dis Child*. 2010;95(1):26–30. doi:10.1136/adc.2008.156497
19. Campanozzi A, Boccia G, Pensabene L, et al. Prevalence and natural history of gastroesophageal reflux: pediatric prospective survey. *Pediatrics*. 2009;123(3):779–783. doi:10.1542/peds.2007-3569
20. Van Howe RS, Storms MR. Gastroesophageal reflux symptoms in infants in a rural population: longitudinal data over the first six months. *BMC Pediatr*. 2010;10:7. doi:10.1186/1471-2431-10-7
21. Orenstein SR, McGowan JD. Efficacy of conservative therapy as taught in the primary care setting for symptoms suggesting infant gastroesophageal reflux. *J Pediatr*. 2008;152(3):310–314. doi:10.1016/j.jpeds.2007.09.009
22. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2005;5:13. doi:10.1186/1471-2288-5-13
23. Baldassare ME, Di Mauro A, Pignatelli MC, et al. Magnesium alginate in gastro-esophageal reflux: a randomized multicentre cross-over study in infants. *Int J Environ Res Public Health*. 2020;17:83. doi:10.3390/ijerph17010083