NCC-WCH

Document information (i.e. version number etc)

Gastro-oesophageal reflux disease in children and young people: Appendix J

Gastro-oesophageal reflux disease: recognition, diagnosis and management in children and young people

.

Appendix J: GRADE tables

Draft for Consultation

Commissioned by the National Institute for Health and Care Excellence

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Copyright
National Collaborating Centre for Women's and Children's Health 2014

Funding
National Institute for Health and Care Excellence 2014

Contents

Appendix J: GRADE Tables 5
J.1 Diagnosing and investigating GORD 5
J.1.1 Natural course of overt regurgitation 5
J.1.2 Signs and symptoms 7
J.1.3 Risk Factors 34
J.1.4 Indications for investigation and treatment 49
J.2 Initial management of GOR and GORD 50
J.2.1 Infant Positioning 50
J.2.2 Feeding changes 59
J.2.3 Alginates and Antacids 69
J.3 Pharmacological treatment of GORD 78
J.4 Enteral feeding for GORD 90
J.5 Surgery for GORD 91

Appendix J: GRADE Tables

J.1 Diagnosing and investigating GORD

J.1.1 Natural course of overt regurgitation

Table 5: GRADE findings for natural history of GOR

Quality assessment							
Number. of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Quality**
Natural history of overt GOR							
1 (Campanozzi et al, 2009)	Prospective cohort	Serious[a]	None	None	None	Some[b]	Moderate
1 (De et al, 2001)	Cross-sectional	Serious[c]	None	None	None	None	Moderate
1 (Gunasekaran et al, 2008)	Cross-sectional	No serious	None	None	None	None	High
1 (Hegar et al, 2004)	Cross-sectional	No serious	None	None	None	None	High
1 (Hegar et al, 2009)	Prospective cohort	Serious[d]	None	None	None	None	Moderate
1 (Hegar et al, 2013)	Prospective cohort	Serious[e]	None	None	None	None	Moderate
1 (Lacono et al, 2005)	Prospective cohort	No serious	None	None	None	None	High
1 (Martin et al, 2002)	Prospective cohort	Serious[a]	None	None	None	None	Moderate
1 (Miyazawa et al, 2002)	Cross-sectional	No serious	None	None	None	None	High
1 (Nelson et al, 1997)	Cross-sectional	Serious[c]	None	None	None	None	Moderate
1 (Nelson et al, 1998)	Case-control	Very serious[a,c]	None	None	None	None	Low
1 (Orenstein et al, 1996)	Case-control	Serious[c]	None	None	None	Some[f]	Moderate
1 (Osatakul et al, 2002)	Prospective cohort	Serious[c]	None	None	None	None	Moderate

Quality assessment							
Number. of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Quality**
1 (Ruigomez et al, 2010)	Retrospective cohort	Very serious[g]	None	Some[h]	None	None	Very low
1 (Van Howe et al, 2010)	Prospective cohort	Very serious [a,c]	None	None	None	None	Low

[a] *Unclear whether loss to follow-up is unrelated to key characteristics*
[b] *Prematurity: 8.6% premature at entry to study*
[c] *Outcome is not clearly defined: definition of regurgitation not reported*
[d] *All dropouts because of excessive symptoms were in the partially breastfed group*
[e] *Presentation of results not particularly clear: it has been assumed that the infants for which data has not been presented are ones that did not regurgitate rather than being considered as missing data or infants lost to follow up (as authors state 4 subjects were lost to follow up). Also, unclear how many subjects were given conservative treatment.*
[f] *Prematurity: 26% of those attending well-baby clinic and 14% of those referred to gastroenterology department premature at entry to study*
[g] *Retrospective study design, based on electronic medical records across a number of GP practices, so variation in tests and treatments, only 15.3% of GERD cohort had a record of a formal diagnostic test being undertaken, none of the children in the control cohort had been tested for GER.*
[h] *This study examines GERD not regurgitation.*

J.1.2 Signs and symptoms

Table 6: GRADE profile of systematic reviews of symptoms and signs.

Quality assessment							
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Quality
Identification of symptoms and signs of GORD							
1 Vandenplas et al, 2009	Systematic Review & Consensus	Very Serious[a,b]	None	None	None	None	Low
1 Tolia et al, 2009	Systematic Review	Serious[a]	None	None	None	None	Moderate
1 Sherman et al, 2009	Systematic Review & consensus	Serious[a]	None	None	None	None	Moderate

[a] *Search strategy not presented*
[b] *Inclusion and exclusion criteria not presented*

J.1.2.1 Distressed behaviour

Table 8: GRADE findings for evaluation of diagnostic value of symptoms of distress for identifying presence of GORD.

Quality assessment								Measure of diagnostic accuracy**						
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	Quality
Cries more than normal in the opinion of the parent used to identify presence of GOR/D														
1 (Orenstein et al, 1996)	Prospective Case-control	Serious[a]	None	None	Serious[b]	None	135	0.54 [0.37, 0.71]	0.86 [0.76, 0.92]	-*	-*	3.88 [2.19, 6.88]	0.53 [0.37, 0.77]	Low
1 (Salvatore et al, 2005)	Prospective Cohort	None	None	None	Serious[b]	None	99	0.62 [0.38, 0.82]	0.52 [0.4, 0.63]	0.25 [0.14, 0.4]	0.84 [0.7, 0.93]	1.29 [0.86, 1.93]	0.73 [0.41, 1.32]	Moderate
Cries for more than 1 hour per day used to identify presence of GOR/D														
1 (Orenstein et al, 1996)	Prospective Case-control	Serious[a]	None	None	Serious[b]	None	135	0.54 [0.37, 0.71]	0.83 [0.75, 0.9]	-*	-*	3.19 [1.88, 5.42]	0.55 [0.38, 0.8]	Low
1 (Salvatore et al, 2005)	Prospective Cohort	None	None	None	Serious[b]	None	99	0.33 [0.15, 0.57]	0.82 [0.72, 0.9]	0.33 [0.15, 0.57]	0.82 [0.72, 0.9]	1.88 [0.87, 4.06]	0.81 [0.59, 1.12]	Moderate
Cries for more than 3 hours per day used to identify presence of GOR/D														
1 (Orenstein et al, 1996)	Case-control	Serious[c]	None	None	Serious[b]	None	135	0.29 [0.15, 0.46]	0.97 [0.71, 0.99]	-*	-*	9.52 [2.78, 32.63]	0.74 [0.6, 0.91]	Low

Quality assessment								Measure of diagnostic accuracy**						
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	Quality
1 (Salvatore et al, 2005)	Prospective Cohort	None	None	None	Serious[b]	None	99	0.57 [0.34, 0.78]	0.61 [0.49, 0.72]	0.28 [0.15, 0.44]	0.84 [0.72, 0.93]	1.46 [0.92, 2.31]	0.71 [0.42, 1.19]	Moderate
Crying when feeding used to identify presence of GOR/D														
1 (Orenstein et al, 1996)	Prospective Case-control	Serious[a]	None	None	Serious[b]	None	135	0.8 [0.63, 0.92]	0.86 [0.85, 0.92]	0.67 [0.5, 0.8]	0.92 [0.85, 0.97]	5.71 [3.42, 9.55]	0.23 [0.12, 0.45]	Low
1 (Salvatore et al, 2005)	Prospective Cohort	None	None	None	Serious[b]	None	99	0.57 [0.34, 0.78]	0.61 [0.72, 0.72]	0.28 [0.15, 0.44]	0.84 [0.72, 0.93]	1.46 [0.92, 2.31]	0.71 [0.42, 1.19]	Moderate
1 (Mathisen et al, 1999)	Prospective cohort	Serious[c]	None	None	Serious[b]	None	40	0.85 [0.62, 0.97]	0.8 [0.6, 0.94]	0.81 [0.58, 0.95]	0.84 [0.6, 0.97]	4.25 [1.74, 10.41]	0.19 [0.06, 0.54]	Low
Back arching or abnormal posturing used to identify presence of GOR/D														
1 (Orenstein et al, 1996)	Prospective Case-control	Serious[a]	None	None	Serious[b]	None	135	0.6 [0.42, 0.76]	0.9 [0.78, 0.95]	-*	-*	6 [3.14, 11.46]	0.44 [0.29, 0.67]	Low
1 (Carr et al, 2000)	Retrospective cohort	Very Serious[d]	None	None	None	None	295	0.03 [0.01, 0.06]	1 [0.96, 1]	1 [0.54, 1]	0.28 [0.23, 0.34]	∞	0.97 [0.95, 0.99]	Low

Quality assessment								Measure of diagnostic accuracy**						
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	Quality
1 (Deal et al, 2005) (1 - 11 months)	Prospective Case-control	Serious[e]	None	None	Serious[b]	None	67	0.66 [0.49, 0.8]	0.78 [0.56, 0.93]	-*	-*	3.03 [1.35, 6.78]	0.44 [0.27, 0.7]	Low
1 (Costa et al, 2004)	Cross-sectional survey	Very serious[f]	None	None	None	None	797	0.45 [0.34, 0.56]	0.97 [0.95, 0.98]	0.63 [0.5, 0.74]	0.93 [0.91, 0.95]	13.26 [8.41, 20.91]	0.57 [0.47, 0.69]	Very Low
Waking > 3/night > 2h/night used to identify presence of GOR/D														
1 (Ghaem et al, 1998)	Case-control	None	None	None	Serious[b]	None	102	0.55 [0.43, 0.67]	0.73 [0.52, 0.88]	-*	-*	2.05 [1.06, 3.99]	0.61 [0.43, 0.86]	Moderate

[a] *Classification of control group was based on not being treated for GORD. The GORD group was based on pH monitoring*
[b] *Wide confidence intervals covering categories from low to high.*
[c] *Children in the control group were not tested for GOR. Small sample size*
[d] *Retrospective chart review based on diagnosis of GERD*
[e] *Presence of GORD was based on clinical judgement, which would include items contained in questionnaire*
[f] *Definition of GORD based on Rome II criteria, no objective measure undertaken*
[*] *Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated.*
[**] *Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.2 Apnoea

Table 9: GRADE findings for evaluation of the temporal association between apnoea for GOR

Quality assessment							Number of children	Temporal association	Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations			
Temporal link between apnoea and reflux in infants									
1 (Mousa et al, 2005)	Cohort	Serious[a]	None	Serious[b]	None	Yes	25	6173 5-minute time events were recorded across the 25 children. 4706 (76.2%) of the time events had no GER or apnoea. 89 had apnoea with GER. 439 apnoea events alone. 939 reflux alone. In 2 of 25 children apnoea and GER events was statistically associated. Across the whole group the association was not statistically significant (p = 0.214).	Low
Temporal link between apnoea and reflux in premature infants									
1 (Peters et al, 2002)	Cohort	Serious[c]	None	Serious[d]	None	No	19	A total of 524 reflux events and 2039 apnoea events were recorded. Apnoea during reflux free periods no different from apnoea during reflux periods (0.19/min [0.00 to 0.85] vs 0.25/min [0.00 to 1.15]); p > 0.05 in 19 infants.	Low

[a] *Small sample size*
[b] *11 of 25 children were premature*
[c] *Small sample size*
[d] *Examining a specific group of AOP*

Table 10: GRADE findings for evaluation of apnoea for identifying GORD

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Apparent Life Threatening Event used to identify presence of GOR/D														

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Sacre et al, 1989)	Case-control study	None	None	None	None	None	449	0.42 [0.3, 0.55]	0.91 [0.88, 0.94]	-*	-*	4.92 [3.17, 7.62]	0.63 [0.51, 0.79]	High
1 (Tolia et al, 2003)	Retrospective chart review	Very Serious[b]	None	None	Serious[c]	Yes[d]	342	0.31 [0.24, 0.38]	0.8 [0.74, 0.86]	0.6 [0.49, 0.71]	0.54 [0.48, 0.61]	1.57 [1.07, 2.28]	0.86 [0.76, 0.98]	Very low
Recurrent apnoea used to identify presence of GOR/D														
1 (Mazliah et al, 2000)	Cross-sectional survey	Serious[e]	None	None	Serious[c]	None	44	0.06 [0.01, 0.21]	1 [0.75, 1]	1 [0.16, 1]	0.31 [0.18, 0.47]	∞	0.94 [0.85, 1.03]	Low
Apnoea ever used to identify presence of GOR/D														
1 (Orenstein et al, 1996)	Case-control	Serious[f]	None	None	Serious[c]	None	135	0.43 [0.26, 0.61]	0.98 [0.93, 1]	-*	-*	21.43 [5.16, 89.04]	0.58 [0.44, 0.78]	Low
Apnoea with cyanosis used to identify presence of GOR/D														
1 (Orenstein et al, 1996)	Case-control	Serious[f]	None	None	Serious[c]	None	135	0.17 [0.07, 0.34]	1 [0.96, 1]	-*	-*	∞	0.83 [0.71, 0.96]	Low

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Salvatore et al, 2005)	Cohort	None	None	None	Serious[c]	None	99	0.11 [0.01, 0.35]	0.85 [0.75, 0.92]	0.15 [0.02, 0.45]	0.8 [0.69, 0.88]	0.75 [0.18, 3.08]	1.04 [0.86, 1.26]	Moderate
Apnoea (not specified) used to identify presence of GOR/D														
1 (Koda et al, 2010)	Retrospective cohort	Very Serious[g]	None	None	None	None	307	0.18 [0.09, 0.3]	0.87 [0.82, 0.91]	0.24 [0.12, 0.39]	0.83 [0.78, 0.87]	1.4 [0.73, 2.68]	0.94 [0.83, 1.07]	Low
1 (Costa et al, 2004)	Cross-sectional	Very serious[h]	None	None	None	None	798	0.35 [0.25, 0.46]	0.97 [0.95, 0.98]	0.58 [0.44, 0.72]	0.92 [0.9, 0.94]	11.21 [6.8, 18.48]	0.67 [0.58, 0.78]	Low
1 (Carret al, 2000)	Retrospective cohort	Very Serious[i]	None	None	None	None	295	0.03 [0.01, 0.06]	0.93 [0.85, 0.97]	0.5 [0.21, 0.79]	0.27 [0.21, 0.32]	0.38 [0.13, 1.14]	1.05 [0.98, 1.12]	Low
1 (Assadamongkol et al, 1993)	Retrospective cohort	Very Serious[j,l]	None	None	Very Serious[c]	None	55	0.12 [0.02, 0.3]	0.97 [0.82, 1]	0.75 [0.19, 0.99]	0.55 [0.4, 0.69]	3.35 [0.37, 30.21]	0.92 [0.78, 1.07]	Very low
1 (Yuksel et al, 2014)	Case-control	Serious[l]	None	Serious[m]	None	None	71	0.05 [0.01, 0.17]	1 [0.89, 1]	-*	-*	∞	0.95 [0.88, 1.02]	Low
Apnoea in preterm infants only used to identify presence of GOR/D														

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Mezzacappa et al, 2008)	Retrospective case-control	Very Serious[k]	None	None	None	None	194	0.94 [0.87, 0.98]	0.13 [0.06, 0.21]	-*	-*	1.08 [0.98, 1.19]	0.45 [0.16, 1.25]	Low

[a] *Children admitted due to ALTE*
[b] *Retrospective chart review based on diagnosis of GERD*
[c] *Wide confidence intervals covering categories from low to high.*
[d] *ALTE as a presenting symptom. ALTE not defined*
[e] *Method of confirming GORD varied between children.*
[f] *Classification of control group was based on not being treated for GORD.*
[g] *Retrospective chart review*
[h] *Definition of GERD included having apnoea*
[I] *Retrospective chart review*
[j] *Retrospective chart review*
[k] *Retrospective chart review*
[l] *Small sample size*
[m] *All children had otitis media*
[*] *Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated.*
[**] *Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.3 Epigastric or chest pain

Table 11: GRADE findings for evaluation of abdominal and chest pain in children and young adults for identifying presence of GORD

Quality assessment							Number of patients	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Chest pain (including heartburn) used to identify presence of GOR/D														
1 (Stordal et al, 2005)	Cohort	None	None	None	Serious[a]	Some[b]	99	0.27 [0.14, 0.44]	0.81 [0.69, 0.9]	0.45 [0.24, 0.68]	0.65 [0.53, 0.75]	1.4 [0.67, 2.91]	0.9 [0.72, 1.14]	Moderate
1 (Stordal et al, 2005)	Case-control	Serious[c]	None	None	Serious[a]	Some[b]	321	0.27 [0.14, 0.44]	0.96 [0.93, 0.98]	-[*]	-[*]	6.98 [3.18, 15.3]	0.76 [0.62, 0.92]	Low
1 Carr et al, 2000	Retrospective case-control	Very Serious[d]	None	None	Very serious[a]	Some[e]	295	0.12 [0.08, 0.17]	0.79 [0.69, 0.87]	-[*]	-[*]	0.58 [0.33, 1.01]	1.11 [0.98, 1.26]	Very low
Abdominal pain or "stomach ache" used to identify presence of GOR/D														
1 (Stordal et al, 2005)	Cohort	None	None	None	Very serious[a]	Some[b]	99	0.62 [0.45, 0.78]	0.16 [0.08, 0.28]	0.31 [0.21, 0.42]	0.42 [0.22, 0.63]	0.74 [0.56, 0.97]	2.35 [1.16, 4.73]	Low
1 (Stordal et al, 2005)	Case-control	Serious[c]	None	None	Very serious[a]	Some[b]	321	0.62 [0.45, 0.78]	0.67 [0.61, 0.72]	0.2 [0.13, 0.28]	0.93 [0.89, 0.96]	1.88 [1.39, 2.54]	0.57 [0.37, 0.86]	Very low

Quality assessment							Number of patients	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Carr et al, 2000)	Retrospective cohort	Very serious[d]	None	None	Very serious[a]	Some[e]	295	0.18 [0.13, 0.24]	0.63 [0.52, 0.73]	0.56 [0.43, 0.68]	0.22 [0.17, 0.28]	0.48 [0.32, 0.72]	1.31 [1.09, 1.56]	Very low
1 (Deal et al, 2005)	Case-control	Serious[f]	None	None	Very serious[a]	None	67	0.43 [0.27, 0.59]	0.96 [0.81, 1]	-*	-*	11.48 [1.62, 81.21]	0.6 [0.45, 0.79]	Very low
1 (Uzun et al, 2012)	Retrospective cohort	Very serious[d]	None	None	Very serious[a]	Some[g]	70	0.23 [0.11, 0.39]	0.87 [0.7, 0.96]	0.69 [0.39, 0.91]	0.47 [0.34, 0.61]	1.79 [0.61, 5.26]	0.88 [0.71, 1.1]	Very low
Epigastric pain used to identify presence of GOR/D														
1 (Stordal et al, 2005)	Cohort	None	None	None	Serious[a]	Some[b]	99	0.27 [0.14, 0.44]	0.56 [0.43, 0.69]	0.27 [0.14, 0.44]	0.56 [0.43, 0.69]	0.62 [0.34, 1.13]	1.29 [0.96, 1.73]	Moderate
1 (Stordal et al, 2005)	Case-control	Serious[c]	None	None	Serious[a]	Some[b]	321	0.27 [0.14, 0.44]	0.93 [0.89, 0.96]	-*	-*	3.84 [1.95, 7.56]	0.79 [0.64, 0.96]	Low

a Wide confidence intervals covering categories from low to high.
b Based on children referred for pH assessment
c Unknown if control group had abnormal pH as not tested.
d Based on retrospective review of medical notes. Based on recorded symptoms rather than questionnaire.
e Mean average age was 4.4 years so accuracy of symptoms reporting is unclear.
f Presence of GORD was based on clinical judgement rather than a diagnostic test.
g Children aged 2 to 17 years – so reliability of reporting across the group is unclear.
** Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated.*
*** Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.4 Hoarseness

Table 12: GRADE findings for evaluation of hoarseness to identify GORD

Quality assessment								Measure of diagnostic accuracy**						
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Number of children	Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	Quality
Hoarseness														
1 (Carr et al, 2000)	Retrospective cohort	Very serious[a]	None	None	None	None	295	0.34 [0.28, 0.41]	0.54 [0.43, 0.65]	0.66 [0.57, 0.75]	0.24 [0.18, 0.31]	0.75 [0.55, 1.01]	1.21 [0.97, 1.51]	Low
1 (Yuksel et al, 2014)	Case-control	Serious[b]	None	Serious[c]	Very serious[d]	None	71	0.08 [0.02, 0.21]	0.97 [0.84, 1]	0.75 [0.19, 0.99]	0.46 [0.34, 0.59]	2.46 [0.27, 22.54]	0.95 [0.85, 1.06]	Very low

[a] *Retrospective chart review*
[b] *Retrospective chart review*
[c] *All children had Otitis Media*
[d] *Confidence intervals cover several categories of usefullness*
** *Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.5 Feeding difficulties

Table 13: GRADE findings for evaluation of feeding difficulties to identify GORD

Quality assessment							Number of patients	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Feeding refusal used to identify presence of GOR/D														
1 (Deal et al, 2005) (1 - 11 months	Case-control	Serious[a]	None	None	Very serious[b]	None	67	0.41 [0.26, 0.58]	0.83 [0.61, 0.95]	-*	-*	2.38 [0.91, 6.24]	0.71 [0.52, 0.97]	Very low
1 (Deal et al, 2005) (12 or older months)	Case-control	Serious[a]	None	None	Very serious[b]	None	67	0.65 [0.48, 0.79]	0.76 [0.56, 0.9]	-*	-*	2.69 [1.36, 5.34]	0.46 [0.29, 0.74]	Very low
1 (Heine et al, 2006)	Cohort	None	None	Serious[c]	None	None	151	0.46 [0.26, 0.67]	0.58 [0.48, 0.66]	0.18 [0.09, 0.3]	0.84 [0.74, 0.91]	1.08 [0.67, 1.75]	0.94 [0.63, 1.4]	Moderate
1 (Orenstein et al, 1996)	Case-control	Serious[c]	None	None	Serious[b]	None	135	0.31 [0.17, 0.49]	0.96 [0.9, 0.99]	-*	-*	7.86 [2.67, 23.08]	0.71 [0.57, 0.9]	Low
1 (Salvatore et al, 2005)	Cohort	None	None	None	Serious[b]	None	99	0.52 [0.3, 0.74]	0.4 [0.29, 0.52]	0.19 [0.1, 0.32]	0.76 [0.6, 0.88]	0.88 [0.56, 1.37]	1.18 [0.7, 2]	Moderate

Quality assessment							Number of patients	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Carret al, 2000)	Retrospective cohort	Very Serious[d]	None	None	Very serious[b]	None	295	0.22 [0.17, 0.28]	0.79 [0.69, 0.87]	0.73 [0.61, 0.84]	0.28 [0.22, 0.34]	1.05 [0.64, 1.71]	0.99 [0.86, 1.13]	Very low
Feeding difficulties used to identify presence of GOR/D														
1 (Heine et al, 2006)	Cohort	None	None	Serious[c]	None	None	151	0.46 [0.26, 0.67]	0.58 [0.48, 0.66]	0.18 [0.09, 0.3]	0.84 [0.74, 0.91]	1.08 [0.67, 1.75]	0.94 [0.63, 1.4]	Moderate
Choking gagging used to identify presence of GOR/D														
1 (Carret al, 2000)	Retrospective cohort	Very Serious[d]	None	None	Very serious[b]	None	295	0.24 [0.18, 0.3]	0.86 [0.77, 0.93]	0.82 [0.7, 0.91]	0.3 [0.24, 0.36]	1.75 [0.96, 3.2]	0.88 [0.79, 0.99]	Very low
Crying when feeding used to identify presence of GOR/D														
1 (Salvatore et al, 2005)	Cohort	None	None	None	Serious[b]	None	99	0.57 [0.34, 0.78]	0.61 [0.49, 0.72]	0.28 [0.15, 0.44]	0.84 [0.72, 0.93]	1.46 [0.92, 2.31]	0.71 [0.42, 1.19]	Moderate
Mathisen et al, 1999	Case-control	Serious	None	None	Serious[b]	None	40	0.85 [0.62, 0.97]	0.8 [0.56, 0.94]	-*	-*	4.25 [1.74, 10.41]	0.19 [0.06, 0.54]	Low
Feeding problems used to identify presence of GOR/D														

Quality assessment							Number of patients	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Mazliah et al, 2000)	Cross-sectional survey	None	None	None	Very serious[b]	None	44	0.06 [0.01, 0.21]	0.92 [0.64, 1]	0.67 [0.09, 0.99]	0.29 [0.16, 0.46]	0.84 [0.08, 8.46]	1.01 [0.84, 1.22]	Very low
Feeding intolerance used to identify presence of GOR/D														
1 (Mezzacappa et al, 2008)	Retrospective Case-control	Very Serious[d]	None	None	Serious[b]	None	174	0.71 [0.61, 0.8]	0.4 [0.3, 0.51]	-*	-*	1.19 [0.96, 1.48]	0.71 [0.47, 1.09]	Very Low
Head aversion when feeding used to identify presence of GOR/D														
1 (Mathisen et al, 1999)	Case-control	Serious	None	None	Serious[b]	None	40	0.2 [0.06, 0.44]	0.9 [0.68, 0.99]	-*	-*	2 [0.41, 9.71]	0.89 [0.68, 1.16]	Low
Facial grimaces when feeding used to identify presence of GOR/D														
1 (Mathisen et al, 1999)	Case-control	Serious	None	None	Serious[b]	None	40	0.35 [0.15, 0.59]	0.8 [0.56, 0.94]	-*	-*	1.75 [0.61, 5.05]	0.81 [0.55, 1.2]	Low
Body withdrawal when feeding used to identify presence of GOR/D														
1 (Mathisen et al, 1999)	Case-control	Serious	None	None	Serious[b]	None	40	0.2 [0.06, 0.44]	0.95 [0.75, 1]	-*	-*	4 [0.49, 32.73]	0.84 [0.66, 1.07]	Low
Feeding complex														

Quality assessment							Number of patients	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Yuksel et al, 2014)	Case-control	Serious[e]	None	Serious[f]	Serious[b]	None	71	0.44 [0.28, 0.6]	0.66 [0.47, 0.81]	-*	-*	1.27 [0.7, 2.3]	0.86 [0.59, 1.25]	Very Low

[a] *Presence of GORD based on clinical judgement*
[b] *Wide confidence intervals covering categories from low to high.*
[c] *Control group not tested for reflux symptoms*
[d] *Based on retrospective review of medical notes. Based on recorded symptoms rather than all symptoms that were present.*
[e] *Retrospective chart review*
[f] *All children had Otitis Media*
[*] *Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated.*
[**] *Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.6 Otitis media

Table 14: GRADE findings for evaluation of otitis media for identifying GORD

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Presence of otitis media for identifying GORD														
1 (El-Serag et al, 2001)	Retrospective case-control	Very serious[a]	None	None	None	None	9900	0.1 [0.07, 0.13]	0.8 [0.79, 0.8]	-*	-*	0.49 [0.37, 0.66]	1.13 [1.09, 1.17]	Low
1 (Kotsis et al, 2009) – Serious OM vs None	Prospective cohort	None	None	None	Serious[b]	None[c]	109	0.32 [0.2, 0.45]	0.88 [0.75, 0.95]	0.76 [0.55, 0.91]	0.51 [0.4, 0.62]	2.59 [1.12, 5.97]	0.78 [0.64, 0.95]	Moderate
1 (Kotsis et al, 2009) – Any OM vs None	Prospective cohort	None	None	None	Serious[b]	None[d]	187	0.22 [0.15, 0.3]	0.88 [0.75, 0.95]	0.83 [0.67, 0.94]	0.28 [0.21, 0.36]	1.78 [0.79, 4.01]	0.89 [0.78, 1.02]	Very low
GOR for identifying OM														
1 (Aydin et al, 2011)	Case-control	Serious[e]	None	Serious[f]	Very serious[b]	None	40	0.3 [0.12, 0.54]	0.85 [0.62, 0.97]	-*	-*	2 [0.58, 6.91]	0.82 [0.59, 1.16]	Very low

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (O'Reilly et al, 2008)	Case-control	Very serious[g]	None	None	None	None	509	0.2 [0.17, 0.24]	0.98 [0.92, 1]	-*	-*	12.95 [1.84, 91.23]	0.81 [0.77, 0.85]	Low

[a] *Retrospective and based on computer records*
[b] *Outcome cover several categories for several items*
[c] *Serious OM vs None*
[d] *Any OM vs none*
[e] *Small sample size*
[f] *adenoid hypertrophy*
[g] *Identification of GORD based on medical records*
[*] *Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated.*
[**] *Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.7 Lower respiratory tract infection

Table 15: GRADE findings for evaluation of pneumonia

Quality assessment							Number of children	Measure of diagnostic accuracy						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Ever had pneumonia used to identify presence of GOR/D														
1 (El-Serag et al, 2001)	Retrospective cohort	Very serious[a]	None	None	None	None	9900	0.06 [0.05, 0.07]	0.98 [0.97, 0.98]	0.41 [0.35, 0.47]	0.81 [0.8, 0.81]	2.76 [2.2, 3.45]	0.96 [0.95, 0.97]	Low
1 (Orenstein et al, 1996)	Case-control	Serious[b]	None	None	None	None	135	0.09 [0.02, 0.23]	1 [0.96, 1]	-*	-*	-	0.91 [0.83, 1.01]	Moderate
1 (Salvatore et al, 2005)	Cohort	None	None	None	Serious[c]	None	99	0.2 [0.06, 0.44]	0.96 [0.89, 0.99]	0.57 [0.18, 0.9]	0.82 [0.73, 0.89]	5.13 [1.25, 21.11]	0.83 [0.67, 1.04]	Moderate
Aspiration Pneumonia used to identify presence of GOR/D														
1 (Assadamongkol et al, 1993)	Retrospective cohort	Very Serious[d]	None	None	None	None	55	0.5 [0.3, 0.7]	0.31 [0.15, 0.51]	0.39 [0.23, 0.58]	0.41 [0.21, 0.64]	0.73 [0.46, 1.14]	1.61 [0.83, 3.13]	Low
Recurrent Pneumonia used to identify presence of GOR/D														
1 (Assadamongkol et al, 1993)	Retrospective cohort	Very Serious[d]	None	None	None	None	55	0.08 [0.01, 0.25]	0.97 [0.82, 1]	0.67 [0.09, 0.99]	0.54 [0.39, 0.68]	2.23 [0.21, 23.19]	0.96 [0.84, 1.09]	Low

Quality assessment							Number of children	Measure of diagnostic accuracy						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Mazliah et al, 2000)	Cross-sectional survey	Serious[e]	None	None	Serious[c]	None	44	0.19 [0.07, 0.37]	0.62 [0.32, 0.86]	0.55 [0.23, 0.83]	0.24 [0.11, 0.42]	0.5 [0.19, 1.36]	1.31 [0.82, 2.08]	Low
Bronchiectasis with or without collapse used to identify presence of GOR/D														
1 (El-Serag et al, 2001)	Retrospective cohort	Very serious[a]	None	None	None	None	9900	0.24 [0.08, 0.47]	0.67 [0.56, 0.77]	0.16 [0.05, 0.34]	0.77 [0.65, 0.86]	0.72 [0.32, 1.65]	1.14 [0.85, 1.51]	Low

[a] *Retrospective and based on computer records*
[b] *Classification of control group was based on not being treated for GORD.*
[c] *Wide confidence intervals covering categories from low to high.*
[d] *Retrospective chart review*
[e] *Method of confirming GORD varied between children*
[*] *Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated*

J.1.2.8 Faltering growth

Table 16: GRADE findings for evaluation of faltering growth.

Quality assessment							Number of patients	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Weight gain problems														
1 (Orenstein et al, 1996)	Case-control	Serious[c]	None	None	Serious[b]	None	135	0.26 [0.12, 0.43]	1 [0.96, 1]	-*	-*	∞	0.74 [0.61, 0.9]	Low
1 (Salvatore et al, 2005)	Cohort	None	None	None	Serious[b]	None	99	0.19 [0.05, 0.42]	0.83 [0.73, 0.91]	0.24 [0.07, 0.5]	0.79 [0.69, 0.87]	1.14 [0.42, 3.14]	0.97 [0.77, 1.22]	Moderate
Failure to thrive														
1 (Carret al, 2000)	Retrospective cohort	Very serious[a]	None	None	Very serious[b]	None	295	0.09 [0.05, 0.14]	1 [0.96, 1]	1 [0.82, 1]	0.29 [0.24, 0.35]	∞	0.91 [0.87, 0.95]	Very low
1 (Tolia et al, 2003)	Retrospective cohort	Very serious[a]	None	None	None	None	342	0.16 [0.11, 0.23]	0.9 [0.84, 0.94]	0.62 [0.47, 0.76]	0.51 [0.45, 0.57]	1.61 [0.92, 2.83]	0.93 [0.86, 1.01]	Low
1 (Costa et al, 2005)	Cross-sectional survey	Very serious[d]	None	None	None	None	797	0.3 [0.21, 0.41]	0.96 [0.94, 0.97]	0.49 [0.35, 0.63]	0.92 [0.89, 0.94]	7.67 [4.74, 12.4]	0.73 [0.63, 0.83]	Low

Quality assessment							Number of patients	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Yuksel et al, 2014)	Case-control	Serious[e]	None	Serious[f]	Serious[g]	None	71	0.44 [0.28, 0.6]	0.66 [0.47, 0.81]	-*	-*	1.27 [0.7, 2.3]	0.86 [0.59, 1.25]	

[a] *Based on retrospective review of medical notes.*
[b] *Wide confidence intervals covering categories from low to high.*
[c] *Control group not tested for reflux symptoms.*
[d] *Classification of cases and controls based on Rome II criteria for adults and not diagnostic tests.*
[e] *Retrospective chart review*
[f] *All children had Otitis Media*
[g] *Wide confidence intervals covering categories from low to moderate.*
[*] *Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated.*
[**] *Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.9 Asthma

Table 17: GRADE findings for evaluation of diagnostic value of asthma for identifying children with GORD

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Using presence of asthma to identify GORD														
1 (El-Serag et al, 2001)	Retrospective cohort	Very serious[a]	None	None	None	None	9900	0.13 [0.12, 0.15]	0.93 [0.93, 0.94]	-*	-*	1.95 [1.7, 2.24]	0.93 [0.91, 0.95]	Low
1 (Ruigomez et al, 2010)	Retrospective cohort	Very serious[b]	None	None	None	None	6677	0.25 [0.23, 0.27]	0.81 [0.8, 0.82]	-*	-*	1.31 [1.19, 1.45]	0.93 [0.9, 0.95]	Low
Using presence of GORD to identify Asthma														
1 (Petersen et al, 1989)	Case-control	Serious[c]	None	None	Very serious[d]	None	39	0.33 [0.16, 0.55]	0.93 [0.68, 1]	-*	-*	5 [0.69, 36.08]	0.71 [0.52, 0.98]	Very Low
1 (Debley et al, 2006)	Case-control	Serious[e]	None	None	None	None	2397	0.19 [0.15, 0.24]	0.97 [0.97, 0.98]	-*	-*	7.65 [5.18, 11.31]	0.83 [0.78, 0.88]	Moderate
1 (Stordal et al, 2006)	Case-control	Serious[f]	None	None	None	None	1136	0.2 [0.17, 0.23]	0.92 [0.88, 0.95]	-*	-*	2.37 [1.55, 3.61]	0.88 [0.83, 0.92]	Moderate

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Chopra et al, 1995)	Case-control	Serious[g]	None	None	Very serious[d]	None	90	0.39 [0.28, 0.5]	1 [0.69, 1]	-*	-*	∞	0.61 [0.51, 0.73]	Very Low
1 (Gustafsson et al, 1990)	Case-control	Very Serious[h]	None	None	Serious[d]	None	69	0.5 [0.34, 0.66]	0.85 [0.66, 0.96]	-*	-*	3.38 [1.3, 8.76]	0.59 [0.42, 0.83]	Very Low

[a] *Retrospective and based on computer records*
[b] *Retrospective and based on computer records. On 15.7% of GORD group had formal test.*
[c] *Definition of GORD was based on barium meal only.*
[d] *Wide confidence intervals means results cover several categories*
[e] *Definition of GORD was based on a questionnaire.*
[f] *Definition of GORD was based on a questionnaire.*
[g] *GORD based on scintiscan. Control group was very small sample size.*
[h] *Results are based on two separate studies using the same methodology. Cases include people age 18 and over.*
[*] *Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated.*
[**] *Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.10 Chronic cough

Table 18: GRADE findings for evaluation of diagnostic value of chronic cough for identifying children with GORD

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Chronic cough used to identify presence of GOR/D														
1 (Uzun et al, 2012)	Retrospective cohort	Very serious[a]	None	None	None	None	70	0.67 [0.5, 0.81]	0.32 [0.17, 0.51]	0.55 [0.4, 0.7]	0.43 [0.23, 0.66]	0.98 [0.71, 1.37]	1.03 [0.53, 2.03]	Low

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Carret al, 2000)	Retrospective cohort	Very Serious[b]	None	None	None	None	214	0.51 [0.44, 0.58]	0.59 [0.48, 0.7]	0.77 [0.69, 0.83]	0.31 [0.24, 0.39]	1.25 [0.93, 1.68]	0.83 [0.66, 1.04]	Low
1 (Chang et al, 2006)	Prospective Cohort	None	None	None	None	None	150	0.43 [0.32, 0.55]	0.51 [0.39, 0.63]	0.48 [0.36, 0.6]	0.46 [0.35, 0.57]	0.87 [0.61, 1.23]	1.13 [0.84, 1.52]	High
1 (Salvatore et al, 2005)	Prospective cohort	Serious[c]	None	None	None	None	99	0.24 [0.08, 0.47]	0.62 [0.51, 0.73]	0.15 [0.05, 0.31]	0.75 [0.63, 0.85]	0.63 [0.28, 1.43]	1.22 [0.91, 1.64]	Moderate
1 (Yuksel et al, 2014)	Case-control	Serious[d]	None	Serious[e]	None	None	71	0.54 [0.37, 0.7]	0.47 [0.29, 0.65]	0.55 [0.38, 0.71]	0.45 [0.28, 0.64]	1.01 [0.66, 1.57]	0.98 [0.6, 1.62]	Low

[a] *Based on presenting symptoms rather than questionnaire, so not all children will have been asked about same symptoms*
[b] *Retrospective chart review*
[c] *Chronic cough based on a single question involving parental assessment*
[d] *Retrospective chart review*
[e] *All children had Otitis Media*
** *Calculated by the NCC technical team based on figures presented within the studies*

J.1.2.11 Dental erosion

Table 19: GRADE findings for evaluation of dental erosion to identify GORD

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
Presence of any type of dental erosion compared to no dental erosion used to identify presence of GOR/D														
1 (Linnett et al, 2002)	Case-control	Serious[a]	None	None	None	None	104	0.46 [0.32, 0.61]	0.6 [0.39, 0.73]	-*	-*	1.14 [0.73, 1.78]	0.9 [0.65, 1.26]	Moderate
1 (Ersin et al, 2006)	Case-control	Serious[b]	None	None	Serious[c]	None	80	0.76 [0.6, 0.89]	0.76 [0.62, 0.88]	-*	-*	3.21 [1.81, 5.66]	0.31 [0.17, 0.56]	Low
1 (Shaw et al, 1998)	Case-control	Serious[d]	None	None	Very serious[c]	None	41	0.81 [0.58, 0.95]	0.85 [0.58, 0.97]	-*	-*	5.4 [1.86, 15.64]	0.22 [0.09, 0.55]	Very low
1 (Wild et al, 2011)	Case-control	Serious[e]	None	None	Serious[c]	None	72	0.76 [0.63, 0.86]	0.43 [0.12, 0.71]	-*	-*	1.33 [0.82, 2.14]	0.56 [0.26, 1.2]	Low
1 (Gonda-Domin et al, 2013)	Case-contorl	None	None	None	None	None	114	0.67 [0.53, 0.79]	0.74 [0.6, 0.84]	-*	-*	2.53 [1.58, 4.06]	0.45 [0.3, 0.67]	High

Quality assessment							Number of children	Measure of diagnostic accuracy**						Quality
Number. of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations		Sensitivity	Specificity	Positive predictive value	Negative predictive value	Positive likelihood ratio	Negative likelihood ratio	
1 (Farahmand et al,	Case-control	Very serious[i]	None	None	Very serious[c]	None	64	0.98 [0.9, 1]	0.81 [0.69, 0.9]	-*	-*	5.18 [3.04, 8.82]	0.02 [0, 0.16]	Very low
Presence of any type of dental erosion compared to no dental erosion in children with cerebral palsy used to identify presence of GOR/D														
1 (Guare et al, 2012)	Case-control	Serious[f]	None	None	Very serious[c]	None	46	0.9 [0.68, 0.99]	0.81 [0.72, 0.93]	-*	-*	4.68 [2.1, 10.43]	0.12 [0.03, 0.47]	Very low
1 (Shaw et al, 1998)	Case-control	Very serious[g]	None	None	Very serious[c]	None	21	0.75 [0.43, 0.95]	0.67 [0.3, 0.93]	-*	-*	2.25 [0.84, 6]	0.38 [0.13, 1.11]	Very low
Presence of GORD compared to no GORD as a cause of dental problems in children with cerebral palsy used to identify presence of GOR/D														
1 (Polat et al, 2013)	Case-control	Very serious[h]	None	None	Very serious[c]	None	37	0.84 [0.6, 0.97]	0.72 [0.54, 0.9]	-*	-*	3.03 [1.4, 6.55]	0.22 [0.07, 0.64]	Very low
Localised vs generalised erosions														
1 (Farahmand et al,	Case-control	Very serious[i]	None	None	Very serious[c]	None	64	0.34 [0.22, 0.48]	0.55 [0.23, 0.83]	-*	-*	0.75 [0.35, 1.58]	1.21 [0.68, 2.15]	Very low

[a] *Control group were not assessed for GORD*
[b] *Unclear how presence of GER was determined in case and control groups*
[c] *Outcome cover several categories for several items*
[d] *Unclear how GER was determined in all children. Children referred to a tertiary dental unit.*
[e] *Unclear if analysis was undertaken on all children or only those who had pH monitoring*
[f] *Small sample size*
[g] *Unclear how GER was determined in all children. Small sample size.*
[h] *Analysis relates to GORD as a risk-factor for dental erosion rather than dental erosion as a marker of GORD*

[l] *excluded children where other sources of erosion were identified*
[*] *Predicative values cannot be calculated from case-control studies as the true prevalence cannot be calculated.*
[**] *Calculated by the NCC technical team based on figures presented within the studies*

J.1.3 Risk Factors

J.1.3.1 Chronic Lung Disease

Table 20: GRADE findings for the association between chronic lung disease and risk of developing GORD

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Bronchopulmonary Dysplasia											
Prevalence and odds ratio for bronchopulmonary dysplasia in children with and without GER[a]/GERD[b]											
1 (Akinola, 2004)	Retrospective cohort	Very serious[c,d]	None	Serious[e]	Very serious[f]	None	64/87 (74%)	38/50 (76%)	OR: 0.88 (0.39 to 1.97)[g]	-	Very low
1 (Mezzacappa, 2008)	Retrospective case-control	Very serious[c,h,i]	None	Serious[j]	Very serious[f]	None	33/87 (38%)	44/87 (51%)	Adjusted OR: 0.89 (0.46 to 1.75)[k]	-	Very low
Cystic Fibrosis											
Prevalence and odds ratio for cystic fibrosis in children with and without GERD[l,m]											
1 (El-Serag, 2001)	Retrospective case-control	Very serious[c,n]	None	No serious	No serious	None	NR/1980	NR/7920	Adjusted OR: 2.89 (1.97 to 4.25)[o]	-	Low
1 (Ruigomez, 2010)	Retrospective cohort	Very serious[c,p]	None	No serious	Very serious[f]	None	5/1700 (0.3%)	2/4977 (0.04%)	Adjusted OR: 3.3 (0.6 to 18.1)[q]	-	Very low
Bronchiectasis											
Prevalence and odds ratio for bronchiectasis (with or without collapse) in children with and without GERD[l]											
1 (El-Serag, 2001)	Retrospectivecase-control	Very serious[c,n]	None	No serious	Serious[f]	None	NR/1980	NR/7920	Adjusted OR: 2.28 (1.14 to 4.57)[o]	-	Very low
Chronic Lung Disease											
Prevalence and odds ratio for chronic lung disease of prematurity in children with and without GER[r]											
1	Retros	Serio	None	Serio	Seri	None	NR	NR	Adjust	-	Very

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
(Fuloria, 2000)	pective case-control	us[c]		us[s]	ous[f]				ed OR: 2.1 (1.1 to 3.5)[t]		low
Severe Chronic Lung Disease											
Prevalence and odds ratio for severe chronic lung disease in children with and without GER[a]											
1 (Akinola, 2004)	Retrospective cohort	Very serious[c,d]	None	Serious[e]	Very serious[f]	None	46/87 (53%)	30/49 (61%)	OR: 0.71 (0.35 to 1.45)[g]	-	Very low

[a] *Akinola 2004: diagnostic criteria for GER - 18 to 24 hour esophageal pH monitoring, infants were identified as positive for GER if there was ≥10% acid reflux with the glucose water feed or ≥5% acid reflux with formula or breast milk*
[b] *Mezzacappa 2008: diagnostic criteria for GERD - prolonged distal intra-esophageal pH monitoring, reflux index ≥10%*
[c] *Retrospective study design*
[d] *Unadjusted ORs*
[e] *Infants less than 32 weeks gestational age admitted to the neonatal intensive care unit*
[f] *Confidence interval spans multiple interpretations*
[g] *NCC-WCH calculation*
[h] *No details of how bronchopulmonary dysplasia was defined/diagnosed*
[i] *Not explained which pH test was selected for inclusion as there seems to be more than one per child (235pH studies in 193 infants)*
[j] *Birthweight <2000g and gestational age ≤37 weeks*
[k] *OR adjusted for birthweight and postconceptional age at time of pH study*
[l] *El-Serag 2001: diagnostic criteria for GERD – subjects identified from electronic medical records, based on ICD-9 coding of GERD (530.81, 530.10, 530.11, 530.19, 530.3)*
[m] *Ruigomez 2010: diagnostic criteria for GERD - identified by Read codes for gastro-oesophageal reflux, reflux esophagitis, esophageal inflammation and heartburn. Non-specific symptoms such as epigastric pain to idenify cases was not used unless they were recorded alongside reflux symptoms.*
[n] *Both the risk factor and outcome based on reliability of coding in medical records*
[o] *OR adjusted for age, gender and ethnicity*
[p] *Only 15.3% of GERD cohort had a record of a formal diagnostic test being undertaken and none of the children in the control cohort had been tested for GER*
[q] *OR adjusted for age, sex, year of diagnosis, visits to primary care physician in the previous year*
[r] *Fuloria 2000: diagnostic criteria for GER - defined as either treatment with anti-reflux medications (metaclopramide, bethanecol, cisparide, cimetidine or ranitidine) or a positive test for GER. Tests for GER included esophageal pH probe, upper gastrointestinal contrast studies and microscopic examination of tracheal aspirates for lipid laden macrophages. Tests for GER were performed and treatment was initiated at the discretion of the attending neonatologist.*
[s] *Very low birth weight premature infants*
[t] *OR adjusted for gestational age, gender, race, days on assisted ventilation and days of hospitalisation*

J.1.3.2 Neurodevelopmental disorders

Table 21: GRADE findings for the association between neurodevelopmental disorders and risk of developing GORD

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Neurodevelopmental disorders											
Prevalence and odds ratio for cerebral palsy in children with and without GER[a]											
1 (Fuloria, 2000)	Retrospective case-control	Very serious[b,c]	None	Serious[d]	Very serious[e]	None	15/111 (14%)	31/235 (13%)	OR: 1.03 (0.53 to 1.99)[f]	-	Very low
Prevalence and odds ratio for neurological disabilities[g] in children with and without GERD[h]											
1 (Ruigomez, 2010)	Retrospective cohort	Very serious[b,i]	None	None	None	None	107/1700 (6.3%)	72/4977 (1.4%)	Adjusted OR: 3.40 (2.50 to 4.70)[j]	-	Low
Prevalence and odds ratio for CNS disease[k] in children with and without GER[l] – total population											
1 (Halpern et al., 1991)	Retrospective review	Very serious[b,c]	None	None	Very serious[e]	None	101/463 (21.8%)	31/149 (20.8%)	OR: 1.06 (0.68 to 1.67)[c]	-	Very low
Prevalence and odds ratio for CNS disease[k] in children with and without GERl – subjects > 1 year of age											
1 (Halpern et al., 1991)	Retrospective review	Very serious[b,c]	None	None	Serious[e]	None	31/69 (44.9%)	14/57 (24.6%)	OR: 2.51 (1.16 to 5.4)[c]	-	Very low
Prevalence and odds ratio for CNS disease[k] in children with and without GER[l] – subjects < 1 year of age											
1 (Halpern et al., 1991)	Retrospective review	Very serious[b,c]	None	None	Very serious[e]	None	70/394 (17.8%)	17/92 (18.5%)	OR: 0.95 (0.53 to 1.71)[c]	-	Very low

[a] *Fuloria 2000: diagnostic criteria for GER - defined as either treatment with anti-reflux medications (metaclopramide, bethanecol, cisparide, cimetidine or ranitidine) or a positive test for GER. Tests for GER included esophageal pH probe, upper gastrointestinal contrast studies and microscopic examination of tracheal aspirates for lipid laden macrophages. Tests for GER were performed and treatment was initiated at the discretion of the attending neonatologist.*
[b] *Retrospective study design*
[c] *Calculated by NCC-WCH, therefore unadjusted odds ratios*
[d] *Very low birth weight premature infants with chronic lung disease*
[e] *Confidence interval spans three possible interpretations*
[f] *Calculated by NCC-WCH technical team based on data reported in the article*

[g] *Included cerebral palsy, neurological syndromes with motor component, chromosomal anomalies, congenital central nervous system anomalies, mental retardation and delayed development, central nervous system neoplasm, and neurological disorders due to neoplasm, trauma, encephalitis and extreme prematurity*
[h] *Ruigomez 2010: diagnostic criteria for GERD - identified by Read codes for gastro-oesophageal reflux, reflux esophagitis, esophageal inflammation and heartburn. Non-specific symptoms such as epigastric pain to identify cases was not used unless they were recorded alongside reflux symptoms.*
[i] *Only 15.3% of GERD cohort had a record of a formal diagnostic test being undertaken and none of the children in the control group had been tested for GERD*
[j] *OR adjusted for age, sex, year of diagnosis, visits to primary care physician in the previous year*
[k] *Includes mental-motor retardation: including cerebral palsy, developmental delay and mental retardation, seizure disorder, hydrocephalus, microcephaly, intracerebral haemorrhage, cortical blindness, abnormal head CT scan only, abnormal EEG without seizures, porencephalic cyst, spastic quadriplegia, cerebral dysgenesis, meningomyelocele, subarachnoid cyst, abnormal brainstem auditory evoked potential only, multiple CNS diseases, syndromes with CNS involvement.*
[l] *Halpern 1991: diagnostic criteria for GER: initial evaluation included an extensive history and physical examination, barium oesophagram, upper gastrointestinal series and 18 to 24 hour esophageal pH monitoring. Documentation of GER by an abnormal pH score derived from 18 to 24 hour esophageal pH monitoring.*

J.1.3.3 Prematurity

Table 22: GRADE findings for the association between prematurity and risk of developing GORD

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Prematurity											
Prevalence and odds ratio for gestational age ≤32 weeks (versus 37 to 41 weeks) in children with and without subsequent oesophagitis[a] at the following ages:											
≤9 years											
1 (Forssell, 2012)	Retrospective case-control	Serious[b,c]	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 6.82 (4.65 to 10.03)[d]	-	Moderate
10 to 19 years											
1 (Forssell, 2012)	Retrospective case-control	Serious[b,c]	No serious	No serious	Serious[e]	None	NR	NR	Adjusted OR: 2.09 (1.18 to 3.70)[d]	-	Low
Prevalence and odds ratio for gestational age 33 to 36 weeks (versus 37 to 41 weeks) in children with and without oesophagitis[a] at the following ages:											
≤9 years											
1 (Forssell, 2012)	Retrospective case-control	Serious[b,c]	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 1.75 (1.42 to 2.14)[d]	-	Moderate
10 to 19 years											
1 (Forssell, 2012)	Retrospective case-control	Serious[b,c]	No serious	No serious	Serious[e]	None	NR	NR	Adjusted OR: 1.41 (1.10 to 1.80)[d]	-	Low
Prevalence and odds ratio for prematurity (25 to 36 weeks of gestation) in children with and without GER[f]											
1 (Kohelet, 2004)	Retrospective cohort	Very serious[b,g]	No serious	No serious	Very serious[e]	None	18/62 (29%)	27/72 (38%)	OR: 0.68 (0.33 to 1.41)[h]	-	Very low
Prevalence and odds ratio for prematurity (<37weeks gestation) in children with and without GOR[i]											

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
1 (Deurloo, 2004)	Retrospective cohort	Very serious[b,g]	No serious	Serious[j]	Serious[e]	None	32/73 (44%)	44/124 (35%)	OR: 1.42 (0.79 to 2.56)[h]	-	Very low

[a] *Forssell 2012: diagnostic criteria for esophagitis - cases of endoscopically verified esophagitis were ascertained through the Patient Register by combining the discharge diagnoses for esophagitis and the procedure codes for upper endoscopy. Confirmation of the diagnosis was based on the explicit diagnosis of esophagitis, combined with the described macroscopic findings at endoscopy that were found in the charts.*
[b] *Retrospective study design*
[c] *Oesophagitis based on unverified clinical coding criteria*
[d] *OR adjusted for birth weight for gestational age, maternal age and birth order*
[e] *Confidence interval spans multiple interpretations*
[f] *Kohelet 2004: diagnostic criteria for GER - 24-hour distal esophageal pH monitoring. Reflux was considered pathologic if the proportion of total time with pH <4 during a 24-hour period exceeded 4%.*
[g] *Unadjusted odds ratios*
[h] *Calculated by NCC-WCH technical team based on data reported in the article*
[i] *Deurloo 2004: diagnostic criteria for GOR - Diagnosed either by clinical symptoms (n=30) or by 24 hour pH measurement (n=43).*
[j] *Infants with oesophageal atresia*

J.1.3.4 Surgical or congenital disorders

Table 23: GRADE findings for the association between surgical/congenital disorders (hiatal hernia, diaphragmatic hernia, oesophageal atresia) and risk of developing GORD

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	GORD	NO GORD	Relative (95% CI)	Absolute (95% CI)	Quality
Hiatal Hernia with Reflux											
Prevalence and odds ratio for hiatal hernia with reflux in children with and without gastrointestinal symptoms[a]											
1 (Abrahams, 1970)	Prospective case-control	Serious[b]	None	Serious[c]	None	None	8/16 (50%)	5/63 (8%)	OR: 11.6 (3.04 to 44.29)[d]	-	Low
Hiatal Hernia											
Prevalence and odds ratio for hiatal hernia in children with and without erosive oesophagitis[e]											
1 (Steward, 1993)	Prospective cohort	Serious[b]	None	No serious	Serious[f]	None	12/20 (60%)	25/75 (33%)	OR: 3.00 (1.09 to 8.28)[d]	-	Low
Hiatal And Diaphragmatic Hernia											
Prevalence and odds ratio for hiatus hernia (congenital and acquired hiatus and diaphragmatic hernia) in children with and without GERD[g]											
1 (Ruigomez, 2010)	Retrospective cohort	Very serious[h,i]	None	None	None	None	13/1700 (0.8%)	6/4977 (0.1%)	Adjusted OR: 7.4 (2.7 to 20.3)[j]	-	Low
Oesophageal Atresia											
Prevalence and odds ratio for congenital oesophageal disorders (oesophageal atresia, stenosis and traque-oesophageal fistula) in children with and without GERD[g]											
1 (Ruigomez, 2010)	Retrospective cohort	Very serious[h,i]	None	None	None	None	8/1700 (0.5%)	5/4977 (0.1%)	Adjusted OR: 4.3 (1.3 to 14.1)[j]	-	Low

[a] *Abrahams 1970: diagnostic criteria for gastrointestinal symptoms - complaints referable to the gastro-intestinal tract (such as vomiting and haematemesis). Each patient was examined fluoroscopically, after the ingestion of 4 to 6 ozs of barium, in the supine position and then prone to see whether a hernia or reflux became visible.*
[b] *Unadjusted odds ratios*
[c] *All children with severe physical disability (cerebral palsy)*
[d] *Calculated by NCC-WCH technical team based on data reported in the article*
[e] *Steward 1993: diagnostic criteria for erosive oesophagitis – endoscopy, oesophagitis was defined by the demonstration of friability, erosions or ulceration of the mucosa*
[f] *Confidence interval spans multiple interpretations*
[g] *Ruigomez 2010: diagnostic criteria for GERD - identified by Read codes for gastro-oesophageal reflux, reflux esophagitis, esophageal inflammation and heartburn. Non-specific symptoms such as epigastric pain to identify cases was not used unless they were recorded alongside reflux symptoms.*
[h] *Retrospective study design*

[i] *Only 15.3% of GERD cohort had a record of a formal diagnostic test being undertaken and none of the children in the control cohort had been tested for GER*
[j] *OR adjusted for age, sex, year of diagnosis and visits to primary care physician in the previous year*

J.1.3.5 Family history of GORD

Table 24: GRADE findings for the association between family history of GORD and risk of developing GORD

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	GORD	NO GORD	Relative (95% CI)	Absolute (95% CI)	Quality
Family History of GORD											
Prevalence and odds ratio for a family history of epigastric pain in adolescents with and without epigastric pain[a] in the following categories:											
Either mother or father has epigastric pain											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Serious[b]	None	14/52 (26.9 %)	189/963 (19.6 %)	Adjusted OR: 1.74 (0.82 to 3.69)c	-	Moderate
Both mother and father have epigastric pain											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Serious[b]	None	4/52 (7.7%)	13/963 (1.3 %)	Adjusted OR: 4.15 (0.78 to 22.2)c	-	Moderate
Prevalence and odds ratio for a family history of heartburn in adolescents with and without heartburna in the following categories:											
Either mother or father has heartburn											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Serious[b]	None	13/32 (40.6 %)	226/988 (22.9 %)	Adjusted OR: 2.47 (0.99 to 6.16)[c]	-	Moderate
Both mother and father have heartburn											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	No serious	None	6/32 (18.8 %)	42/988 (4.3 %)	Adjusted OR: 5.71 (1.62 to 20.1)[c]	-	High
Prevalence and odds ratio for a family history of acid regurgitation in adolescents with and without acid regurgitationa in the following categories:											
Either mother or father has acid regurgitation											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Serious[b]	None	15/49 (30.6 %)	147/965 (15.2 %)	Adjusted OR: 2.54 (1.16 to	-	Moderate

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
									5.60)[c]		
Both mother and father have acid regurgitation											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	No serious	None	4/49 (8.2%)	10/965 (1.0%)	Adjusted OR: 6.89 (1.32 to 35.7)[c]	-	High

[a] *Murray 2007: diagnostic criteria - both adolescents and their parents completed a questionnaire including the following questions:*
1) how often in the last 3 months have you had pain or discomfort in the place shown in the picture? (a diagram was included showing the epigastric area)
2) how often in the last 3 months have you had heartburn? (burning or ache behind the breastbone)
3) how often in the last 3 months have you got a very sour or acid tasting fluid at the back of your throat?
[b] *Confidence interval spans multiple interpretations*
[c] *OR adjusted for adolescent's age, sex, social class, household density (persons per room), BMI category, alcohol intake and smoking status. Analysis was also restricted to children living with both natural parents.*

J.1.3.6 Obesity

Table 25: GRADE findings for the association between obesity and risk of developing GORD

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Overweight											
Prevalence and odds ratio for overweight in children with and without GERD[a]											
1 (Stordal, 2006)	Prospective case control	Serious[b]	No serious	Serious[c]	Serious[d]	None	NR	NR	Adjusted OR: 1.6 (1.1 to 2.4)[e]	-	Very low
Prevalence and odds ratio for overweight in children with and without epigastric pain[f]											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Very serious[d]	None	NR	NR	Adjusted OR: 1.09 (0.49 to 2.40)[g]	-	Low
Prevalence and odds ratio for overweight in children with and without heartburn[f]											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Very serious[d]	None	NR	NR	Adjusted OR: 1.06 (0.35 to 3.21)[g]	-	Low
Prevalence and odds ratio for overweight in children with and without acid regurgitation[f]											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Very serious[d]	None	NR	NR	Adjusted OR: 1.64 (0.72 to 3.72)[g]	-	Low
Prevalence and odds ratio for overweight in children with and without GERD[h] at the following ages:											
2 to 5 years											
1 (Koebnick, 2011)	Retrospective cross sectional	Serious[i]	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 0.95 (0.85 to 1.07)[j]	-	Moderate
6 to 11 years											
1 (Koebnick, 2011)	Retrospective cross	Serious[i]	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 0.99	-	Moderate

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
	sectional								(0.87 to 1.12)[j]		
12 to 19 years											
1 (Koebnick, 2011)	Retrospective cross sectional	Serious[i]	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 1.08 (1.01 to 1.15)[j]	-	Moderate
Overweight/Obesity											
Prevalence and odds ratio for overweight/obesity in children with and without a positive reflux score[k]											
1 (Quitadamo, 2012)	Prospective cohort	Serious[l]	No serious	No serious	No serious	None	29/49 (59%)	30/104 (29%)	OR: 3.58 (1.76 to 7.28)[m]	-	Moderate
Prevalence and odds ratio for overweight/obesity in children with and without GERD[n]											
1 (Elitsur, 2009)	Retrospective chart review	Very serious[i,o]	No serious	No serious	Serious[d]	None	237/491 (48%)	108/247 (44%)	OR: 1.2 (0.88 to 1.63)[m]	-	Very low
Obesity											
Prevalence and odds ratio for obesity in children with and without a positive reflux symptom score[p]											
1 (Pashankar, 2009)	Prospective case-control	No serious	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 7.4 (1.7 to 32.5)[q]	-	High
Prevalence and odds ratio for obesity in children with and without epigastric pain[f]											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Very serious[d]	None	NR	NR	Adjusted OR: 0.84 (0.20 to 3.65)[g]	-	Low
Prevalence and odds ratio for obesity in children with and without heartburn[f]											
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Very serious[d]	None	NR	NR	Adjusted OR: 0.84 (0.11 to 6.60)[g]	-	Low
Prevalence and odds ratio for obesity in children with and without acid regurgitation[f]											

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**GORD**	**NO GORD**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
1 (Murray, 2007)	Prospective cross-sectional	No serious	No serious	No serious	Serious[d]	None	NR	NR	Adjusted OR: 3.46 (1.24 to 9.69)[g]	-	Moderate
Moderate Obesity (BMI for age ≥95th percentile or a BMI ≥30kg/m²)											
Prevalence and odds ratio for moderate obesity in children with and without GERD[h] at the following ages:											
2 to 5 years											
1 (Koebnick, 2011)	Retrospective cross sectional	Serious[i]	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 0.92 (0.80 to 1.06)[j]	-	Moderate
6 to 11 years											
1 (Koebnick, 2011)	Retrospective cross sectional	Serious[i]	No serious	No serious	Serious[d]	None	NR	NR	Adjusted OR: 1.16 (1.02 to 1.32)[j]	-	Low
12 to 19 years											
1 (Koebnick, 2011)	Retrospective cross sectional	Serious[i]	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 1.16 (1.07 to 1.25)[j]	-	Moderate
Extreme/Morbid Obesity											
Prevalence and odds ratio for extreme obesity in children with and without GERD[h] at the following ages:											
2 to 5 years											
1 (Koebnick, 2011)	Retrospective cross sectional	Serious[i]	No serious	No serious	Serious[d]	None	NR	NR	Adjusted OR: 1.26 (0.95 to 1.68)[j]	-	Low
6 to 11 years											
1 (Koebnick, 2011)	Retrospective cross	Serious[i]	No serious	No serious	Serious[d]	None	NR	NR	Adjusted OR: 1.32	-	Low

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	GORD	NO GORD	Relative (95% CI)	Absolute (95% CI)	Quality
	sectional								(1.13 to 1.56)[j]		
12 to 19 years											
1 (Koebnick, 2011)	Retrospective cross sectional	Serious[i]	No serious	No serious	No serious	None	NR	NR	Adjusted OR: 1.40 (1.28 to 1.52)[j]	-	Moderate
Prevalence and odds ratio for morbid obesity in children with and without GERD[r]											
1 (El-Serag, 2001)	Retrospective case-control	Very serious[i,s]	No serious	No serious	Serious[d]	None	NR/1980	NR/7920	Adjusted OR: 1.90 (1.17 to 3.02)[t]	-	Very low

[a] *Stordal 2007: diagnostic criteria for GERD - 7-item GERD questionnaire developed and validated by the author. GERD if 3 or more points on a questionnaire. A score of 3 or more points (positive symptom score) has a 75% sensitivity and 96% specificity for GERD defined by an abnormal pH monitoring. Overweight: BMI calculated as weight divided by height² and compared to international age-adjusted percentiles. Overweight and obesity were defined as BMI corresponding to an adult BMI above 25 and 30, respectively.*
[b] *Presence of GORD based on questionnaire rather than objective diagnostic test*
[c] *Population included children with asthma*
[d] *Confidence interval spans multiple interpretations*
[e] *Odds ratio adjusted for age, gender and asthma*
[f] *Murray 2007: diagnostic criteria - both adolescents and their parents completed a questionnaire including the following questions:*
1) how often in the last 3 months have you had pain or discomfort in the place shown in the picture? (a diagram was included showing the epigastric area)
2) how often in the last 3 months have you had heartburn? (burning or ache behind the breastbone)
3) how often in the last 3 months have you got a very sour or acid tasting fluid at the back of your throat?
BMI was calculated as body weight (kg) divided by the square of standing height (m). Adolescent BMI was categorised into normal, overweight and obese according to the age-sex specific thresholds of Cole et al).
[g] *Odds ratio adjusted for age, sex, social class, household density (persons per room), smoking, alcohol and passive smoking*
[h] *Koebnick 2011: diagnostic criteria - International Classification of Disease codes (ICD-9 code 530.81). GERD diagnosis was validated in a random subsample of about 5% of cases (n=480) by confirming diagnosis codes for GERD from physician's notes in the electronic medical record. Overweight and obesity was defined based on the sex-specific BMI for age growth charts developed by the CDC and WHO definitions for overweight and obesity in adults. Normal weight: BMI for age ≥5th and <85th percentile. Overweight: BMI for age ≥85th percentile or a BMI ≥25kg/m². Moderately obese: BMI for age ≥95th percentile or a BMI ≥30kg/m². Extremely obese: BMI for age ≥1.2 x 95th percentile or a BMI ≥35kg/m²*
[i] *Retrospective study design*
[j] *Odds ratio adjusted for sex, race and age within each age group*
[k] *Quitadamo 2012: diagnostic criteria for positive reflux score- during the clinic visit, children's esophageal symptoms (heartburn, epigastric pain, vomiting and regurgitation, irritability with meals, dysphagia and/or odynophagia, respiratory symptoms and hematemesis) during the preceding 2 months were recorded using a standardized questionnaire. The severity and frequency of symptoms were classified into different grades based on a scale used in previous studies. A score for each symptom and a total symptom score were calculated. Overweight/obesity: height, weight, BMI and waist circumference were determined for each participant. Based on the Institute of Medicine definitions, subjects were classified according to BMI as underweight - BMI <5th percentile, normal weight - BMI 5th to 85th percentile, overweight - BMI 85th to 95th percentile and obese - BMI >95th percentile and according to waist circumference in children with waist circumference <75th percentile, from 75th to 90th percentile and >90th percentile*

[l] *Positive reflux score not defined*
[m] *NCC-WCH calculation*
[n] *Elitsur 2009: diagnostic criteria for GERD – histology, the histological reports were based on assessment of at least 3 biopsies obtained from the distal esophagus. BMI status was defined as follows: normal weight - BMI <85th percentile, overweight - BMI between 85th and 95th percentiles, obese - BMI >95th percentile*
[o] *Unadjusted odds ratios*
[p] *Pashakanar 2009 diagnostic criteria: All children were interviewed in person using a standard questionnaire (completed by parents if child younger than 10 years). The questionnaire consists of a history of any sickness in the last 2 weeks and 5 symptoms experienced over the last week (vomiting, nausea, heartburn, regurgitation and dysphagia). A score was given for each symptom and a validated total score of 3 or more was considered a positive reflux symptom score. Obesity: weight and height were measured by experienced nursing assistants. BMI calculated as weight divided by height². Obesity defined as BMI greater than 95th percentile for age and sex on growth charts from the Centre for Disease Control*
[q] *Odds ratio was adjusted for age, sex, race and caffeine exposure.*
[r] *El-Serag 2001: diagnostic criteria for GERD - based on ICD-9 coding of GERD (530.81, 530.10, 530.11, 530.19, 530.3). Morbid obesity diagnosed according to ICD-9 codes.*
[s] *Both the risk factor and outcome based on reliability of coding in medical records*
[t] *Odds ratio adjusted for age, gender and ethnicity*

J.1.4 Indications for investigation and treatment

None

J.2 Initial management of GOR and GORD

J.2.1 Infant Positioning

Table 26: GRADE findings for comparison of prone with supine positioning

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Prone**	**Supine**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux index (% of time with pH < 4.0)											
1 (Bhat et al 2007)	RCT – crossover	Very serious[a, b, c,d]	NA	Serious	Non[e]	Yes[f]	n=21 Median (range): 0 (0 to 11.4)	n=21 Median (range): 3 (0 to 15.4)	NA	p=0.002	Very low
1 (Tobin et al 1997)	RCT – crossover	Serious[c,d]	NA	Non[e]	Non[e]	None	n=24 Mean (Standard deviation [SD]): 6.72 (5.2)	n=24 Mean (SD): 15.33 (11.4)	Mean Difference [MD]: -8.00 (-12.83 to -3.17)[g]	p <0.05	Moderate

[a] *Method of randomisation not reported*
[b] *Unclear whether there was adequate concealment of allocation*
[c] *Unclear whether investigators were blinded to intervention*
[d] *Unclear whether investigators were blinded to confounding factors*
[e] *12/21 subjects were oxygen dependent and had or subsequently fulfilled the diagnosis of BPD (oxygen dependency beyond 36 weeks postmenstrual age)*
[f] *Infants born premature*
[g] *Calculated by NCC-WCH technical team based on data reported in the article*

Table 27: GRADE findings for comparison of prone head elevated (at 30 to 45 degrees) positioning in harness with infant seat elevated at 60 degrees

Quality assessment							Number of infants		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Prone head elevate position in harness	Infant seat	Relative (95% CI)	Absolute (95% CI)	Quality
Reflux index (% of time with pH <4.0)											
1 (Orenstein et al 1983[a])	RCT – crossover	Serious[a,b,c]	None	None	None	None	n=15 Mean (SD): 7.9 (8.9)	n=15 Mean (SD): 37.4 (24)	MD: -29.50 (-42.46 to -16.54)[d]	p <0.001	Moderate
Number of episodes with pH <4.0											
1 (Orenstein et al 1983[a])	RCT – crossover	Serious[a,b,c]	None	None	None	None	n=15 Mean (SD): 5.2 (4.3)	n=15 Mean (SD): 19.6 (13.6)	MD: -14.40 (-21.59 to -7.21)[d]	p < 0.001	Moderate
Number of such episodes lasting longer than 5 minutes											
1 (Orenstein et al 1983[a])	RCT – crossover	Serious[a,b,c]	None	None	Serious[e]	None	n=15 Mean (SD): 0.6 (0.8)	n=15 Mean (SD): 1.9 (2.3)	MD: -1.30 (-2.54 to -0.06)[d]	p<0.05	Low
Duration of the longest episode in each 2 hour postprandial period											
1 (Orenstein et al 1983[a])	RCT – crossover	Serious[a,b,c]	None	None	Serious[e]	None	n=15 Mean (SD): 5.0 (6.6)	n=15 Mean (SD): 13.1 (19.4)	MD: -8.10 (-18.45 to 2.25)[d]	p<0.05	Low

[a] *Unclear whether there was adequate concealment of allocation*
[b] *Unclear whether investigators were blinded to intervention*
[c] *Unclear whether investigators were blinded to confounding factors*
[d] *Calculated by NCC-WCH technical team based on data reported in the article*
[e] *Confidence interval of standardised mean difference crosses 2 zones (wide confidence interval)*

Table 28: GRADE findings for comparison of head elevated prone positioning with flat prone positioning

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Head elevated prone**	**Flat prone**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux index (% of time with pH <4.0)											
1 (Orenstein et al 1990)	RCT – crossover	Serious[a,b]	None	None	Serious[c]	None	n= 90 Mean (SD): 27.8 (30.4)	n= 90 Mean (SD): 34.6 (31.3)	MD: -6.80 (-15.81 to 2.21)[d]	p= NS[e]	Low
Number of episodes with pH <4.0											
1 (Orenstein et al 1990)	RCT – crossover	Serious[a,b]	None	None	Serious[c]	None	n= 90 Mean (SD): 6.2 (5.7)	n= 90 Mean (SD): 7.8 (7.6)	MD: -1.60 (-3.56 to 0.36)[d]	p= NS[e]	Low
Mean duration of reflux episodes											
1 (Orenstein et al 1990)	RCT – crossover	Serious[a,b]	None	None	None	None	n= 90 Mean (SD): 6.1 (9.5)	n= 90 Mean (SD): 6.2 (8.5)	MD: -0.10 (-.2.74 to 2.54)[d]	p= NS[e]	Moderate
Number of reflux episodes lasting longer than 5 minutes											
1 (Orenstein et al 1990)	RCT – crossover	Serious[a,b]	None	None	None	None	n= 90 Mean (SD): 1.3 (1.9)	n= 90 Mean (SD): 1.5 (1.9)	MD: -0.20 (-0.75 to 0.35)d	p= NS[e]	Moderate
Duration of the longest reflux episode											
1 (Orenstein et al 1990)	RCT - crossover	Serious[a,b]	None	None	None	None	n= 90 Mean (SD): 17.1 (22.8)	n= 90 Mean (SD): 17.9 (20.9)	MD: -0.80 (-7.18 to 5.58)[d]	p= NS[e]	Moderate

NS – not significant
[a] *Unclear whether there was adequate concealment of allocation*
[b] *Unclear whether investigators were blinded to confounding factors*
[c] *Confidence interval of standardised mean difference crosses 2 zones (wide confidence interval)*

[d] *Calculated by NCC-WCH technical team based on data reported in the article*
[e] *Significance defined as p<0.05*

Table 29: GRADE findings for comparison of infant seat elevated at 60 degrees with horizontal prone positioning

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Infant seat**	**Horizontal prone**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux index (% of time with pH <4.0)											
1 (Orenstein et al 1983[b])	RCT – crossover	Serious[a,b,c]	None	None	Serious[d]	None	n= 9 Mean (SD): 28.2 (19.2)	n= 9 Mean (SD): 12.8 (11.1)	MD: 15.00 (0.66 to 29.34)[e]	p= 0.023	Low
Number of episodes with pH <4.0											
1 (Orenstein et al 1983[b])	RCT – crossover	Serious[a,b,c]	None	None	Serious[d]	None	n= 9 Mean (SD): 16.0 (7.2)	n= 9 Mean (SD): 10.1 (6.9)	MD: 6.00 (-0.47 to 12.47)[e]	p= 0.002	Low
Number of reflux episodes lasting longer than 5 minutes											
1 (Orenstein et al 1983[b])	RCT – crossover	Serious[a,b,c]	None	None	Serious[d]	None	n= 9 Mean (SD): 1.7 (1.8)	n= 9 Mean (SD): 0.6 (0.9)	MD: 1.00 (-0.46 to 2.46)[e]	p= 0.093	Low
Duration of the longest reflux episode in each 2 hour postprandial period											
1 (Orenstein et al 1983[b])	RCT – crossover	Serious[a,b,c]	None	None	Serious[d]	None	n= 9 Mean (SD): 6.7 (3.9)	n= 9 Mean (SD): 4.0 (2.4)	MD: 3.00 (0.08 to 5.92)[e]	p= 0.079	Low

[a] *Unclear whether there was adequate concealment of allocation*
[b] *Unclear whether investigators were blinded to intervention*
[c] *Unclear whether investigators were blinded to confounding factors*
[d] *Confidence interval of standardised mean difference crosses 2 zones (wide confidence interval)*
[e] *Calculated by NCC-WCH technical team based on data reported in the article*

Table 30: GRADE findings for comparison of supine reversed-Trendelenburg position of 10 degrees with flat supine positioning

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Supine reversed Trendelenburg**	**Flat supine**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux index (% of time with pH <4.0)											
1 (Bagucka et al 1999)	RCT – crossover	Very serious[a,b,c,d]	None	None	Serious[e]	None	n= 10 Mean (SD): 19.08 (13.10)	n= 10 Mean (SD): 10.62 (6.40)	MD: 8.00 (-0.87 to 16.87)[f]	p=0.08	Very low
Number of episodes with pH <4.0											
1 (Bagucka et al 1999)	RCT – crossover	Very serious[a,b,c,d]	None	None	Very serious[g]	None	n= 10 Mean (SD): 32.3 (8.00)	n= 10 Mean (SD): 33.9 (15.6)	MD: -2.00 (-13.09 to 9.09)[f]	p=0.95	Very low
Duration of the longest reflux episode											
1 (Bagucka et al 1999)	RCT – crossover	Very serious[a,b,c,d]	None	None	Serious[e]	None	n= 10 Mean (SD): 38.9 (46.81)	n= 10 Mean (SD): 17 (6.34)	MD: 22.00 (-7.37 to 51.37)[f]	p=0.16	Very low

[a] *Method of randomisation not reported*
[b] *Unclear whether there was adequate concealment of allocation*
[c] *Unclear whether investigators were blinded to intervention*
[d] *Unclear whether investigators were blinded to confounding factors*
[e] *Confidence interval of standardised mean difference crosses 2 zones (wide confidence interval)*
[f] *Calculated by NCC-WCH technical team based on data reported in the article*
[g] *Confidence interval of standardised mean difference crosses 3 zones (very wide confidence interval)*

Table 31: GRADE findings for comparison of prone with right lateral positioning

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Prone**	**Right lateral**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux index (% of time with pH <4.0)											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	None	Yes[c]	n=18 Mean (SD): 6.3 (7.2)	n=18 Mean (SD): 29.4 (13.6)	MD: -23.10 (-30.20 to -16.00)[d]	p<0.05	Moderate
1 (Tobin et al 1997)	RCT – crossover	Serious[a,b]	None	None	None	None	n=24 Mean (SD): 6.72 (5.2)	n=24 Mean (SD): 12.02 (6.8)	MD: -5.00 (-8.44 to -1.56)[d]	p<0.05	Moderate
Number of episodes with pH <4.0											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	None	Yes[c]	n=18 Mean (SD): 15.4 (11.9)	n=18 Mean (SD): 41.6 (19.5)	MD: -26.20 (-36.75 to -15.65)[d]	p<0.05	Moderate
Number of reflux episodes lasting longer than 5 minutes											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	None	Yes[c]	n=18 Mean (SD): 1.1(1.7)	n=18 Mean (SD): 4.5 (3.4)	MD: -3.40 (-5.15 to -1.65)[d]	p<0.05	Moderate
Duration of the longest reflux episode											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	None	Yes[c]	n=18 Mean (SD): 8.6 (9.3)	n=18 Mean (SD): 26 (16.5)	MD: -17.4 (-26.18 to -8.62)[d]	p<0.05	Moderate

[a] *Unclear whether investigators were blinded to intervention*
[b] *Unclear whether investigators were blinded to confounding factors*
[c] *Infants born premature*
[d] *Calculated by NCC-WCH technical team based on data reported in the article*

Table 32: GRADE findings for comparison of left lateral with right lateral positioning

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Left lateral**	**Right lateral**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux index (% of time with pH <4.0)											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	None	Yes[c]	n=18 Mean (SD): 11 (9.3)	n=18 Mean (SD): 29.4 (13.6)	MD: -18.4 (-26.01 to -10.79) d	p<0.05	Moderate
1 (Tobin et al 1997)	RCT – crossover	Serious[a,b]	None	None	None	None	n=24 Mean (SD): 7.69 (5)	n=24 Mean (SD): 12.02 (6.8)	MD: -4 (-7.44 to -0.56)[d]	p<0.05	Moderate
Number of episodes with pH <4.0											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	None	Yes[c]	n=18 Mean (SD): 24.6 (14.8)	n=18 Mean (SD): 41.6 (19.5)	MD: -17.00 (-28.33 to -5.67)[d]	p<0.05	Moderate
Number of reflux episodes lasting longer than 5 minutes											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	None	Yes[c]	n=18 Mean (SD): 1.8 (2.1)	n=18 Mean (SD): 4.5 (3.4)	MD: -2.70 (-4.55 to -0.85)[d]	p<0.05	Moderate
Duration of the longest reflux episode											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	None	Yes[c]	n=18 Mean (SD): 10 (10.2)	n=18 Mean (SD): 26 (16.5)	MD: -16 (-24.98 to -7.02)[d]	p<0.05	Moderate

[a] *Unclear whether investigators were blinded to intervention*
[b] *Unclear whether investigators were blinded to confounding factors*
[c] *Infants born premature*
[d] *Calculated by NCC-WCH technical team based on data reported in the article*

Table 33: GRADE findings for comparison of prone with left lateral positioning

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Prone**	**Left lateral**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux index (% of time with pH <4.0)											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	Serious[c]	Yes[d]	n=18 Mean (SD): 6.3 (7.2)	n=18 Mean (SD): 11 (9.3)	MD: -4.70 (-10.15 to 0.75)[e]	p<0.05	Low
1 (Tobin et al 1997)	RCT – crossover	Serious[a,b]	None	None	Serious[c]	None	n=24 Mean (SD): 6.72 (5.2)	n=24 Mean (SD): 7.69 (5.0)	MD: -1.00 (-3.83 to 1.83)[e]	NS	Low
Number of episodes with pH <4.0											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	Serious[c]	Yes[d]	n=18 Mean (SD): 15.4 (11.9)	n=18 Mean (SD): 24.6 (14.8)	MD: -9.20 (-17.98 to -0.42)[e]	p<0.05	Low
Number of reflux episodes lasting longer than 5 minutes											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	Serious[c]	Yes[d]	n=18 Mean (SD): 1.1 (1.7)	n=18 Mean (SD): 1.8 (2.1)	MD: -0.70 (-1.95 to 0.55)[e]	p>0.05[f]	Low
Duration of the longest reflux episode											
1 (Ewer et al 1999)	RCT – crossover	Serious[a,b]	None	None	Very serious[g]	Yes[d]	n=18 Mean (SD): 8.6 (9.3)	n=18 Mean (SD): 10 (10.2)	MD: -1.40 (-7.78 to 4.98)[e]	p>0.05[f]	Very low

[a] *Unclear whether investigators were blinded to intervention*
[b] *Unclear whether investigators were blinded to confounding factors*
[c] *Confidence interval of standardised mean difference crosses 2 zones (wide confidence interval)*
[d] *Infants born premature*
[e] *Calculated by NCC-WCH technical team based on data reported in the article*
[f] *Unclear reporting but seems as though p >0.05*
[g] *Confidence interval of standardised mean difference crosses 3 zones (very wide confidence interval)*

Table 34: GRADE findings for comparison of left lateral with supine positioning

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Left lateral**	**Supine**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux index (% of time with pH <4.0)											
1 (Tobin et al 1997)	RCT – crossover	Serious[a,b]	None	None	None	None	n=24 Mean (SD): 7.69 (5.0)	n=24 Mean (SD): 15.33 (11.4)	MD: -7.00 (-11.83 to -2.17)[c]	p <0.05	Moderate

[a] *Unclear whether investigators were blinded to intervention*
[b] *Unclear whether investigators were blinded to confounding factors*
[c] *Calculated by NCC-WCH technical team based on data reported in the article*

J.2.2 Feeding changes

Table 35: GRADE findings for comparison of thickened feeds with standard formula feeds for reduction in GOR related symptoms.

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Thickened feed	Standard /comparator feed	Relative (95% CI)	Absolute (95% CI)	Quality
Reduced frequency of overt regurgitation: pH and/or impedance monitoring											
Number of infants without regurgitation											
1 (Iacono et al, 2002)	RCT	Serious[a, b]	None	None	None	None	28 of 82	12 of 84	Relative Risk: 2.39 [1.31, 4.37]	N/A	Moderate
Number of episodes of regurgitation (per day or week)											
3 (4 arms) Moukarzel et al, 2007 Xinias et al, 2005 Miyazawa et al, 2006)	Meta analysis of RCTs	Serious[a, b]	Serious[c]	None	None	Yes[d]	-	-	Mean Difference: -2.00 [-4.65, 0.65]	N/A	Low
Change in regurgitation frequency from baseline at one week											
1 (Vanderhoof et al, 2003)	RCT	Very Serious[a, e]	None	None	Serious	No	Change -6 (range +/- 1)[g]	Change -6 (range +/- 1)[g]	Non-significant[g]	N/A	Very low
Episodes of emesis over a 90 minute period											
1 (Orenstein et al, 1986)	RCT; crossover	Very Serious[h]	None	None	Serious	No	1.2 (SD +/- 0.7)[g]	3.9 (SD +/- 0.9)[g]	p = 0.015[g]	N/A	Very low
1 (Wenzl et al, 2003)	RCT; crossover	None	None	None	Serious[f]	No	1.07 (SD +/- 1.69)[g]	4.86 (SD +/- 4.05[g]	p < 0.003[g]	N/A	Moderate
Frequency of regurgitation per day, median (IQR)											
1 (Miyaz	RCT; crosso	Serio	None	Non	Seri	Yes[i]	HL-350	HL-00	p =	N/A	Low

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Thickened feed**	**Standard /comparator feed**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
awa et al, 2004)	ver within arms	us[a]		e	ous[f]		Median 1.6 (IQR 0.8 to 2.0)[g]	Median 3.5 (IQR 2.3 to 4.9)[g]	0.021[g]		
1 (Miyazawa et al, 2004)	RCT; crossover within arms	Serious[a]	None	None	Serious[f]	Yes[i]	HL-450 Median 1.3 (IQR 0.6 to 2.3)[g]	HL-00 Median 2.9 (IQR 2.0 to 3.2)[g]	p = 0.0003[g]	N/A	Low
1 (Miyazawa et al, 2007)	RCT;	Serious[a]	None	None	Serious[f]	Yes[i]	HL-350 Median 2.3 (IQR 1.6 to 3.6)[g]	HL-00 Median 5.2 (IQR 3.7 to 7.8)[g]	p < 0.01)[g]	N/A	Low
Number of episodes of vomiting per day											
2 (Moukarzel et al, 2007 Xinias et al, 2005)	Meta analysis of RCTs	Serious[a]	None	None	None	Yes	-	-	Mean Difference: -0.97 [-1.54, -0.39]	N/A	Moderate
Reflux measured using oesophageal pH or impedance monitoring											
Reflux Index (% time pH < 4.0)											
3 (Moukarzel et al, 2007 Xinias et al, 2005 Vandenplas et al, 1994)	Meta-analysis of RCTs	Serious[a, b]	None	None	None	Yes[d]	-	-	Mean Difference: -3.38 [-5.28, -1.48]	N/A	Moderate

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Thickened feed**	**Standard /comparator feed**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Resolution of faltering growth											
Weight gain (grams per day)											
4 (Chao & Vandenplas, 2007a Chao & Vandenplas, 2007b Xinias et al, 2005)	Meta analysis of RCTs	Very serious[a]	Serious[c]	None	None	Yes[d]	-	-	Mean Difference: 3.99 [1.66, 6.31]	N/A	Low
Adverse events											
Discontinued due to diarrhoea											
1 (Iacono et al, 2002)	RCT	Serious[a, b]	None	None	None	No	14 of 82	0 of 84	∞	N/A	Moderate
Reported adverse events (not specified)											
1 (Vanderhoof et al, 2003)	RCT	Very Serious[a]	None	None	Serious[f]	No	-	-	No difference between groups[g]	N/A	Very Low
1 (Miyazawa et al, 2004)	RCT; crossover within arms	Serious[a]	None	None	Serious[f]	No	-	-	No difference between groups[g]	N/A	Low
1 (Xinias et al, 2005)	RCT;	Serious[a]	None	None	Serious[f]	No	-	-	No difference between groups[g]	N/A	Low

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Thickened feed	Standard /comparator feed	Relative (95% CI)	Absolute (95% CI)	Quality
Parent reported reduction in infant distress – not reported											
Improvement in validated reflux questionnaire – not reported											
Parent satisfaction with this intervention – not reported											

[a] *Method of randomisation not described in detail*
[b] *High discontinuation rate*
[c] *High heterogeneity between studies*
[d] *Variation in viscosity of formulas and nutritional value of formulas*
[e] *Children assessed at one week and some given further treatment*
[f] *Imprecision could not be investigated due to way result have been reported and cross-over design*
[g] *Result as reported in study*
[h] *Study based on response to a single feed; Method of investigation was scintigraphically*
[I] *It is unclear how these studies are linked. Numbers in each arm differ.*
N/A Not Applicable

Table 36: GRADE findings for comparison of thickened feeds with standard formula feeds for reduction in GOR related symptoms in children with Cerebral Palsy

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Thickened feed	Standard/comparator feed	Relative (95% CI)	Absolute (95% CI)	Quality
Reduced frequency of overt regurgitation											
1 (Miyazawa et al, 2008)	RCT; crossover within arms	Serious[a]	None	None	Serious[b]	No	High pectin Median 2.5 (IQR 1.0 to 5.0)	Standard feed median 1.0 (IQR 1.0 to 1.5)	P < 0.05	N/A	Low
1 (Miyazawa et al, 2008)	RCT; crossover within arms	Serious[a]	None	None	Serious[b]	No	Low Pectin median 0.0 (0.0 to 0.5)	Standard feed median 0.0 (0.0 to 0.1)	NS	N/A	Low
Reflux measured using oesophageal pH or impedance monitoring – not reported											
Resolution of faltering growth – not reported											
Adverse events – not reported											
Parent reported reduction in infant distress – not reported											
Improvement in validated reflux questionnaire – not reported											
Parent satisfaction with this intervention – not reported											

[a] *Method of randomisation not described in detail*
[b] *Could not be calculated*
NS Not significant at p < 0.05

Table 37: GRADE findings for comparison of thickened feeds (Soy milk and fibre) with standard formula feeds for reduction in GOR related symptoms.

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Thickened feed**	**Standard /comparator feed**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reduced frequency of overt regurgitation											
Number of infants without regurgitation											
1 (Ostorm et al, 2006)	RCT	Very Serious[a, b]	None	Very serious[c]	Serious[d]	No	11 of 67	3 of 66	Relative Risk: 3.61 [1.06, 12.36]	N/A	Very Low
Number of episodes of regurgitation											
1 (Ostorm et al, 2006)	RCT	Very Serious[a, b]	None	Very serious[c]	None	No		-	Mean difference: -0.40 [-0.49, -0.31]	N/A	Very Low
Reflux measured using oesophageal pH or impedance monitoring – not reported											
Resolution of faltering growth – not reported											
Adverse events – not reported											
Parent reported reduction in infant distress – not reported											
Improvement in validated reflux questionnaire – not reported											
Parent satisfaction with this intervention – not reported											

[a] *Effect of cow's milk intolerance not controlled for in analysis*
[b] *25% discontinuation rate across study*
[c] *Wide confidence intervals*
N/A Not Applicable

Table 38: GRADE findings for comparison of thickened feeds with standard formula feeds plus positional management for reduction in GOR related symptoms

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Thickened feed**	**Standard /comparator feed**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reduced frequency of overt regurgitation											
Number of episodes of regurgitation and vomiting per day											
1 (Chao & Vandenplas, 2007b)	RCT	Serious[a, b]	None	Serious[c]	Serious[d]		-	-	Mean Difference: -0.77 [-1.16, -0.38]	N/A	Very low
Reflux measured using oesophageal pH or impedance monitoring – not reported											
Resolution of faltering growth – not reported											
Adverse events – not reported											
Parent reported reduction in infant distress – not reported											
Improvement in validated reflux questionnaire – not reported											
Parent satisfaction with this intervention – not reported											

[a] *Randomisation and concealment not described in detail*
[b] *20% discontinuation from study*
[c] *Comparison group had positional management*
[d] *Wide confidence intervals*
N/A Not Applicable

Table 39: GRADE findings for comparison of thickened feeds with 25% strengthened regular formula for reduction in GOR related symptoms.

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Thickened feed**	**Standard /comparator feed**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reduced frequency of overt regurgitation											
Number of episodes of regurgitation and vomiting per day											
1 (Chao & Vandenplas, 2007a)	RCT	Serious[a]	None	Serious[b]	None	Yes	-	-	Mean Difference -1.96 [-2.34, -1.58]	N/A	Very low
Reflux measured using oesophageal pH or impedance monitoring – not reported											
Resolution of faltering growth – not reported											
Adverse events – not reported											
Parent reported reduction in infant distress – not reported											
Improvement in validated reflux questionnaire – not reported											
Parent satisfaction with this intervention – not reported											

[a] *Randomisation and concealment not described in detail*
[b] *Comparison group had partially strengthened formula.*
N/A Not Applicable

Table 40: GRADE findings for comparison of cow's milk protein elimination with continued cow's milk diet on the symptoms of GER

Quality assessment							Number of infants		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Intervention**	**Comparator**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reduced frequency of overt regurgitation – not reported											
Reflux measured using oesophageal pH or impedance monitoring											
Total number of reflux episodes											
1 (Borrelli et al, 2012)	Non-randomised clinical trial	Very serious[a]	None	None	Serious[b]	No	Amino acid formula: Median 65 (range 39 to 87.5)	Standard cow's milk: Median 105 (range 58 to 127.5)	p < 0.001	N/A	Very low
Reflux Index (% of time pH < 4.0)											
1 (Borrelli et al, 2012)	Non-randomised clinical trial	Very serious[a]	None	None	Serious[b]	No	Amino acid formula: Median 3.4 (SD +/- 2.6)	Standard cow's milk: Median 3.6 (SD +/- 2.7)	NS	N/A	Very low
Resolution of faltering growth – not reported											
Adverse events – not reported											
Parent reported reduction in infant distress – not reported											
Improvement in validated reflux questionnaire – not reported											
Parent satisfaction with this intervention – not reported											

[a] *Non-randomised study design & all children were known to have CMA*
[b] *Could not be calculated*
N/A Not Applicable

Table 41: GRADE findings for comparison of differing feeding volumes on symptoms of GER

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Intervention**	**Comparator**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reduced frequency of overt regurgitation – not reported											
Reflux measured using oesophageal pH or impedance monitoring											
Total number of reflux episodes											
1 (Sutphen & Dillard, 1988)	Non-randomised crossover clinical trial	Very serious[a, b]	None	None	Serious[c]	No	9 ml/kg mean 8.1 (SD 13.9)	18 ml/kg mean 14.3 (SD 12.5)	p = 0.004	N/A	Very low
1 (Sutphen & Dillard, 1988)	Non-randomised crossover clinical trial	Very serious[a, b]	None	None	Serious[c]	No	9 ml/kg = mean 9.6 (SD 7.2)	27.3 ml/kg = mean 24.4 (SD 20.2)	p = 0.007	N/A	Very low
Resolution of faltering growth – not reported											
Adverse events – not reported											
Parent reported reduction in infant distress – not reported											
Improvement in validated reflux questionnaire – not reported											
Parent satisfaction with this intervention – not reported											

[a] *Non-randomised study design*
[b] *Variation in how study protocol was applied.*
[c] *Could not be calculated*
N/A Not Applicable

J.2.3 Alginates and Antacids

Table 42: GRADE findings for comparison of aluminium-free infant Gaviscon (sodium alginate) with placebo in infants aged less than 6 months.

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Aluminium free Infant Gaviscon (sodium alginate)**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Cessation (or symptom free days) of overt regurgitation											
Reported as at least 10% symptom free days, %											
1 (Miller et al 1999)	RCT	Very serious[a,b,c,d,e]	None	None	Serious[f]	None	13/42 (31%)	5/46 (11%)	p=0.027[g] Odds ratio [OR] (95% CI): 3.68 (1.18 to 11.44)[h]	-	Very low
Reduced frequency of overt regurgitation											
Reported as median number of vomiting/regurgitation episodes in the previous 24 hours											
1 (Miller et al 1999)	RCT	Very serious[a,b,c,d,e]	None	None	Not assessed[i]	None	n=42 Median (range): 3.0 (0 to 22)	n=46 Median (range): 5.0 (0 to 37)	p=0.009[g]	-	Low
Reported as mean frequency of vomiting/regurgitation episodes after 14 days											
1 (Miller et al 1999)	RCT	Very serious[a,b,c,d,e]	None	None	Not assessed[i]	None	n=42 Mean: 4.5 (Standard deviation [SD] not reported)	n=46 Mean: 6.2 (SD not reported)	p=0.056[g]	-	Low
Adverse outcomes, n (%)											
Functional diarrhoea											
1 (Miller	RCT	Very serio	None	Non	Very serio	None	6/42 (14.3	5/46 (10.9	p>0.1[k]	-	Very

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Aluminium free Infant Gaviscon (sodium alginate)	Placebo	Relative (95% CI)	Absolute (95% CI)	Quality
et al 1999)		us[a,b,c,d,e]		e	us[j]		%)	%)	OR (95% CI): 1.37 (0.38 to 4.86)[h]		low
Teething syndrome											
1 (Miller et al 1999)	RCT	Very serious[a,b,c,d,e]	None	None	Very serious[j]	None	5/42 (11.9 %)	3/46 (6.5 %)	p>0.1[k] OR (95% CI): 1.94 (0.43 to 8.66)[h]	-	Very low
Diarrhoea not otherwise specified											
1 (Miller et al 1999)	RCT	Very serious[a,b,c,d,e]	None	None	Very serious[j]	None	1/42 (2.4%)	4/46 (8.7 %)	p>0.1[k] OR (95% CI): 0.26 (0.03 to 2.39)h	-	Very low
Constipation											
1 (Miller et al 1999)	RCT	Very serious a,b,c,d,e	None	None	Very serious[j]	None	4/42 (9.5%)	1/46 (2.2 %)	p>0.1[k] OR (95% CI): 4.74 (0.51 to 44.20)[h]	-	Very low
Acute nasopharyngitis											
1 (Miller et al 1999)	RCT	Very serious[a,b,c,d,e]	None	None	Very serious[j]	None	3/42 (7.1%)	1/46 (2.2 %)	p>0.1[k] OR (95% CI): 3.46 (0.35		Very low

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Aluminium free Infant Gaviscon (sodium alginate)	Placebo	Relative (95% CI)	Absolute (95% CI)	Quality
									to 34.64)[h]		
Colic[l]											
1 (Miller et al, 1999)	RCT	Very serious[a,b,c,d,e]	None	None	Very serious[j]	None	2/42 (4.8%)	3/46 (6.5%)	p>0.1[k] OR (95% CI): 0.72 (0.11 to 4.51)[h]	-	Very low
Parent reported reduction in infant distress											
Reported as parent/guardian assessment of symptoms, n (%)											
1 (Miller et al, 1999)	RCT	Very serious[a,b,c,d,e]	None	None	Serious[i]	None	Very good + good: 33/41 Acceptable, poor + very poor: 8/41	Very good + good: 21/44 Acceptable, poor + very poor: 23/44	Chi squared equals 8.468[g] p= 0.0036[g]	-	Very Low

NA - not applicable
[a] *Randomisation not described in detail*
[b] *Unclear whether there was adequate allocation concealment*
[c] *Unclear whether investigators were blinded to intervention*
[d] *Unclear whether investigators were blinded to confounding factors*
[e] *20 withdrawals (alginate, n=7; placebo, n=13; p>0.2) due primarily to adverse events (alginate, n=4; placebo, n=7) and lack of efficacy (alginate, n=2; placebo, n=3)*
[f] *Wide confidence interval (CI crosses 2 zones)*
[g] *As reported in the study (Wilcoxon rank sum test)*
[h] *Calculated by NCC-WCH technical team based on data reported in the article*
[i] *Imprecision could not be investigated due to way result has been reported*
[j] *Very wide confidence interval (CI spans 3 zones)*
[k] *As reported in article (chi square or Fisher's exact test, as appropriate)*
[l] *Reported as adverse event in paper*

Table 43: GRADE findings for comparison of Gaviscon (alginate) with placebo in children aged up to 3 years

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Gaviscon (alginate)**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reflux measured using oesophageal pH-metry											
Total number of reflux episodes (oesophageal pH <4 for at least 25 seconds) in 24 hours											
1 (Butset al 1987)	RCT	Serious[a,b,c,d,e]	None	None	Serious[f]	None	n=10 Mean (SD): 56.0 (53.1)	n=10 Mean (SD): 90.6 (46.5)	p-value for after Gaviscon versus before Gaviscon: p<0.05[g] p-value for after placebo versus before placebo: NS[g] Mean Difference [MD] (95% CI): -35.00 (-78.50 to 8.50)[h]	-	Low
Number of reflux episodes greater than 5 minutes											
1 (Butset al 1987)	RCT	Serious[a,b,c,d,e]	None	None	None	None	n=10 Mean (SD): 1.2 (0.6)	n=10 Mean (SD): 4.6 (2.8)	p-value for after Gaviscon versus before Gaviscon: p<0.05[g]	-	Moderate

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Gaviscon (alginate)**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
									p-value for after placebo versus before placebo: NS[g] MD (95% CI): -4.00 (-5.96 to -2.04)[h]		
Percent total reflux (Reflux Index)											
1 (Butset al 1987)	RCT	Serious[a,b,c,d,e]	None	None	None	None	n=10 Mean (SD): 6.1 (0.9)	n=10 Mean (SD): 10.1 (4.4)	p value for after Gaviscon versus before Gaviscon: p<0.05[g] p value for after placebo versus before placebo: NS[g] MD (95% CI): -4.00 (-6.56 to -1.44)[h]	-	Moderate
Adverse outcomes (events not specified), n (%)											

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Gaviscon (alginate)**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
1 (Buts et al 1987)	RCT	Serious[a,b,c,d,e]	None	None	Serious[i]	None	n=100/10 (0%)	n=100/10 (0%)	-	-	Low

NS – not significant
[a] *Randomisation method not described in detail*
[b] *Alternate allocation to treatments*
[c] *Not all subjects endoscoped*
[d] *Unclear whether investigators were blinded to intervention*
[e] *Unclear whether investigators were blinded to confounding factors*
[f] *Wide confidence interval (confidence interval of SMD crosses 2 zones)*
[g] *As reported in study*
[h] *Calculated by NCC-WCH technical team based on data reported in the article*
[i] *Imprecision could not be investigated due to way result have been reported*

Table 44: GRADE findings for Gaviscon infant liquid (alginic acid with antacid) with placebo in children and young adults aged up to 17 years.

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Gaviscon infant liquid (alginic acid with antacid)	Placebo	Relative (95% CI)	Absolute (95% CI)	Quality
Reflux measured using oesophageal pH-metry											
Number of episodes of GER (esophageal pH <4) in 24 hours											
1 (Forbes et al 1986)	RCT	Serious[a,b,c,d]	None	None	Serious[e]	None	n=10 Mean (SD): 81 (72.7)	n=10 Mean (SD): 49 (34.8)	p: NS[f] MD (95% CI): 32.00 (-18.18 to 82.18)[g]	-	Low
Total duration of acid reflux in minutes											
1 (Forbes et al 1986)	RCT	Serious[a,b,c,d]	None	None	Very serious[h]	None	n=10 Mean (SD): 74 (123.3)	n=10 Mean (SD): 96 (34.8)	p: NS[f] MD (95% CI): -22.00 (-101.26 to 57.26)[g]	-	Very low
Adverse outcomes (events not specified), n (%)											
1 (Forbes et al 1986)	RCT	Very serious[a,b,c,d]	None	None	Not assessed[i]	None	n=10 0/10 (0%)	n=10 0/10 (0%)	-	-	Low

NS – not significant
[a] *Method of randomisation not described in detail*
[b] *Unclear whether there was adequate allocation concealment*
[c] *Not all subjects endoscoped*
[d] *Unclear whether investigators were blinded to confounding factors*
[e] *Wide confidence interval (confidence interval of SMD crosses 2 zones)*
[f] *As reported in the study (Wilcoxon signed rank test)*
[g] *Calculated by NCC-WCH technical team based on data reported in the article*
[h] *Very wide confidence interval (confidence interval of SMD crosses 3 zones)*
[i] *Imprecision could not be investigated due to way result have been reported*

Table 45: GRADE findings for infant Gaviscon (sodium and magnesium alginate and mannitol but no bicarbonate) with placebo in infants aged up to 12 months.

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Infant Gaviscon (sodium and magnesium alginate and mannitol)	Placebo	Relative (95% CI)	Absolute (95% CI)	Quality
Reflux measured using intra-oesophageal impedance and dual channel pH monitoring											
Number of reflux events per hour											
1 (Del Buono et al 2005)	RCT	Very Serious[a,b,c,d]	None	None	Not assessed[e]	None	-	Median difference (placebo -– Gaviscon infant), range: 0.06 (-1.20 to 3.80)	P = 0.784[f]	-	Low
Number of acid reflux events per hour											
1 (Del Buono et al 2005)	RCT	Very Serious[a,b,c,d]	None	None	Not assessed[e]	None	-	Median difference (placebo -– Gaviscon infant), range: -0.02 (-0.55 to 3.94)	p = 0.940[f]	-	Low
Total reflux time per hour (seconds per hour)											

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Infant Gaviscon (sodium and magnesium alginate and mannitol)**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
1 (Del Buono et al 2005)	RCT	Serious[a,b,c,d]	None	None	Not assessed[e]	None	-	Median difference (placebo -– Gaviscon infant), range: -7.6 (-38.5 to 111.8)	p = 0.096[f]	-	Low

[a] *Method of randomisation not described in detail*
[b] *Unclear whether groups were comparable at baseline (baseline characteristics not reported)*
[c] *Unclear whether groups were comparable for dropout (numbers not reported)*
[d] *Unclear whether groups were comparable for missing data (numbers not reported)*
[e] *Imprecision could not be investigated due to way result have been reported*
[f] *As reported in study (Wilcoxon signed rank test) a Method of randomisation not described in detail*

J.3 Pharmacological treatment of GORD

Table 46: GRADE findings for comparison of PPIs with placebo for the management of GORD in infants.

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Proton Pump Inhibitor**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reduced frequency of overt regurgitation											
Regurgitation (Change % of feeds per week)											
1 (Orenstein et al, 2009)	RCT	Serious[a]	None	None	Not assessed[b]	None	Lansoprazole: N = 81, -14%	n = 81, -11%	NS[c]	N/A	Moderate
Frequency of vomiting											
1 (Omari et al, 2007)	RCT, Crossover	Very Serious [a,d]	None	None	Not assessed[b]	None	Omeprazole: Median 8.5 (IQR 7 to 22.8)	Median 6.5 (IQR 3 to 14.3)	NS [c]	N/A	Low
Vomiting											
1 (Davidson et al, 2013)	RCT	Serious [e]	None	Serious [f]	Very serious [g]	None	Esomeprazole: Mean 5.21 (SD 6.75)	Mean 4.87 (SD 5.93)	MD 0.34 [-3.15, 3.83]	N/A	Very low
Frequency of regurgitation											
1 (Hussain et al, 2014)	RCT	Very Serious [a,h]	None	Serious[i]	Not assessed[b]	None	Raberprazole: NR	NR	NS[c]	N/A	Very Low
Reflux measured using oesophageal pH-monitoring or impedance monitoring											
Number of acid GER episodes											
1 (Omari et al, 2007)	RCT, Crossover	Very Serious [a,d]	None	None	Not assessed[b]	None	Omeprazole: 59.6 (SE 26.7)	119.4 (SE 20.9)	p < 0.05[c]	N/A	Moderate
Number of acid GER episodes lasting longer than 5 minutes											
1 (Omari et al, 2007)	RCT, Crossover	Very Serious [a,d]	None	None	Not assessed[b]	None	Omeprazole: 3.0 (SE	8.0 (SE 2.1	a < 0.01[c]	N/A	Moderate

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Proton Pump Inhibitor**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
							2.0))				
Longest acid GER episode (minutes)											
1 (Omari et al, 2007)	RCT, Crossover	Very Serious [a,d]	None	None	Not assessed [b]	None	Omeprazole: 16.3 (SE 8.0)	48.6 (SE 10.1)	p < 0.01[c]	N/A	Moderate
% time pH < 4.0											
1 (Omari et al, 2007)	RCT, Crossover	Very Serious [a,d]	None	None	Not assessed [b]	None	Omeprazole: 4.9 (SE 3.4)	19.0 (SE 4.5)	p < 0.01[c]	N/A	Moderate
1 (Moore et al, 2003)	RCT, Crossover	Serious [a,j]	None	Serious[k]	Not assessed [b]	None	Omeprazole: 1.0 (SD 1.3)	5.3 (SD 4.9)	p < 0.01[c]	N/A	Low
Resolution of oesophagitis – not reported											
Resolution of faltering growth – not reported											
Adverse outcomes											
Adverse events											
1 (Orenstein et al, 2009)	RCT	Serious[a]	None	None	None	None	Lansoprazole: 50[l]	37[l]	NS[c]	N/A	Moderate
1 (Hussain et al, 2014)	RCT	Very Serious [a, h]	None	Serious[i]	None	None	Rebeprazole: 83/178	42/89	NS[c]	N/A	Very Low
Serious adverse events											
1 (Orenstein et al, 2009)	RCT	Serious[a]	None	None	None	None	Lansoprazole: 10[m]	2[m]	p = 0.032 [c]	N/A	Moderate
1 (Omari et al, 2007)	RCT, Crossover	Very Serious [a,d]	None	None	None	None	Omeprazole: 0	0	NS[c]	N/A	Low
1 (Davidson et al, 2013)	RCT	SeriousI[e]	None	Serious[f]	None	None	Esomeprazole: 6	9	NS[c]	N/A	Low
Parent reported reduction in infant distress											

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Proton Pump Inhibitor**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Global severity index (parent reported improved at 4 weeks)											
1 (Orenstein et al, 2009)	RCT	Serious[a]	None	None	Not assessed[b]	None	Lansoprazole: 45	44	NS[c]	N/A	Moderate
Improvement in validated reflux questionnaire											
Visual Analogue Scale by parents of infants irritability											
1 (Moore et al)	RCT, Crossover	Serious[j]	None	Serious[k]	Not assessed[b]	None	Omeprazole: 5.0 (SD 3.1)	5.9 (SD 2.1)	p = 0.214[c]	N/A	High
I-GERQ-R											
1 (Hussain et al, 2014)	RCT	Very Serious[j]	None	Serious[i]	Not assessed[b]	None	Raberprazole: NR	NR	NS[c]	N/A	Very Low
Parent satisfaction with this intervention											
Responder rate (>50% reduction in feeding or crying symptoms from baseline)											
1 (Orenstein et al, 2009)	RCT	Serious[a]	None	None	Not assessed[b]	None	Lansoprazole 44%	44%	NS[c]	N/A	Moderate
Discontinued due to non-efficacy											
1 (Orenstein et al, 2009)	RCT	Serious[a]	None	None	Very serious[o]	None	Lansoprazole: 28 of 81	29 of 81	0.97 [0.64, 1.47]	N/A	Low
Discontinued due to worsening symptoms											
1 (Winter et al, 2012)	RCT	Serious[n]	None	Serious[p]	Very serious[o]	None	Esomeprazole: 15 of 39	20 of 41	0.79 [0.48, 1.31]	N/A	Very Low

CI confidence interval; RCT randomised controlled trial; NS not significant, NA not applicable; MD mean difference; NR not reported; SE standard error; GER gastro-esophageal reflux
NS Non significant at p < 0.05.
N/A Not applicable – could not be calculated on data available.

a Poor reporting of results that not all GRADE items could be assessed
b Reporting of results did not allow imprecision to be calculated.
c As reported in the study.
d Small sample size; no washout period during crossover between treatments.

e Groups unbalanced at baseline; small sample size
f Study included neonates only
g SMD confidence intervals cross several categories on Cohen effect size. MD presented in table as more relevant.
h Method of randomisation not described in detail
i Only included infants in whom PPIs were effective in a pre-randomisation phase.
j Method of randomisation not explained in detail; no washout period; results from before crossover
k Infants had GERD and were irritable.
l Reported events were: Infection – URI, ear, LRTI, viral, constipation, eczema, fever, respiratory tract congestion, rhinorrhea, candidiasis, diarrhea, vomiting.
m Reported events were: Lower respiratory infection, diarrhea, Ileua, dehydration, otitis media, upper respiratory infection, epididymal infection, arachnoid cyst, febrile convulsion, klebsiella infection.
n Method of randomisation and concealment not explained in detail.
o Confidence intervals cross several +/- 0.25 RR
p Infants had to respond to treatment to enter the randomised part of the study.

Table 47: GRADE findings for comparison of H_2 receptor antagonists with placebo for the management of GORD in infants.

Quality assessment							Number of Children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	H_2RA	Comparator	Relative (95% CI)	Absolute (95% CI)	Quality
Reduced frequency of overt regurgitation											
Regurgitation at 4 weeks											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	Not assessed[k]	None	Nizatidine: Mean 1.3 (SD 1.1)	Mean 2.2 (SD 1.3)	N/A[b]	N/A	Very low
Vomiting at 4 weeks											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	Not assessed[k]	None	Nizatidine: Mean 0.8 (SD 0.9)	Mean 2.1 (SD 1.1)	N/A[c]	N/A	Very low
Regurgitation at 8 weeks											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	Not assessed[k]	None	Nizatidine: Mean 0.3 (SD 0.7)	Mean 1.7 (SD 1.4)	N/A[b]	N/A	Very Low
Vomiting at 8 weeks											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	Not assessed[k]	None	Nizatidine: Mean 0.4 (SD 0.7)	Mean 1.6 (SD 1.7)	N/A[c]	N/A	Very Low
Reflux measured using oesophageal pH-monitoring or impedance monitoring											
% of reflux episodes (Reflux Index)											
1 (Simeone et al, 1997)	RCT	Very serious a	None	None	Not assessed[k]	None	Nizatidine: Median 4.3 (range 1.5 to 11.2)	Median 10.4 (4.1 to 18.8)	N/A[d]	N/A	Very Low
Number of reflux episodes											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	Not assessed[k]	None	Nizatidine: Median 85.8 (range 42	Median 123 (range 32 to 360)	N/A[d]	N/A	Low

Quality assessment							Number of Children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**H_2RA**	**Comparator**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
							to 227)				
Number of reflux episodes > 5 minutes											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	Not assessed[k]	None	Nizatidine: Median 1.7 (range 0 to 6)	Median 5.4 (range 2 to 10)	N/A[d]	N/A	Very low
Duration time of longest episode											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	Not assessed[k]	None	Nizatidine: Median 11.8 (range 4 to 40)	Median 25.1 (range 3 to 73)	N/A[d]	N/A	Very low
Resolution of oesophagitis - endoscope											
Esophagitis score[g]											
1 (Cucchiara et al, 1989)	RCT	Serious[e]	None	None	Not assessed[k]	None	Cimetidine: Mean 1.6 (SD 2.43)	Mean SD 5.43 (3.81)	N/A[f]	N/A	Low
Esophagitis score improved[g]											
1 (Cucchiara et al, 1989)	RCT	Serious[e]	None	None	None	None	Cimetidine: 16 of 17	9 of 15	RR 1.57 [1.02, 2.41]	N/A	Moderate
Endoscopy score normal[h]											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	None	None	Nizatidine: 5 of 12	2 of 12	RR 2.50 [0.60, 10.46]	N/A	Low
Histology score normal[i]											
1 (Simeone et al, 1997)	RCT	Very serious[a]	None	None	None	None	Nizatidine: 9 of 12	3 of 12	RR 3.00 [1.07, 8.43]	N/A	Low
Resolution of faltering growth – Not reported											
Adverse outcomes											
1 (Cucc	RCT	Serious[e]	None	None	None	None	Cimetidine:	0	NS[j]	N/A	Moderate

Quality assessment							Number of Children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	H_2RA	Comparator	Relative (95% CI)	Absolute (95% CI)	Quality
hiara et al, 1989)							0				
Parent reported reduction in infant distress – Not reported											
Improvement in validated reflux questionnaire											
Clinical score											
1 (Cucchiara et al, 1989)	RCT	Serious[e]	None	None	Not assessed[k]	None	Cimetidine: Mean 5.00 (SD 4.36)	Mean 9.46 (SD 4.86)	N/A[f]	N/A	Low
% improvement in clinical score from baseline											
1 (Cucchiara et al, 1989)	RCT	Very Serious[e]	None	None	Not assessed[k]	None	Cimetidine: Mean -67.39% (SD 23.17)	Mean -29.57% (SD 30.31)	p < 0.01[j]	N/A	Low
Parent satisfaction with this intervention – Not reported											

H_2RA H_2 receptor antagonists; CI confidence interval; RCT; randomised controlled trial; SD standard deviation; NA not applicable; RR relative risk
NS Non significant at p< 0.05
N/A Not applicable – could not be calculated on data available
a Method of randomisation not explained in detail. Small sample size. High dropout rate (26%). Poor reporting of study results so GRADE items could not be assessed.
b Based on a categorical score 0 to 3 so cannot be analysed as a continuous variable. Reduced from baseline in intervention group but not placebo.
c Based on a categorical score 0 to 3 so cannot be analysed a continuous variable. Significantly reduced from baseline in both groups by 8 weeks.
d No comparative results presented. Significantly reduced in treatment group compared to baseline, but not the placebo group.
e Method of randomisation and allocation concealment not explained in detail. Poor reporting of study results so GRADE items could not be assessed.
f Based on a categorical score 0 to 9 so cannot be analysed a continuous variable. Reduced from baseline in intervention group but not placebo.
g Scored from 0 to 9 – normal mucosa, mild degree, moderate degree, severe degree
h Classified as "Normal, erithema and edema, erythema and friability, erosions."
I Classified as "Normal, mild or moderate histology."
j As reported by authors
k Reporting of results did not allow imprecision to be calculated.

Table 48: GRADE findings for comparison of prokinetics (metoclopramide and domperidone) with placebo for the management of GORD in infants.

Quality assessment							Number of Children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Prokinetic	Comparator	Relative (95% CI)	Absolute (95% CI)	Quality
Reduced frequency of overt regurgitation											
1 (Leung et al, 1984)	RCT	Very serious[a]	None	None	Not assessed[b]	None	Metoclopramide: 1.6 (SD 2.0)	Not reported	p < 0.05[c]	N/A	Very low
Reflux measured using oesophageal pH-monitoring or impedance monitoring											
% of reflux episodes < 4.0											
1 (Bines et al, 1992)	RCT	Very serious[d]	None	None	Not assessed[b]	None	Domperidone: Mean 11.8 (SD not reported)	Mean 15.9 (SD not reported)	NS[c]	N/A	Very Low
1 (Carroccio et al, 1993)	RCT	Serious[e]	None	None	Not assessed[b]	None	Domperidone: Median 8 (range 2 to 35)	Median 9 (range 3 to 40)	NS[c]	N/A	Low
1 (Bellissant et al, 1997)	RCT	Serious[e]	None	None	Very serious[f]	None	Metoclopramide: Mean 6.7 (SD 9.2)	Mean 8.1 (SD 11.7)	MD -1.40 [-7.99, 5.19]	N/A	Low
1 (Tolia et al, 1989)	RCT, crossover	Very serious[g]	None	None	Not assessed[b]	None	Metoclopramide: Median 10.3 (range 2.4 to 22.8)	Median 13.4 (2.8 to 30.5)	p < 0.001[c]	N/A	Low
Number of reflux episodes < 4.0											
1 (Bines et al, 1992)	RCT	Very serious[d]	None	None	Not assessed[b]	None	Domperidone: 26 (SD not	28 (SD not reported)	p = 0.001[b]	N/A	Very Low

Quality assessment							Number of Children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Prokinetic	Comparator	Relative (95% CI)	Absolute (95% CI)	Quality
							reporter)				
1 (Carroccio et al, 1993)	RCT	Serious[e]	None	None	Not assessed[b]	None	Domperidone: median 48.5 (range 2 to 181)	Median 68 (range 38 to 130)	N/S[b]	N/A	Moderate
1 (Cresi et al, 2008)	RCT	Serious[e]	None	None	Not assessed[b]	None	Domperidone: NR	NR	p < 0.05[c]	N/A	Low
1 (Bellissant et al, 1997)	RCT	Serious[e]	None	None	Very serious[f]	None	Metoclopramide: 63 (SD 136)	43 (SD 26)	MD 20.00 [-42.20, 82.20]	N/A	Moderate
1 (Tolia et al, 1989)	RCT, crossover	Very serious[g]	None	None	Not assessed[b]	None	Metoclopramide: 25.0 (SD 3.4)	22.4 (SD 2.5)	NS[c]	N/A	Moderate
Duration time of longest episode											
1 (Bines et al, 1992)	RCT	Very serious[d]	None	None	Not assessed[b]	None	Domperidone: 12.6	20.9	NS[c]	N/A	Very low
1 (Carroccio et al, 1993)	RCT	Serious[e]	None	None	Not assessed[b]	None	Domperidone: Median 16 (range 2 to 51)	Median 33.5 (range 8 to 103)	NS[c]	N/A	Low
1 (Bellissant et al, 1997)	RCT	Serious[e]	None	None	Very serious[g]	None	Metoclopramide: Mean 18 (SD 30)	Mean 15 (SD 17)	MD 3.00 [-12.41, 18.41]	N/A	Moderate
Number of reflux episodes > 5 minutes											
1 (Carroccio et	RCT	Serious[e]	None	None	Not assessed	None	Domperidone:	Median 6 (rang	NS[c]	N/A	Low

Quality assessment							Number of Children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision[b]	Other considerations	Prokinetic	Comparator	Relative (95% CI)	Absolute (95% CI)	Quality
al, 1993)							Median 7.5 (range 0 to 16)	e 1 to 20)			
1 (Bellissant et al, 1997)	RCT	Serious[e]	None	None	Serious[h]	None	Metoclopramide: Mean 1.9 (SD 3.0)	Mean 3.0 (SD 3.5)	MD -1.10 [-3.14, 0.94]	N/A	Moderate
1 (Tolia et al, 1989)	RCT, crossover	Very serious[f]	None	None	Not assessed[b]	None	Metoclopramide: 2.6 (SD 0.5)	2.0 (SD 0.3)	NS[c]	N/A	Low
Resolution of oesophagitis – Not reported											
Resolution of faltering growth – Not reported											
Adverse outcomes											
Diarrhea											
1 (Bines et al, 1992)	RCT	Very serious[d]	None	None	None	None	Domperidone: 4	2	NS[c]	N/A	Low
Any adverse event											
1 (Carroccio et al, 1993)	RCT	Serious[e]	None	None	None	None	Domperidone: 0	0	NS[c]	N/A	Moderate
1 (Tolia et al, 1989)	RCT, crossover	Very serious[f]	None	None	None	None	Metoclopramide: 0	0	NS[c]	N/A	Low
Any adverse event leading to discontinuation											
1 (Bellissant et al, 1997)	RCT	Serious[e]	None	None	None	None	Metoclopramide: 3 of 19	1 of 20	NS[c]	N/A	Moderate
Parent reported reduction in infant distress – Not reported											
Improvement in validated reflux questionnaire – Not reported											
Parent satisfaction with this intervention – Not reported											

CI confidence interval; RCT randomised controlled trial; SD standard deviation; NA not applicable; NS not significant;
NS Non significant at $p < 0.05$
N/A Not applicable – could not be calculated on data available

a Method of randomisation and concealment not described. Control group treatment not explained. Reason for unbalanced groups not explained. Poor reporting of data so not all GRADE items could be assessed.
b Data not reported so imprecision could not be calculated
c As reported in the study
d Method of randomisation and concealment not described in detail. Small sample size (<10 per arm). Poor reporting of data so not all GRADE items could be assessed.
e Method of concealment not described in detail. Poor reporting of data so not all GRADE items could be assessed.
f wide confidence intervals - SMD crosses +/- 0.5 effect size
g No washout period between cross-over. Method of randomisation and allocation not explained in detail. Individual periods not reported so reanalysis could not be undertaken.
h wide confidence intervals – SMD crosses -0.5 and 0 effect size

Table 49: GRADE findings for comparison of Proton pump inhibitors compared with H_2 receptor antagonists for managing gastro-oesophageal reflux symptoms

Quality assessment							Number of patients		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**H_2RA**	**PPI**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reduced frequency of overt regurgitation – Not reported											
Reflux measured using oesophageal pH-monitoring or impedance monitoring											
Oesophageal pH <4.0 % improvement from baseline (; measured with: 24-hour combined intraoesophageal and intragastric pH monitor; Better indicated by higher values)											
1 (Cucchiara et al, 1993)	randomised trials	very serious[a,b,c,d, e]	None	Serious[f]	Not assessed[g]	None	Median 59.6 (range 2 to 83.4)	Median 61.9 (range 34 to 99)	NS[h]	-	Very Low
Intragatric pH < 2.0 (minutes) % improvement from baseline (measured with: 24-hour combined intraoesophageal and intragastric pH monitor; Median range of scores: 0-100; Better indicated by higher values)											
1 (Cucchiara et al, 1993)	randomised trials	very serious[a,b,c,d]	None	Serious[f]	Not assessed[g]	None	Median 26.2 (range 0.35 to 95.6)	Median 61.5 (range 7.2 to 98.4)	NS[h]	-	Very Low
Intragatric pH < 4.0 % improvement from baseline (measured with: 24-hour combined intraoesophageal and intragastric pH monitor; range of scores: 0-100; Better indicated by higher values)											
1 (Cucchiara et al, 1993)	randomised trials	very serious[a,b,c,d]	None	Serious[f]	Not assessed[g]	None	Median 22.3 (range 2.1 to 72.8)	Median 29.0 (range 16.4 to 62.8)	NS[h]	-	Very Low
Median intragstric pH % improvement from baseline (Better indicated by higher values)											
1 (Cucchiara	randomised trials	very serious[a,b,]	None	Serious[f]	Not assessed	None	Median 37.4	Median 60.1	P < 0.05[h]	-	Very Low

Quality assessment							Number of patients		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**H_2RA**	**PPI**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
et al, 1993)		[c,d]			[g]		(range 0 to 56.7)	(range 9.3 to 81)			
Resolution of oesophagitis											
Healing of oesophagitis (grade 0 to 2 on histology score) - Ranitidine vs Omeprazole											
1 (Cucchiara et al, 1993)	randomised trials	very serious[a,b]	None	Serious[f]	Very serious[i]	None	8/13 (61.5 %)	9/12 (75%)	RR 0.82 (0.48 to 1.41)	-	Very Low
Resolution of faltering growth – Not reported											
Adverse events requiring discontinuation											
1 (Cucchiara et al, 1993)	randomised trials	very serious[a,b]	none	Serious[f]	None	none	0/13 (0%)	0/12 (0%)	NS[h]	-	Very Low
Parent reported reduction in infant distress – Not reported											
Improvement in validated reflux questionnaire											
60% or more decrease in symptom score - Ranitidine vs. Omeprazole											
1 (Cucchiara et al, 1993)	randomised trials	very serious[a,b]	None	Serious[f]	very serious[f]	None	9/13 (69.2 %)	10/12 (83.3 %)	RR 0.83 (0.53 to 1.29)	-	Very Low
GOR symptoms score (; range of scores: 0-45; Better indicated by lower values)											
1 (Cucchiara et al, 1993)	randomised trials	very serious[a,b,c,d]	None	Serious[f]	Not assessed[g]	None	Median 9.0	Median 9.0	NS[h]	-	Very Low
Parent satisfaction with this intervention – Not reported											

H_2RA H_2 receptor antagonists; PPI protein pump inhibitor; CI confidence interval; NS not significant; RR relative risk; GOR gastro-oesophageal reflux
a High dropout rate
b Method of randomisation not defined
c Small sample size
d Data reported as medians due to skewness
e poor reporting
f Study examining children who had failed previous treatment
g imprecision not assessed
h as reported in study
i Wide confidence intervals crossing no effect and +/- 0.25

J.4 Enteral feeding for GORD

None

J.5 Surgery for GORD

Table 50: GRADE findings for RCT comparison of Open Nissen Fundoplication (ONF) with Laparoscopic Nissen Fundoplication (LNF)

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**ONF**	**LNF**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Cessation (or symptom free days) of overt regurgitation											
Reported as late postoperative recurrence of GORD, n/N, % (exact follow-up time point not reported)											
1 (McHoney et al., 2011)	RCT	Serious[a]	None	Serious	Very serious[b]	Yes[c]	3/18 (16.7 %)	1/14 (7.1 %)	Odds ratio [OR] (95% CI): 2.60 (0.24-28.14)[d]	9.5% (-17.1 to 32.8)[e]	Very low
Adverse outcomes											
Reported as early postoperative incidence of infection, n/N, % (exact follow-up time point not reported)											
1 (McHoney et al., 2011)	RCT	Serious[f]	None	None	Very serious[b]	Yes[c]	1/20 (5%)	3/19 (16 %)	OR (95% CI): 0.28 (0.03-2.97)[d]	-10.8 (-33 to 10.5)[e]	Very low
Reported as patients with complications occurring in the first 30 days after surgery, n/N, %											
1 (Knatten et al., 2012)	RCT	Very serious[g,h,i,j]	None	None	Very serious[b]	Yes[k]	24/44 (55%)	24/44 (55 %)	OR (95% CI): 1 (0.43-2.31)[d]	-	Very low
Reported as postoperative complications (total number of complications) occurring in the first 30 days, n (44 children in each arm)											
1 (Knatten et al., 2012)	RCT	Very serious[g,h,i,j]	None	None	Not assessed[q]	Yes[k]	31	34	NA	-	Low
Reported as postoperative grade I complications l (number of complications) occurring in the first 30 days, n; (44 children in each arm)											
1 (Knatten et al., 2012)	RCT	Very serious[g,h,i,j]	None	None	Not assessed[q]	Yes[k]	11	11	NA	-	Low
Reported as postoperative grade II complications m (number of complications) occurring in the first 30 days, n: (44 children in each arm)											

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**ONF**	**LNF**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
1 (Knatten et al., 2012)	RCT	Very serious[g,h,i,j]	None	None	Not assessed[q]	Yes[k]	18	17	NA	-	Low
Reported as postoperative grade IIIb complications n (number of complications) occurring in the first 30 days, n; (44 children in each arm)											
1 (Knatten et al., 2012)	RCT	Very serious[g,h,i,j]	None	None	Not assessed[q]	Yes[k]	2	6	NA	-	Low
Reported as patients readmitted to hospital because of complications after discharge, n/N, %											
1 (Knatten et al., 2012)	RCT	Very serious[g,h,i]	None	None	Very serious[b]	Yes[k]	11/44 (25%)	12/44 (27%)	OR (95% CI): 0.89 (0.34-2.30)[d]	-	Very low
Reported as early postoperative incidence of gastric paresis, n/N, % (exact follow-up time point not reported)											
1 (McHoney et al., 2011)	RCT	Serious[f]	None	None	Very serious[b]	Yes[c]	2/20 (16%)	3/19 (11%)	OR (95% CI): 1.42 (0.21-9.52)[d]	-5.8% (-28.7 to 16.8)[e]	Very low
Reported as late postoperative incidence of dysphagia, n/N, % (exact follow-up time point not reported)											
1 (McHoney et al., 2011)	RCT	Serious[f]	None	None	Not assessed[q]	Yes[c]	0/16 (0%)	1/16 (6.3%)	-	-6.3% (-28.3 to 13.8)[e]	Moderate
Reported as late postoperative incidence of retching o, n/N, %											
1 (McHoney et al., 2011)	RCT	Very serious[a,p]	None	None	Very serious[b]	Yes[c]	10/18 (55.6)	1/16 (6.3%)	OR (95% CI): 18.75 (2.02 - 173.94)[d]	49.3% (18.3 to 69.8)[e]	Very low
Reported as mean time to full feed in days, mean (CI)											
1 (McHoney et al., 2011)	RCT	None	None	Serious	Not assessed[q]	None	2 (2 to 4)	2 (2 to 4)	P = 0.85[e]	-	Moderate
Resolution of erosive oesophagitis (endoscopic and histologic) – Not reported											
Resolution of reflux symptoms – for example, heartburn, retrosternal or epigastric pain, waterbrash – Not reported											

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	ONF	LNF	Relative (95% CI)	Absolute (95% CI)	Quality
Resolution of faltering growth – Not reported											
Parent reported reduction in infant distress – Not reported											
Resolution of faltering growth – Not reported											
Parent reported reduction in infant distress – Not reported											
Oesophageal reflux measured by oesophageal pH-metry or impedance monitoring – Not reported											
Improvement in validated reflux questionnaire – Not reported											
Parent satisfaction with the intervention – Not reported											

NA-not applicable or not calculable on the data
[a] *Unbalanced drop-out in the LNF arm, reasons not reported*
[b] *Wide confidence interval (CI crosses three zones)*
[c] *The study was not adequately powered for the clinical outcomes*
[d] *NCC-WCH calculation*
[e] *As reported by study authors*
[f] *Unclear whether a valid and reliable method was used to assess outcome*
[g] *No adequate concealment*
[h] *No blinding of the patients or postoperative care staff*
[i] *Unclear whether the groups received same level of care*
[j] *Unclear whether a valid and reliable method was used to assess outcome*
[k] *The study was not adequately powered for its primary outcome reoccurrence and result not reported; for adverse outcomes, a post hoc power calculation was performed*
[l] *Graded according to the Clavien-Dindo classification. Grade I complications do not require pharmacologic treatment, including dislocated gastrostomy, hematoma at the epigastric port site, gastroenteritis, wound infection, and feeding problems*
[m] *Graded according to the Clavien-Dindo classification. Grade II complications require pharmacologic treatment with drugs other than those allowed for Grade I, including airway complications, gastrostomy infection, blood transfusion, urinary tract infection, and gastroenteritis*
[n] *Graded according to the Clavien-Dindo classification. Grade IIIb complications implies that surgical, endoscopic, or radio logic intervention has been performed. Grade IIIb denotes interventions with the use of general anaesthesia, including food impaction, port site hernia/wound rupture, and redo gastrostomy*
[o] *Continued beyond the first six weeks after surgery*
[p] *Subjective outcome reported by parents postoperatively*
[q] *Data was not presented in a way that allowed imprecision to be calculated.*

Table 51: GRADE findings for observational comparison of Laparoscopic Nissen Fundoplication (LNF) with Open Nissen Fundoplication (LNF

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	LNF	ONF	Relative (95% CI)	Absolute (95% CI)	Quality
Adverse outcomes											
Reported as patients undergoing reoperation, n/N (%)											
1 (Diaz et al., 2005)	Retrospective cohort study	Serious[a, b]	None	Serious	Very Serious[c]	None	43/306 (14%)	12/150 (8%)	Odds ratio [OR] (95% CI): 1.88 (0.96-3.68)[d,e]	-	Very low

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**LNF**	**ONF**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Reported as frequency of short-term acute bleeding, n (%),											
1 (Diaz et al., 2005)	Retrospective cohort study	Very serious[a,b,f]	None	None	Not assessed[k]	None	1 (0.8%)	0	P = 0.67[d]	-	Very low
Reported as frequency of short-term acute respiratory problem, n (%)											
1 (Diaz et al., 2005)	Retrospective cohort study	Very serious[a,b,f]	None	None	Not assessed[k]	None	4 (1.3%)	12 (8%)	P = 0.046[d]	-	Very low
Reported as frequency of acute infection, n (%)											
1 (Diaz et al., 2005)	Retrospective cohort study	Very serious[a,b,f]	None	None	Not assessed[k]	None	3 (0.9%)	2 (1.3%)	P = 0.53[d]	-	Very low
Reported as frequency of acute prolonged ileus, n (%)											
1 (Diaz et al., 2005)	Retrospective cohort study	Very serious[a,b,f]	None	None	Not assessed[k]	None	4 (1.3%)	14 (9.3%)	P = 0.0003[d]	-	Very low
Reported as acute other, n (%)											
1 (Diaz et al., 2005)	Retrospective cohort study	Very serious[a,b,f]	None	None	Not assessed[k]	None	6 (1.9)	6 (4%)	P = 0.2[d]	-	Very low
Reported as total frequency of acute complications, n (%)											
1 (Diaz et al., 2005)	Retrospective cohort study	Very serious[a,b,f]	None	None	Not assessed[k]	None	18 (5.9%)	34 (22.7%)	P = 0.0001[d]	-	Very low
Reported as potential risk factors (LNF versus ONF) associated with reoperation											
1 (Diaz et al., 2005)	Retrospective cohort study	Serious[g]	None	None	Very serious[c]	None	-	-	OR (95% CI): 1.68 (0.84-3.3) P = 0.1427[d,g]	-	Very low
Reported as the probability of survival (defined as those who did not require reoperation) and respective reoperation rate at 12 months after initial operation (LNF versus ONF)											
1	Retros	Very	None	Non	Very	None	Survi	Survi	OR	-	Very

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**LNF**	**ONF**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
(Diaz et al., 2005)	pective cohort study	serious[a,b]		e	serious[c]		val/reoperation, n (%): 274 (89.5 %)/ 32 (10.5 %)[h]	val/reoperation, n (%): 144 (96%)/ 6 (4.0 %)[h]	(95% CI): 2.80 (1.15-6.86)[i]		low
Reported as the probability of survival (defined as those who did not require reoperation) and respective reoperation rate at 24 months after initial operation (LNF versus ONF)											
1 (Diaz et al., 2005)	Retrospective cohort study	Very serious[a,b]	None	None	Very serious[c]	None	Survival/reoperation, n (%): 265 (86.6 %)/41 (13.4 %)[h]	Survival/reoperation, n (%): 140 (93.3 %)/10 (6.7 %)[h]	OR (95% CI): 2.17 (1.05-4.45)[i]	-	Very low
Reported as the probability of survival (defined as those who did not require reoperation) and respective reoperation rate at 36 months after initial operation (LNF versus ONF)											
1 (Diaz et al., 2005)	Retrospective cohort study	Very serious[a,b]	None	None	Very serious[c]	None	Survival/reoperation, n (%): 262 (85.6 %)/44 (14.4 %)[h,j]	Survival/reoperation, n (%): 138 (91.9 %)/12 (8.1 %)[h]	OR (95% CI): 1.93 (0.99-3.78)[i]	-	Very low
Cessation (or symptom free days) of overt regurgitation – Not reported											
Resolution of erosive oesophagitis (endoscopic and histologic) – Not reported											
Resolution of reflux symptoms – for example, heartburn, retrosternal or epigastric pain, waterbrash – Not reported											
Resolution of faltering growth – Not reported											
Parent reported reduction in infant distress – Not reported											
Resolution of faltering growth – Not reported											
Parent reported reduction in infant distress – Not reported											

Quality assessment							Number of children		Effect		
Number of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	LNF	ONF	Relative (95% CI)	Absolute (95% CI)	Quality
Oesophageal reflux measured by oesophageal pH-metry or impedance monitoring – Not reported											
Improvement in validated reflux questionnaire – Not reported											
Parent satisfaction with the intervention – Not reported											

NA-not applicable or not calculable on the data
[a] *Intervention groups were not comparable at baseline in terms of undergoing diagnoses*
[b] *Unclear whether there were systematic differences between groups in the care provided*
[c] *Confidence interval crosses three zones*
[d] *As reported by study authors*
[e] *Unadjusted odds ratio*
[f] *Unclear how outcomes were ascertained, diagnosed or verified*
[g] *Odds ratio adjusted for age, gender, neurological impairment, chronic respiratory condition, cardiac disease, prematurity, and reflux alone*
[h] *Percentage as reported by study authors, number of patients calculated by NCC-WCH*
[i] *NCC-WCH calculation*
[j] *Number of patients undergoing reoperation at 36 months different from what previously reported, which was 43, due to discrepancies in percentage reported by study authors and rounding in calculations.*
[k] *Data was not presented in the paper in a format that allowed imprecision to be assessed.*

Table 52: GRADE findings for observational comparison of fundoplication with gastro-jejunal feeding tubes (GJT)

Quality assessment							Number of children		Effect		
Number of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Fundoplication**	**GJT**	**Relative (95% CI)**	**Absolute (95% CI)**	**Quality**
Adverse outcome											
Reported as death a during the following 10 years (median length of follow-up 3.4 years), n/N (%)											
1 (Srivastava et al. 2009)	Retrospective cohort study[d]	Very serious[c,d,e]	None	None	None	None	40/323 (12%)	9/43 (21%)	Hazard ratio [HR], (95% CI): 0.30 (0.12-0.73)[f]	-	Very low
Reported as aspirational pneumonia (AP) during the following years (median length of follow-up 3.4 years), n/N, (%)											
1 (Srivastava et al. 2009)	Retrospective cohort study[b]	Very serious[c,d,e,g]	None	None	Very serious[h]	None	48/323 (15%)	7/43 (16%)	HR (95% CI): 0.71 (0.21-1.69)[i]	-	Very low
Cessation (or symptom free days) of overt regurgitation – Not reported											
Resolution of erosive oesophagitis (endoscopic and histologic) – Not reported											
Resolution of reflux symptoms – for example, heartburn, retrosternal or epigastric pain, waterbrash – Not reported											
Resolution of faltering growth – Not reported											
Parent reported reduction in infant distress – Not reported											
Resolution of faltering growth – Not reported											
Parent reported reduction in infant distress – Not reported											
Oesophageal reflux measured by oesophageal pH-metry or impedance monitoring – Not reported											
Improvement in validated reflux questionnaire – Not reported											
Parent satisfaction with the intervention – Not reported											

[a] *The study was underpowered to detect true differences in this infrequent outcome*
[b] *Study subjects were children with neurologic impairment and GORD*
[c] *Intervention groups were not comparable at baseline in terms of comorbidities*
[d] *Confounders including propensity for surgical indication were adjusted for in analyses, but there still could be other unmeasured confounders*
[e] *Unclear whether the groups received same level of care before and after surgery*
[f] *Adjusted hazard ratio : the Cox model was stratified by age (patients > 1 year versus patients ≤ 1 year) while adjusting for propensity scores for surgery indication and baseline heterogeneities*
[g] *The distinction between AP caused by primary aspiration (e.g., secretions) or secondary aspiration (e.g., refluxed GERD) could not be made because of the nature of the retrospective study*
[h] *Confidence interval crosses three zones*
[i] *Adjusted hazard ratio from Cox model adjusting for propensity scores for surgery indication and baseline heterogeneities*