Supplemental File 3 – Risk of bias assessment of included studies

Study, setting	Study Participation	Study attrition	Prognostic factor measurement	Outcome measurement	Study confounding	Statistical analysis and reporting
El-Serag et al. (12)	MODERATE	HIGH	HIGH	LOW	NA	LOW
Orenstein et al. (18)	HIGH	HIGH	NA	LOW	NA	LOW
Ruigomez et al. (19)	LOW	MODERATE	LOW	MODERATE	NA	MODERATE
Shepherd et al. (20)	HIGH	HIGH	NA	MODERATE	NA	LOW

NA = Not applicable

Author and year of publication	Sheperd 1987			
Date	25-7-2016			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Adequacy of Reporting**	**Risk of bias**
Instructions to assess the risk of each potential bias:	These issues will guide your thinking and judgment about the overall risk of bias within each of the 6 domains. Some 'issues' may not be relevant to the specific study or the review research question. These issues are taken together to inform the overall judgment of potential bias for each of the 6 domains.	Provide comments or text excerpts in the white boxes below, as necessary, to facilitate the consensus process that will follow.	Yes, partial, no or unsure	High, Moderate, or Low (in Summary column) considering all relevant issues
1. Study Participation	**Goal: To judge the risk of selection bias (likelihood that relationship between PF and outcome is different for participants and eligible non-participants).**			
Source of target population	The source population or population of interest is adequately described for key characteristics	Children with GER	YES	
Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (number and type used, e.g., referral patterns in health care)	A series of consecutive cases attending the Royal Children's Hospital in Brisbane. Between Jan 1980 and Dec 1981 clinical investigative and management data were collected according to a precoded protocol.	YES	
Recruitment period	Period of recruitment is adequately described	Jan 1980-Dec 1981	YES	
Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	The Royal Children's Hospital in Brisbane	YES	
Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (e.g., including explicit diagnostic criteria or "zero	Not described	NO	

	time" description).			
Adequate study participation	There is adequate participation in the study by eligible individuals	Not reported	NO	
Baseline characteristics	The baseline study sample (i.e., individuals entering the study) is adequately described for key characteristics.	Gender, age, clinical features	YES	
Summary Study participation	**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias of the observed relationship between PF and outcome.**	Very limited description, likely that patient selection was not performed against inclusion and exclusion criteria. Definition GER not clear.		**~~MODERATE~~ HIGH**
2. Study Attrition	**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and outcome are different for completing and non-completing participants).**			
Proportion of baseline sample available for analysis	Response rate (i.e., proportion of study sample completing the study and providing outcome data) is adequate.	From results it seems that there are follow up data from all patients. Not clear if only children with follow up data were included in study population.	UNSURE	
Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	NA	NA	
Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided	NA	NA	
Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics. There are no important differences between key characteristics and outcomes in participants who completed the study and those who did not.	NA	NA	
Study Attrition Summary	**Loss to follow-up (from baseline sample to study population analyzed)**			**HIGH**

	is not associated with key characteristics (i.e., the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.			
3. Prognostic Factor Measurement	**Goal: To judge the risk of measurement bias related to how PF was measured (differential measurement of PF related to the level of outcome).**			
Definition of the PF	A clear definition or description of 'PF' is provided (e.g., including dose, level, duration of exposure, and clear specification of the method of measurement).	NA, aim was to describe clinical course	NA	
Valid and Reliable Measurement of PF	Method of PF measurement is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall). Continuous variables are reported or appropriate cut-points (i.e., not data-dependent) are used.	NA	NA	
Method and Setting of PF measurement	The method and setting of measurement of PF is the same for all study participants.	NA	NA	
Proportion of data on PF available for analysis	Adequate proportion of the study sample has complete data for PF variable.	NA	NA	
Method used for missing data	Appropriate methods of imputation are used for missing 'PF' data.	NA	NA	
PF Measurement Summary	**PF is adequately measured in study participants to sufficiently limit potential bias.**	NA		**NA**

4. Outcome Measurement	**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up and level and extent of the outcome construct.	Esophagitis, symptoms, complications Limited description of definition and measurement (precoded protocol). Clear description for esophagitis.	PARTIAL	
Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	If measurement was adequate is difficult to assess because of limited detail in reporting	UNSURE	
Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Yes, precoded protocol	YES	
Outcome Measurement Summary	**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**MODERATE**
5. Study Confounding	**Goal: To judge the risk of bias due to confounding (i.e. the effect of PF is distorted by another factor that is related to PF and outcome).**			
Important Confounders Measured	All important confounders, including treatments (key variables in conceptual model), are measured.	NA	NA	
Definition of the confounding factor	Clear definitions of the important confounders measured are provided (e.g., including dose, level, and duration of exposures).	NA	NA	
Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (e.g., may include relevant outside sources of information on measurement properties, also	NA	NA	

	characteristics, such as blind measurement and limited reliance on recall).			
Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	NA	NA	
Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	NA	NA	
Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (e.g., matching for key variables, stratification, or initial assembly of comparable groups). Important potential confounders are accounted for in the analysis (i.e., appropriate adjustment).	NA	NA	
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**	NA		**NA**
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Yes	YES	
Model development strategy	The strategy for model building (i.e., inclusion of variables in the statistical model) is appropriate and is based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	NA	NA	
Reporting of results	There is no selective reporting of results.	**Difficult to assess because of limited detail in Methods section**	NO	

Statistical Analysis and Presentation Summary	**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**LOW**

Author and year of publication	El-Serag 2004			
Date	21-7-2016			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Adequacy of Reporting**	**Risk of bias**
Instructions to assess the risk of each potential bias:	These issues will guide your thinking and judgment about the overall risk of bias within each of the 6 domains. Some 'issues' may not be relevant to the specific study or the review research question. These issues are taken together to inform the overall judgment of potential bias for each of the 6 domains.	Provide comments or text excerpts in the white boxes below, as necessary, to facilitate the consensus process that will follow.	Yes, partial, no or unsure	High, Moderate, or Low (in Summary column) considering all relevant issues
1. Study Participation	**Goal: To judge the risk of selection bias (likelihood that relationship between PF and outcome is different for participants and eligible non-participants).**			
Source of target population	The source population or population of interest is adequately described for key characteristics	GERD in children without comorbid illnesses (neurological deficits, congenital esophageal anomalies, chronic obstructive airway conditions)	YES	
Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (number and type used, e.g., referral patterns in health care)	Administrative and endoscopic database Texas Children's Hospital (all medical diagnoses since 1990); GERD defined as erosive esophagitis (530.1) who underwent upper endoscopic procedure (CPT-4 codes 43234, 43235, 43239) between 1990-1996; children >= 5 years	YES	
Recruitment period	Period of recruitment is adequately described	1990-1996 (data from database)	YES	
Place of recruitment	Place of recruitment (setting and geographic location) are adequately described	Texas Children's Hospital	YES	
Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (e.g., including explicit diagnostic criteria or "zero time" description).	Inclusion: children >=5 yrs with GERD (erosive esophagitis - ICD 9 code 530.1) who underwent upper endoscopic procedure Exclusion: Cerebral palsy, mental retardation,	YES	

		tracheoesophageal fistula, congenital esophageal stenosis, severe comorbid illness such as solid organ or bone marrow transplant, cancer or cystic fibrosis; residence outside Houston area		
Adequate study participation	There is adequate participation in the study by eligible individuals	222 potentially eligible based on database; inclusion of children of >=5 (target pop = all children)	YES	
Baseline characteristics	The baseline study sample (i.e., individuals entering the study) is adequately described for key characteristics.	Age, gender, racial/ethnic distribution, education, marital status, BMI, smoking, excessive alcohol, treatment => characteristics are reported for interviewed participants (=end of follow up)	NO	
Summary Study participation	**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias of the observed relationship between PF and outcome.**	Retrospective identification of cohort Inclusion of children of >=5 (target pop = all children), no info on baseline characteristics		**MODERATE**
2. Study Attrition	**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and outcome are different for completing and non-completing participants).**			
Proportion of baseline sample available for analysis	Response rate (i.e., proportion of study sample completing the study and providing outcome data) is adequate.	222 potentially eligible based on database; 9 died, 6 had comorbid disease. 127 (61%) of the 207 eligible participants declined to participate => high rate of non-participation (this is lost to follow up)	YES	
Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	Limited: no stat sign differences in age, gender and race (data not shown) Sensitivity analysis in which prevalence of GERD symptoms was calculated incl all eligible children and assuming they were symptom free => also analysis needed assuming all non-respondents had GERD symptoms (best/worst case scenario)	NO	
Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided	Not reported	NO	
Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics.	Not reported	NO	

	There are no important differences between key characteristics and outcomes in participants who completed the study and those who did not.			
Study Attrition Summary	**Loss to follow-up (from baseline sample to study population analyzed) is not associated with key characteristics (i.e., the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**	Very high dropout rate, unclear how sample analyzed differs from baseline sample		**HIGH**
3. Prognostic Factor Measurement	**Goal: To judge the risk of measurement bias related to how PF was measured (differential measurement of PF related to the level of outcome).**			
Definition of the PF	A clear definition or description of 'PF' is provided (e.g., including dose, level, duration of exposure, and clear specification of the method of measurement).	Aim of study was to determine if GERD in childhood persists in adolescence and young adulthood. Age (at fu?); sex; race; family history; BMI (at fu); age onset GERD	PARTIAL	
Valid and Reliable Measurement of PF	Method of PF measurement is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall). Continuous variables are reported or appropriate cut-points (i.e., not data-dependent) are used.	BMI self-report, age onset GERD based on database and self-report	PARTIAL	
Method and Setting of PF measurement	The method and setting of measurement of PF is the same for all study participants.	Yes	YES	

Proportion of data on PF available for analysis	Adequate proportion of the study sample has complete data for PF variable.	Yes	YES	
Method used for missing data	Appropriate methods of imputation are used for missing 'PF' data.	NA	NA	
PF Measurement Summary	**PF is adequately measured in study participants to sufficiently limit potential bias.**	Age and BMI measured only at fu, age at onset for some participants (% not reported) based on self-report at follow up assessment		**HIGH**
4. Outcome Measurement	**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up and level and extent of the outcome construct.	GERQ measured symptoms: at least monthly symptoms (any GERD) or at least weekly symptoms (frequent GERD); heartburn; acid regurgitation; symptom severity on 4-point scale Erosive esophagitis BE	YES	
Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	Detailed definitions	YES	
Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Yes	YES	
Outcome Measurement Summary	**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**LOW**
5. Study Confounding	**Goal: To judge the risk of bias due to confounding (i.e. the effect of PF is distorted by another factor that is related to PF and outcome).**			
Important Confounders	All important confounders, including	Not applicable, only univariate analyses	NA	

Measured	treatments (key variables in conceptual model: LIST), are measured.			
Definition of the confounding factor	Clear definitions of the important confounders measured are provided (e.g., including dose, level, and duration of exposures).	Not applicable, only univariate analyses	NA	
Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	Not applicable, only univariate analyses	NA	
Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	Not applicable, only univariate analyses	NA	
Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	Not applicable, only univariate analyses	NA	
Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (e.g., matching for key variables, stratification, or initial assembly of comparable groups). Important potential confounders are accounted for in the analysis (i.e., appropriate adjustment).	Not applicable, only univariate analyses	NA	
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**NA**

6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.		YES	
Model development strategy	The strategy for model building (i.e., inclusion of variables in the statistical model) is appropriate and is based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	Not applicable, only univariate analyses	NA	
Reporting of results	There is no selective reporting of results.	Selective reporting not likely	YES	
Statistical Analysis and Presentation Summary	**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**LOW**

Author and year of publication	Orenstein 2006			
Date	22-7-2016			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Adequacy of Reporting**	**Risk of bias**
Instructions to assess the risk of each potential bias:	These issues will guide your thinking and judgment about the overall risk of bias within each of the 6 domains. Some 'issues' may not be relevant to the specific study or the review research question. These issues are taken together to inform the overall judgment of potential bias for each of the 6 domains.	Provide comments or text excerpts in the white boxes below, as necessary, to facilitate the consensus process that will follow.	Yes, partial, no or unsure	High, Moderate, or Low (in Summary column) considering all relevant issues
1. Study Participation	**Goal: To judge the risk of selection bias (likelihood that relationship between PF and outcome is different for participants and eligible non-participants).**			
Source of target population	The source population or population of interest is adequately described for key characteristics	Children with esophagitis	YES	
Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (number and type used, e.g., referral patterns in health care)	The “parent” study was a randomized, double blind, placebo controlled trial evaluating a histamine-2 receptor antagonist, a prokinetic agent, or both as therapy for symptomatic reflux esophagitis in 100 infants younger than 12 months of age. => trial population probably limits generalizability	YES	
Recruitment period	Period of recruitment is adequately described	Between July 1, 1994, and October 13, 1999, infants between the ages of 28 and 366 days (corrected gestational age) who were referred to the Pediatric Gastroenterology Division of Children’s Hospital of Pittsburgh because of a clinical suspicion of GERD, who did not respond symptomatically to a 2-wk trial of conservative therapy, and who demonstrated histologic morphometric reflux esophagitis on a distal esophageal suction biopsy, were recruited to participate in the two-by-two factorial pharmacotherapy study, which had	YES	

		been approved by the Children's Hospital of Pittsburgh Institutional Review Board.		
Place of recruitment	Place of recruitment (setting and geographic location) are adequately described		YES	
Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (e.g., including explicit diagnostic criteria or "zero time" description).	Exclusion criteria: histologic evidence for infectious or eosinophilic esophagitis; gastrointestinal structural abnormalities or prior surgery; unacceptable risk for dual placebo (history of severe apparent life-threatening event or hematemesis); unacceptable risk from the study drugs; or inability to complete the study as predicted by the investigators. Inclusion: not meeting exclusion criteria	YES	
Adequate study participation	There is adequate participation in the study by eligible individuals	24 randomized to placebo, 19 returned for visit at month 2 and are included in present study => n=24 is baseline sample	UNSURE	
Baseline characteristics	The baseline study sample (i.e., individuals entering the study) is adequately described for key characteristics.	Table 1	YES	
Summary Study participation	**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias of the observed relationship between PF and outcome.**	Probably selective sample		**HIGH**
2. Study Attrition	**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and outcome are different for completing and non-completing participants).**			
Proportion of baseline sample available for analysis	Response rate (i.e., proportion of study sample completing the study and providing outcome data) is adequate.	N=24 placebo group = baseline sample, 5 did not return for follow up, 3 withdrawn: available for analysis n=16 8/24 = 33% dropout rate (= high)	YES	
Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.		YES	

Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided	Failed to return and withdrawn => reasons unknown	PARTIAL	
Outcome and prognostic factor information on those lost to follow-up	Participants lost to follow-up are adequately described for key characteristics. There are no important differences between key characteristics and outcomes in participants who completed the study and those who did not.	Only two characteristics, initial weight (p = 0.01) and "arching" (p = 0.02) differed significantly between those two groups [return for fu, y/n] , with greater weight and more arching in the follow-up babies, although the small sample size and multiple comparisons make type II and type I errors, respectively, possible.	YES	
Study Attrition Summary	**Loss to follow-up (from baseline sample to study population analyzed) is not associated with key characteristics (i.e., the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**	High dropout rate and differences between follow-up and not follow up babies [at study entrance],		**HIGH**
3. Prognostic Factor Measurement	**Goal: To judge the risk of measurement bias related to how PF was measured (differential measurement of PF related to the level of outcome).**			
Definition of the PF	A clear definition or description of 'PF' is provided (e.g., including dose, level, duration of exposure, and clear specification of the method of measurement).	An in-depth examination of comprehensive data on 19 infants repetitively evaluated while participating in the placebo arm of a 12-month pharmacotherapy study thus provides a unique opportunity to describe in detail the natural history of both symptoms and histology of infantile esophagitis. PF not aim of study.	NA	
Valid and Reliable Measurement of PF	Method of PF measurement is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).		NA	

	Continuous variables are reported or appropriate cut-points (i.e., not data-dependent) are used.			
Method and Setting of PF measurement	The method and setting of measurement of PF is the same for all study participants.		NA	
Proportion of data on PF available for analysis	Adequate proportion of the study sample has complete data for PF variable.		NA	
Method used for missing data	Appropriate methods of imputation are used for missing 'PF' data.		NA	
PF Measurement Summary	**PF is adequately measured in study participants to sufficiently limit potential bias.**			**NA**
4. Outcome Measurement	**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up and level and extent of the outcome construct.	Follow up was 12 months or earlier if children needed medication (‘rescue’, n=6) or were withdrawn (n=3) Symptoms (I-GERQ and parent global score) Esophagitis (by suction biopsy)	YES	
Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	Yes	YES	
Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Yes	YES	
Outcome Measurement Summary	**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**			**LOW**

5. Study Confounding	**Goal: To judge the risk of bias due to confounding (i.e. the effect of PF is distorted by another factor that is related to PF and outcome).**			
Important Confounders Measured	All important confounders, including treatments (key variables in conceptual model: LIST), are measured.	NA	NA	
Definition of the confounding factor	Clear definitions of the important confounders measured are provided (e.g., including dose, level, and duration of exposures).	NA	NA	
Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).	NA	NA	
Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.	NA	NA	
Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.	NA	NA	
Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (e.g., matching for key variables, stratification, or initial assembly of comparable groups). Important potential confounders are accounted for in the analysis (i.e., appropriate adjustment).	NA	NA	
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the**			**NA**

	relationship between PF and outcome.			
6. Statistical Analysis and Reporting	**Goal: To judge the risk of bias related to the statistical analysis and presentation of results.**			
Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the analysis.	Yes	YES	
Model development strategy	The strategy for model building (i.e., inclusion of variables in the statistical model) is appropriate and is based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	NA	NA	
Reporting of results	There is no selective reporting of results.	No selective reporting	YES	
Statistical Analysis and Presentation Summary	**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**LOW**

Author and year of publication	Ruigomez 2010a			
Date	25-7-2016			
Biases	**Issues to consider for judging overall rating of "Risk of bias"**	**Study Methods & Comments**	**Adequacy of Reporting**	**Risk of bias**
Instructions to assess the risk of each potential bias:	These issues will guide your thinking and judgment about the overall risk of bias within each of the 6 domains. Some 'issues' may not be relevant to the specific study or the review research question. These issues are taken together to inform the overall judgment of potential bias for each of the 6 domains.	Provide comments or text excerpts in the white boxes below, as necessary, to facilitate the consensus process that will follow.	Yes, partial, no or unsure	High, Moderate, or Low (in Summary column) considering all relevant issues
1. Study Participation	**Goal: To judge the risk of selection bias (likelihood that relationship between PF and outcome is different for participants and eligible non-participants).**			
Source of target population	The source population or population of interest is adequately described for key characteristics	Children and adolescents with GERD managed in primary care	YES	
Method used to identify population	The sampling frame and recruitment are adequately described, including methods to identify the sample sufficient to limit potential bias (number and type used, e.g., referral patterns in health care)	The source population comprised individuals recorded in THIN [database]who were registered with a collaborating primary care practice (PCP) for at least 1 year before the start of the study period (January 1, 2000) (Figure 1). All cases with a diagnosis of GERD but no recorded reflux esophagitis or other esophageal injury were followed from the day after the initial GERD diagnosis date (index date) until the earliest occurrence of one of the following endpoints: a new diagnosis of reflux esophagitis or a GERD-related esophageal complication (including esophageal ulcer, esophageal stricture, Barrett's esophagus and esophageal cancer), death or the end of the follow-up period (November 30, 2008).	YES	
Recruitment period	Period of recruitment is adequately described	Yes, see above	YES	
Place of recruitment	Place of recruitment (setting and	United Kingdom	YES	

	geographic location) are adequately described			
Inclusion and exclusion criteria	Inclusion and exclusion criteria are adequately described (e.g., including explicit diagnostic criteria or "zero time" description).	Inclusion: All cases with a diagnosis of GERD but no recorded reflux esophagitis or other esophageal injury Excluded: For the present study we further excluded individuals with any record of reflux esophagitis or other esophageal injury (e.g. ulcer, stricture or Barrett's esophagus) at their initial diagnosis. Pregnant girls were also excluded.	YES	
Adequate study participation	There is adequate participation in the study by eligible individuals	Database: participation rate not reported Incident GERD without esophagitis: n=1,242	UNSURE	
Baseline characteristics	The baseline study sample (i.e., individuals entering the study) is adequately described for key characteristics.	Limited. Age, gender, diagnosis (heartburn or reflux)	PARTIAL	
Summary Study participation	**The study sample represents the population of interest on key characteristics, sufficient to limit potential bias of the observed relationship between PF and outcome.**			**LOW**
2. Study Attrition	**Goal: To judge the risk of attrition bias (likelihood that relationship between PF and outcome are different for completing and non-completing participants).**			
Proportion of baseline sample available for analysis	Response rate (i.e., proportion of study sample completing the study and providing outcome data) is adequate.	Complete data of all included participants. [only participants with complete fu included?]	YES	
Attempts to collect information on participants who dropped out	Attempts to collect information on participants who dropped out of the study are described.	NA	NA	
Reasons and potential impact of subjects lost to follow-up	Reasons for loss to follow-up are provided	NA	NA	
Outcome and prognostic factor information on those lost to	Participants lost to follow-up are adequately described for key	NA	NA	

follow-up	characteristics. There are no important differences between key characteristics and outcomes in participants who completed the study and those who did not.			
Study Attrition Summary	**Loss to follow-up (from baseline sample to study population analyzed) is not associated with key characteristics (i.e., the study data adequately represent the sample) sufficient to limit potential bias to the observed relationship between PF and outcome.**	NA	NA	**MODERATE**
3. Prognostic Factor Measurement	**Goal: To judge the risk of measurement bias related to how PF was measured (differential measurement of PF related to the level of outcome).**			
Definition of the PF	A clear definition or description of 'PF' is provided (e.g., including dose, level, duration of exposure, and clear specification of the method of measurement).	Sex, age at initial GERD diagnosis, visit to PCP in previous year, initial diagnosis based on Read codes (database), use of acid suppressants (within 30 days of initial diagnosis) No explanation why these factors are important.	NO	
Valid and Reliable Measurement of PF	Method of PF measurement is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall). Continuous variables are reported or appropriate cut-points (i.e., not data-dependent) are used.	Cut-of points unclear for age. Measurement quality depends on quality database. Validated for pharmacoepidemiology, no further information	UNSURE	
Method and Setting of PF measurement	The method and setting of measurement of PF is the same for all study participants.	Yes (based on database)	YES	

Proportion of data on PF available for analysis	Adequate proportion of the study sample has complete data for PF variable.	Yes (all)	YES	
Method used for missing data	Appropriate methods of imputation are used for missing 'PF' data.	NA	NA	
PF Measurement Summary	**PF is adequately measured in study participants to sufficiently limit potential bias.**			**LOW**
4. Outcome Measurement	**Goal: To judge the risk of bias related to the measurement of outcome (differential measurement of outcome related to the baseline level of PF).**			
Definition of the Outcome	A clear definition of outcome is provided, including duration of follow-up and level and extent of the outcome construct.	Esophagitis (not further defined) Follow up: mean 4 years, sd 1.9 years	NO	
Valid and Reliable Measurement of Outcome	The method of outcome measurement used is adequately valid and reliable to limit misclassification bias (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and confirmation of outcome with valid and reliable test).	Unclear	NO	
Method and Setting of Outcome Measurement	The method and setting of outcome measurement is the same for all study participants.	Yes	YES	
Outcome Measurement Summary	**Outcome of interest is adequately measured in study participants to sufficiently limit potential bias.**	Difficult to assess because lack of detailed reporting.		**MODERATE**
5. Study Confounding	**Goal: To judge the risk of bias due to confounding (i.e. the effect of PF is distorted by another factor that is related to PF and outcome).**			
Important Confounders Measured	All important confounders, including treatments (key variables in conceptual model: LIST), are measured.	NA – univariate results	NA	

Definition of the confounding factor	Clear definitions of the important confounders measured are provided (e.g., including dose, level, and duration of exposures).			
Valid and Reliable Measurement of Confounders	Measurement of all important confounders is adequately valid and reliable (e.g., may include relevant outside sources of information on measurement properties, also characteristics, such as blind measurement and limited reliance on recall).			
Method and Setting of Confounding Measurement	The method and setting of confounding measurement are the same for all study participants.			
Method used for missing data	Appropriate methods are used if imputation is used for missing confounder data.			
Appropriate Accounting for Confounding	Important potential confounders are accounted for in the study design (e.g., matching for key variables, stratification, or initial assembly of comparable groups). Important potential confounders are accounted for in the analysis (i.e., appropriate adjustment).			
Study Confounding Summary	**Important potential confounders are appropriately accounted for, limiting potential bias with respect to the relationship between PF and outcome.**			**NA**
<u>6. Statistical Analysis and Reporting</u>	**<u>Goal: To judge the risk of bias related to the statistical analysis and presentation of results.</u>**			
Presentation of analytical strategy	There is sufficient presentation of data to assess the adequacy of the	Only p-value levels, no CI or ORs	PARTIAL	

	analysis.			
Model development strategy	The strategy for model building (i.e., inclusion of variables in the statistical model) is appropriate and is based on a conceptual framework or model. The selected statistical model is adequate for the design of the study.	NA	NA	
Reporting of results	There is no selective reporting of results.	Risk factors not defined in Methods section	UNSURE	
Statistical Analysis and Presentation Summary	**The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid or spurious results.**			**MODERATE**