Cochrane Australia

Systematic Review of Safe Level of Gluten for People with Coeliac Disease

FINAL REPORT
5 February 2016

Trusted evidence.
Informed decisions.
Better health.

Table of Contents

Review team

This was review prepared by Jane Reid, Kelly Allen and Steve McDonald. Additional support was provided by Madeleine Hill and Sue Brennan. All contributors declare they have no relevant declarations of interest. We acknowledge Jason Tye-Din for his help in providing clarification of clinical issues.

Abstract

Background

Coeliac disease is a systemic autoimmune disease characterised by histological abnormalities of the small bowel mucosa. The mainstay of treatment of coeliac disease is a strict lifelong 'gluten-free' diet, although total avoidance of gluten can be difficult and there is individual variability in tolerance to gluten.

Objectives

To undertake a systematic review to determine if there is a threshold level of intake of gluten that is safe for people with coeliac disease to consume.

Search methods

We searched MEDLINE, Embase, Cochrane Library, Scopus, Web of Science, trial registers, grey literature sources and key websites for relevant systematic reviews and primary studies. We conducted forward citation searches of key systematic reviews and eligible studies, and checked reference lists. Searches were conducted in November and December 2015.

Selection criteria

We selected studies that assessed the effects of different amounts of gluten on small-bowel histology, serology and clinical symptoms in adults or children with confirmed coeliac disease. Dietary intervention or gluten exposure were of any duration. Gluten source included gliadin and wheat-starch based gluten-free products but excluded oats. Randomised trials, non-randomised, cohort and cross-sectional studies that investigated safe levels of gluten were included. We excluded challenge studies that assessed mucosal change using high levels of gluten.

Data collection and analysis

Two reviewers were involved in study selection, data collection and risk of bias assessment. Discrepancies were resolved by discussion within the review team. In the absence of quantitative analysis, results were synthesised narratively and grouped according to study design. The GRADE principles were used in considering the overall quality of the evidence.

Main results

We included 18 studies involving 1754 participants; nine were intervention studies (including five randomised trials) and nine were observational. Studies were very heterogeneous with respect to the type, amount and duration of exposure to gluten, populations studied, how outcomes were assessed and types of study design included. The risk of bias of the randomised trials was judged to be low or unclear, and high for the non-randomised studies. Only one study aimed to establish the safety threshold of prolonged exposure to small amounts of gluten and thus provided the most direct evidence for this review. This was a placebo-controlled randomised trial of 39 patients that found mucosal damage occurred in patients receiving 50 mg gluten/day (percentage change in villous height/crypt depth: -20% (95% CI: -22% to -13%) 13 participants; moderate quality evidence) but was of uncertain effect in those receiving 10 mg gluten/day (-1% (95% CI: -18% to 68%) 13 participants). From the limited randomised and other evidence available, mucosal deterioration was commonly present in patients receiving about 50 mg gluten/day but robust evidence was lacking for the effects of gluten consumption in the critical range of 2 to 10 mg gluten/day. The evidence in relation to the effect on clinical symptoms was inconsistent and was difficult to interpret because of the uncertainty and variability over how this outcome was measured across studies.

Authors' conclusions

There is moderate evidence based on one randomised trial that patients with coeliac disease develop mucosal damage following consumption of 50 mg gluten/day, but the evidence of effect is uncertain for lower levels of consumption. The studies highlight the individual variability in tolerance to gluten and the difficulty this raises for setting a safe threshold. The overall evidence included in the review is

characterised by a high level of heterogeneity and consequently most studies provide only low quality indirect evidence of an association between gluten consumption and mucosal change. In the absence of larger randomised trials that compare small amounts of gluten intake, the current evidence precludes establishing a definitive threshold level of gluten that is safe for all people with coeliac disease to consume.

1 Background

1.1 Description of the condition and treatment

Coeliac disease is a systemic autoimmune disease characterised by histological abnormalities of the small bowel mucosa and intestinal malabsorption, triggered by gluten ingestion in genetically susceptible individuals (Lebwohl 2015). People with coeliac disease have an increased risk of cancer, including small intestinal adenocarcinoma and non-Hodgkin's lymphoma, and experience a range of clinical symptoms, including gastrointestinal symptoms, metabolic bone disease, infertility, and many other manifestations (Lebwohl 2015). Prevalence of the condition has risen in recent times and is currently estimated to be 1% worldwide, although most people with the condition remain undiagnosed (Hall 2009; Lebwohl 2015).

The mainstay of treatment of coeliac disease is a strict lifelong gluten-free diet. Adherence to a gluten-free diet, for most individuals, results in histological and clinical remission, and reduced risk of gastrointestinal malignancies, osteoporosis, infertility, fatigue and depression (Hall 2009). However, total avoidance of gluten can be difficult to achieve for a number of reasons, including contamination of "gluten-free" products. Therefore, the term "gluten-free" generally refers to a level of gluten that is meant to be harmless when consumed indefinitely, rather than to a total absence of gluten (Akobeng 2008).

1.2 Rationale for undertaking this review

Gluten-free labelling laws in Australia are based on a "no-detectable" non-fixed gluten limit, as opposed to the fixed value adopted in Europe (Codex) and the USA (FDA) (defined as no more than 20ppm of gluten). In Australia the sensitivity of testing is now able to detect as little as 3ppm gluten, which means food labelled "gluten-free" has to be under that limit to be considered gluten-free. However, as gluten food testing is becoming increasingly sensitive, there is concern that in the future the availability of "gluten-free food", the only treatment for people with coeliac disease, may be compromised.

Coeliac Australia took an initial position to support a proposed application by the Australian Food and Grocery Council (AFGC) to Food Standards Australia New Zealand (FSANZ) to bring the Australian gluten-free standard in line with the rest of the world and support a fixed limit of less than 20ppm. Concerns were raised regarding the safety of 20ppm for all people with coeliac disease.

The Medical Advisory Committee (MAC) of Coeliac Australia concluded that 20ppm is likely to be safe for most patients with coeliac disease. It also noted that the medical literature was not robust, and it was not possible to guarantee the safety of a 20ppm limit for all patients with coeliac disease. As a result, Coeliac Australia withdrew in principle support for the AFGC application pending a decision about Coeliac Australia's position on the issue.

In the interests of being thorough and transparent, Coeliac Australia has commissioned this external systematic review to determine whether a level of gluten can be identified which represents a safe threshold of intake that persons with coeliac disease can safely consume.

2 Objectives

To undertake a systematic review of studies to answer the following question: For people with coeliac disease, is there a threshold level of intake of gluten that is safe to consume?

3 Methods

3.1 Criteria for considering studies for this review

3.1.1 Types of participants

Adults and/or children with confirmed diagnosis of coeliac disease. The method of confirmation or use of a particular consensus definition of coeliac disease was noted. Studies involving populations of non-coeliac gluten sensitivity were excluded.

3.1.2 Types of interventions/exposures

Diets that compare differing amounts of gluten (whether estimated or measured). Comparison groups could include a no-gluten group. Studies that included gliadin or wheat-starch-based gluten-free products as the intervention, or part of the intervention, were included. Studies (or study arms) in which oats were part of the comparison diets were excluded. No minimum was applied to the duration of the dietary intervention or gluten exposure.

3.1.3 Types of outcome measures

Primary outcome

1. Histological: changes in small bowel histology, as reported in the individual study.

Secondary outcomes

1. Serological: changes to serum antibodies.
2. Symptomatic: changes in clinical and non-clinical symptoms, including, but not limited to, intestinal disturbances, fatigue, mood and quality of life.

3.1.4 Types of studies

Because of the expected limited evidence available from clinical trials, the inclusion criteria was intentionally broad and comprised randomised trials, non-randomised studies, cohort studies and cross-sectional studies. Single case reports were excluded.

To be eligible for inclusion, the purpose of the study needed to include investigating levels of intake of gluten that are safe for people with coeliac disease to consume. This last aspect of the eligibility criteria was interpreted liberally since very few studies specified the aim of establishing a safety threshold. The decision to include studies was based on the levels of gluten being evaluated. We included micro-challenge studies but excluded gluten challenge studies (where the aim was to assess mucosal changes) that only involved high doses of gluten (generally above 2 g gluten/day).

3.2 Search methods for identification of studies

We developed a comprehensive strategy to identify relevant studies, including searches of traditional bibliographic databases, grey literature sources, prospective registers of trials and reviews, websites of key organisations and forward citation searching. All studies, irrespective of publication status, format or date were considered for inclusion. The only restriction we applied was to exclude studies reported in languages other than English.

3.2.1 Search terms

We used the Yale MeSH Analyzer (http://mesh.med.yale.edu) to identify relevant Medical Subject Headings (MeSH) for MEDLINE by retrieving the PubMed IDs of the included studies from the Akobeng review. The Yale MeSH Analyzer sorts and groups MeSH terms alphabetically for ease of scanning across the citations entered.

In addition to obvious terms, such as ‘Celiac Disease’, ‘Glutens’ and ‘Diet, Gluten-Free’, we included terms that captured aspects of the outcome, such as ‘Toxicity Tests’ and ‘Intestinal Mucosa/pathology’. Based on our assessment of the Akobeng studies, we also included ‘adverse effects’ as a floating subheading; this retrieves any citation in which adverse effects has been applied to any MeSH term as a subheading.

As well as identifying relevant MeSH terms, we included free-text terms, ensuring variant spellings, truncation and synonyms were applied appropriately. To minimise the possibility of missing potentially

relevant studies, we did not restrict our search to particular publication types or study designs. The only restriction we applied was to limit to studies in humans published in English.

3.2.2 Bibliographic and grey literature databases

We searched MEDLINE (via Ovid), Embase (via Ovid), Cochrane Library, Scopus and Web of Science. The relevant MeSH terms were matched to the relevant Emtree terms in Embase. Searches were not restricted by date of publication. We supplemented these searches with searches of PubMed (limited to the subset of citations not indexed in MEDLINE) and the grey literature databases, OpenGrey and Grey Literature Report. Searches were conducted on 11 November 2015. The search strategies and retrieval for each database are presented in **Appendix 1**.

3.2.3 Organisations

Websites of the following organisations were searched for relevant guidelines, systematic reviews and individual studies. The purpose of searching these organisational websites was to identity reports that may have referenced or referred to relevant primary studies. These searches were conducted on 16 December 2015.

Name	Website
Agency for Healthcare Research and Quality (AHRQ)	www.ahrq.gov
Australian Food and Grocery Council	www.afgc.org.au
CODEX International Food Standards	www.codexalimentarius.org
Food Standards Agency (UK)	www.food.gov.uk
Food Standards Australia New Zealand	www.foodstandards.gov.au
National Institutes for Healthcare and Excellence	www.nice.org.uk
US Food and Drug Administration (FDA)	www.fda.gov
World Health Organization (WHO)	www.who.int/en

3.2.4 Guidelines, systematic reviews and clinical trials

We searched the following guidelines, systematic reviews and clinical trials websites for references to any relevant published, unpublished or ongoing studies, including systematic reviews. These searches were conducted between 10-15 December 2015.

Name	Website
Clinical practice guidelines portal (NHMRC)	www.clinicalguidelines.gov.au
National guideline clearing house	www.guideline.gov
PROSPERO (prospective register of SRs)	www.crd.york.ac.uk/PROSPERO/
WHO International Clinical Trials Registry Platform (ICTRP)	www.who.int/ictrp/
ClinicalTrials.gov	www.clinicaltrials.gov

3.2.5 Other sources

We used Scopus, Web of Science and Google Scholar to undertake a forward citation search of the systematic review by Akobeng, and Web of Science alone to undertake forward citation searches of the 13 studies included in the review. These searches were conducted on 31 October 2015.

The reference lists of all eligible studies and relevant systematic reviews were checked for additional studies. Finally, we checked the references of potentially relevant literature submitted to us by a member of the Medical Advisory Committee of Coeliac Australia.

3.3 Data collection and analysis

3.3.1 Selection of studies

Citations identified from the literature searches and citation checking were imported to EndNote and duplicates removed. One reviewer (JR) initially screened the titles and abstracts to exclude those citations that were clearly outside the scope of the review. The full-text of the remaining citations were obtained and the citations re-screened by the same reviewer using a pre-designed eligibility form based on the inclusion criteria. Citations excluded at this stage were categorised according the reason for exclusion

(**Figure 1**). A second reviewer (KA) verified the inclusion/exclusion decisions. Disagreements about study eligibility were resolved by discussion among the review team.

3.3.2 Data extraction and management

For each included study, data were extracted by one reviewer (JR) in structured summary tables. A second reviewer (KA or SM) verified the accuracy of the data extraction and confirmed any manual estimations of levels of gluten made by JR. Disagreements were resolved by discussion among the review team.

We collected information on the following characteristics:

- design and location
- aim
- participants (age, gender and duration of diagnosis)
- duration of diet/exposure
- exposure (amount and type of gluten)
- outcomes
- results

3.3.3 Assessment of risk of bias of included studies

We used the Cochrane risk of bias tool to assess the randomised trials and the new ROBIS tool (Whiting 2015) to assess the systematic reviews. For the non-randomised intervention studies we used a modified version of the Cochrane Effective Practice and Organisation of Care (EPOC) Group risk of bias tool. Because of the well-recognised limitations of observational studies to address questions of causation when assessing the effects of different interventions, we did not conduct any risk of bias assessment on the observational studies. In the case of some studies that did not contain sufficient information on the risk of bias domains, we sought clarification and further information from the authors.

3.3.4 Data synthesis and overall quality of the evidence

If we had identified sufficient studies with the same study design and comparable interventions measured outcomes in similar ways, we planned to conduct a meta-analysis. Because of the heterogeneity with respect to study design, type of intervention/exposure and outcomes measured, we tabulated data and synthesised the results narratively. To address the concern over the inclusion of different types of study design in the review, we reported the results by study design, grouping the randomised evidence separately from the non-randomised evidence and observational studies.

We used the principles of GRADE criteria as a framework to assess the overall quality, completeness and applicability of the evidence (Balshem 2011). Based on this assessment, the body of evidence was rated as high, moderate, low or very low.

3.3.5 Subgroup analysis

In the protocol we had planned to analyse the results for adults and children separately, where this was feasible. Although there were studies that included only adults or only children, several studies included both children/adolescents and adults. In the studies with mixed populations it wasn't possible to analyse the results separately by population group. We therefore chose not to analyse the results separately but to include details of the study populations in the tables and comment where appropriate in the text.

4. Results

4.1 Search results

4.1.1 Bibliographic databases

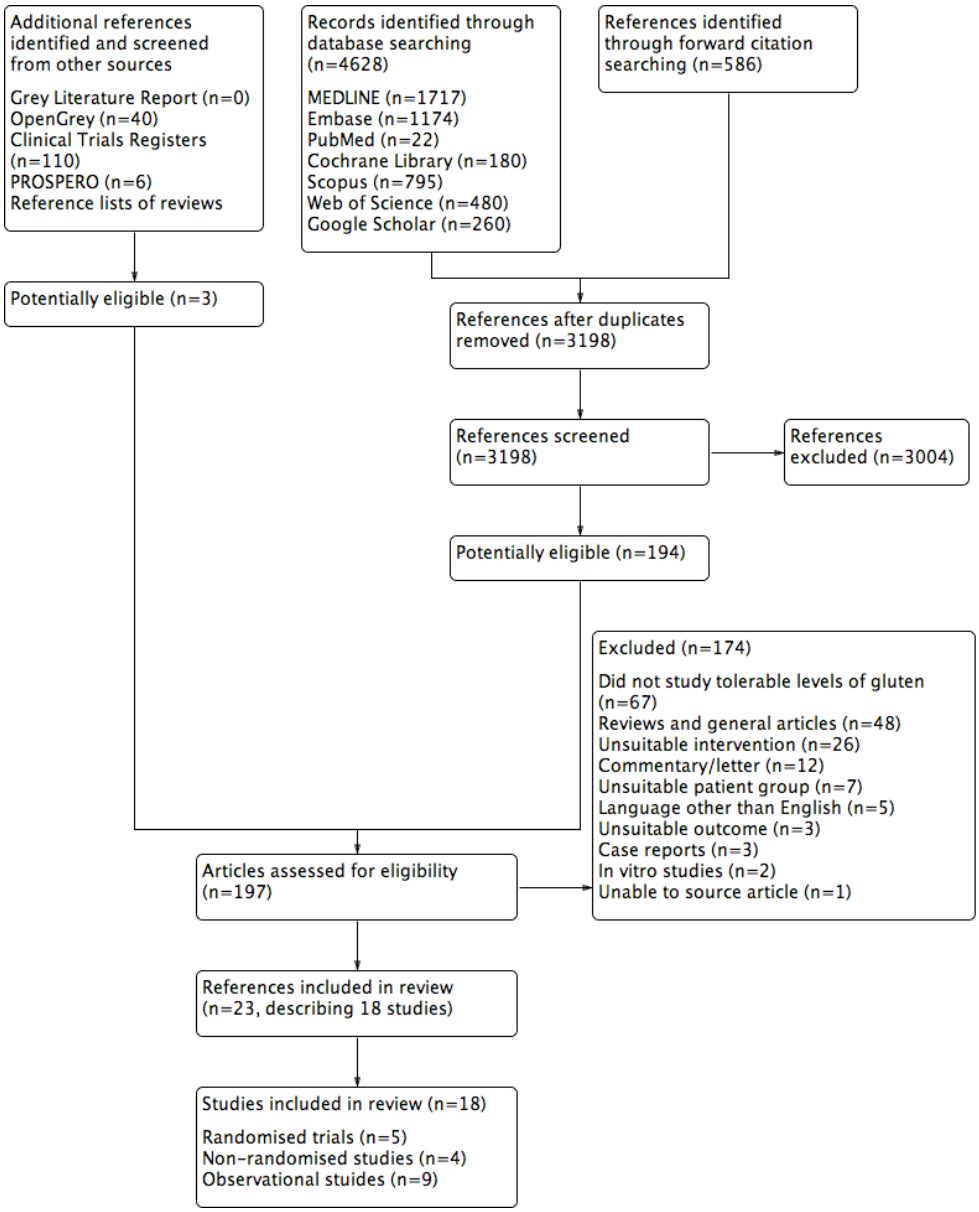
The searches of MEDLINE, Embase, Cochrane Library, PubMed, Scopus, Web of Science and Google Scholar were conducted on 11 November 2015 and yielded 4628 citations. (See **Appendix 1** for list of

search strategies and citations retrieved.) An additional 586 citations were identified as a result of the forward citation searching of the Akobeng review and its included studies. Following de-duplication, 3198 citations were initially screened by one reviewer (JR) and 3004 excluded as being clearly out of scope for this review.

The remaining 194 citations were carefully considered by the same reviewer, obtaining the full-text of articles where necessary. All eligible and potentially eligible citations, plus those citations deemed irrelevant because of the dose of gluten given, were checked for eligibility by a second reviewer (KA).

Of the 194 citations screened, 23 were included (reporting 18 individual studies). Our searches retrieved the 13 studies included in Akobeng 2008 and all were deemed eligible for this review. We only identified one study (Greco 2011) published since Akobeng 2008. The majority of excluded citations concerned gluten challenge studies in which the amounts of gluten given far exceeded what is considered tolerable or safe. **Figure 1** shows the flow of citations through the review.

Figure 1 Study flow diagram



4.1.2 Grey literature and other sources

We screened 40 records from OpenGrey; these were all theses or dissertations (20 in French and five in Czech) and none met the inclusion criteria. No records were retrieved from Grey Literature Report.

Searches of the WHO International Clinical Trials Registry Platform and ClinicalTrials.gov yielded 110 trial records (23 unique to WHO; 39 unique to CT.gov; 48 in both registers). There was a record in CT.gov of the randomised trial by Catassi 2007. The only other relevant record was of an Indian trial by Nayer. According to the description in WHO Registry, this trial investigated the effect of trace amounts of dietary gluten on nutritional, clinical, serological, biochemical and histological parameters of children with coeliac disease (Nayar 2009). Despite following up the publications listed and writing to the author, we were unable to get any additional information about the trial results.

None of the six records retrieved from PROSPERO, the prospective register of systematic reviews, met our inclusion criteria. We identified one additional citation (Ejderhamn 1988) by checking the reference lists of relevant reviews. This study of 11 patients was included in the review by Hischenhuber but was deemed ineligible for this review because there was no comparative dietary exposure.

4.1.3 Submissions

We received one submission (from John O'Brien of Rebellion Brewing); this was a study of the proteomic characterisation and relative quantification of hordein (gluten) in beer (Colgrave 2012). This study presents spectral evidence for prolamin proteins and is not conducted on individuals with coeliac disease, as such it was not considered further for this review.

4.2 Characteristics of included studies

Eighteen studies met the eligibility criteria. These included nine intervention studies and nine observational studies. Five of the intervention studies were randomised trials and four were non-randomised (two crossover and two before-and-after studies). The observational studies were cross-sectional with the exception of one cohort study. The studies were conducted in Australia, Canada, England, Finland, Italy and Sweden, and were published between 1974 and 2011. A total of 1754 participants were included in the studies, although one study accounted for over half the participants. See **Table 3** for the detailed characteristics of the included studies, including the results.

4.2.1 Aims

The randomised trial that most closely matched the objective of this review (Catassi 2007) was the only included study that specifically aimed to establish the safety threshold of exposure to small amounts of gluten. The remaining four randomised trials covered aspects of safety of ingesting small amounts of gluten or gluten-containing products but which were not designed to establish a safe threshold. These four trials aimed to: evaluate the safety of wheat flour hydrolysed during food processing (Greco 2011); the safety of wheat-starch based hydrolysate products (Kaukinen 2008); the response to a wheat-starch based gluten-free diet (Peräaho 2003); and the effects of chronic ingestion of small amounts of gliadin (Catassi 1993).

The two crossover studies evaluated the effects of small amounts of gliadin-containing products (Ciclitira 1984; Ciclitira 1985). One of the before-and-after studies monitored the influence of gluten exposure (Laurin 2002), the other evaluated the tolerance of prolonged consumption of small amounts of gliadin in wheat-starch based products (Chartrand 1997). None of the observational studies were designed to establish a safe threshold of gluten; instead they were concerned with evaluating the response of small levels (estimated) of gluten or gliadin, or evaluating the safety of wheat-starch based products.

4.2.2 Participants

Nine studies included adults only, two included children only, one included adolescents only and six included both adults and adolescents/children. In most studies the diagnosis of coeliac disease was

confirmed by small bowel biopsy. Participants had been on a gluten-free diet for at least a year, with the exception of two studies that included newly diagnosed patients (Kaukinen 1999; Peräaho 2003). In all but one study (Selby 1999) coeliac disease was well-controlled (i.e. patients were asymptomatic). The studies were generally small, with the median number of participants being 40 (range 7 to 965). The subset of randomised trials included 219 participants (median 39 participants; range 13 to 90). The largest study (Stuart 1997) was a cross-sectional study conducted in Australia that only investigated the presence of symptoms and was based on questionnaire responses.

4.2.3 Interventions and exposure

The studies were very heterogeneous with respect to the type, amount and duration of exposure to gluten. In addition to studies that evaluated gluten, a range of gluten-containing products were also assessed, including wheat-starch based products (five studies), gliadin (four studies), hydrolysed wheat flour (one study) and wheat-based starch hydrolsate products (one study).

The amount of gluten (or equivalent) tested in the studies ranged from trace amounts (as little as 0.005 mg) to 16 g a day. **Table 1** shows the wide variation in both the types, duration and amounts of gluten evaluated just among the randomised trials. The duration of exposure to gluten (or equivalent) ranged from one week to one year among the intervention studies, and from four months to 11 years among the observational studies. Two of the randomised trials included a no-gluten placebo group (Catassi 2007; Kaukinen 2008).

Table 1 Types, duration and amounts (/day) of gluten evaluated in the randomised trials

Study ID	Intervention	Duration	Group 1	Group 2	Group 3
Greco 2011	Hydrolysed wheat flour	60 days	16,025 mg	496 mg	1.6 mg
Kaukinen 2008	Wheat-based starch hydrolsate products	24 weeks	0.12 mg	0.005 mg	'no gluten'
Catassi 2007	Gluten	90 days	50 mg	10 mg	'no gluten'
Peräaho 2003	Wheat-starch based gluten	1 year	~40 mg	~40 mg	[n/a]
Catassi 1993	Gliadin	28 days	1000 mg	200 mg	[n/a]

Information concerning the exact amount of gluten (or equivalent) given to participants was provided in two of the randomised trials (Catassi 2007; Catassi 1993) and the four non-randomised studies. For the remaining three randomised trials we calculated the amount of gluten based on other information provided in the paper (Greco 2011; Kaukinen 2008; Peräaho 2003). In the observational studies comparison groups were formed on the basis of the amount of gluten in the diet derived from estimates obtained through dietary interviews, food diaries and/or questionnaires.

4.2.4 Outcomes

All studies measured aspects of small bowel histology, serology and/or clinical symptoms. Small bowel biopsies were conducted in all but two studies (Chartrand 1997; Stuart 1997), 13 studies assessed coeliac serology and 13 measured gastrointestinal symptoms. Small bowel histology was most commonly reported as villous height/crypt depth (10 studies) and intraepithelial lymphocytes (IEL) count (12 studies). Measures of serology most frequently reported were anti-gliadin antibodies (AGA) (9 studies); endomysial antibodies (EMA) (7 studies); and anti-tissue transglutaminase (IgA) (4 studies). Clinical symptoms were reported in a number of different ways, including the Gastrointestinal Symptom Rating Scale (3 studies) plus a variety of other unspecified questionnaires.

4.3 Excluded studies

Of the 197 potentially eligible articles assessed for inclusion, 174 were excluded due to reasons described in the study flow diagram (**Figure 1**). All excluded papers were assigned to a single category even if some papers had multiple reasons for exclusion, e.g. a study that was conducted in an unsuitable population

and also tested an intolerable level of gluten. Articles were most commonly excluded because they described gluten-challenge studies that evaluated levels of gluten known to exceed safe or tolerable levels, such as several grams of gluten per day. Other excluded studies tested an unsuitable intervention (e.g. oats) or did not administer the gluten orally. Thirteen studies, which at first screening of title and/or abstract seemed particularly relevant to the review but were excluded, are listed in the excluded studies table below (**Table 2**). (See section 7.2 for the references to these studies.)

Table 2 Excluded studies

Study ID	Reason for exclusion
Biagi 2004	Case report
Buchanan 2008	Review
Ejderhamn 1988	Unsuitable intervention
Gibert 2006	Unsuitable outcome
Gibert 2013	Modelling study
Hamilton 1972	Did not study a tolerable level of gluten (2.25 g per day)
Jansson 2001	Did not study a tolerable level of gluten (estimated minimum 2.2 g per day)
Kumar 1988	Did not study a tolerable level of gluten (2.5 vs 10 g per day)
Lahdeaho 2011	Did not study a tolerable level of gluten (1.25-3 vs 3-5 g per day)
Mazzarella 2013	Unsuitable intervention (transamidation of gluten)
Nayar 2009	Unable to access article or contact investigators
Pyle 2005	Did not study a tolerable level of gluten (5 vs 10 g per day)
Srinvasan 1996	Unsuitable intervention (oat challenge)

Table 3 Characteristics of Included Studies (for abbreviations see end of table)

Study ID	Design	Aim	Participants	Duration	Exposure	Outcomes	Results
Randomised trials							
Greco 2011	RCT, Italy	To investigate safety of goods made of wheat flour hydrolysed during food processing.	**13 adolescents** (median age, 19 years; range, 12–23 years) on GFD for ≥ 5 years	60 days	- natural flour (16,025 mg gluten/day) (n=6) - extensively hydrolysed flour (496 mg gluten/day) (n=2) - fully hydrolysed flour (1.6 mg gluten/day) (n=5)[1]	1. small-bowel histology (duodenum) 2. serology 3. clinical symptoms	1. mucosal atrophy developed in 100%, 100% and 0%, respectively; IELs (CD3 and TCRγδ) increased in 100%, 100% and 0%, respectively 2. IgA anti-Tg2 antibodies increased in 100%, [NS], and 0%, respectively; IgA anti-EmA antibodies increased in 100% in 16 g group, not reported for 496 mg or 1.6 mg groups 3. clinical complaints in 33%, 0% and 0%, respectively
Kaukinen 2008	RCT, Finland	To investigate safety of wheat-based starch hydrolysate products.	**90 adults** (57 F, 33 M; median age, 55-61 years; range, 24-81 years) on GFD for ≥ 1 year (median duration 8-11 years, range 1-42 years)	24 weeks	- glucose syrups (0.12 mg gluten/day) (n=30) - maltodextrins (0.005 mg gluten/day) (n=30)[2] - placebo (no gluten) (n=30)	1. small-bowel histology (duodenum) 2. serology 3. clinical symptoms 4. adverse events	1. no significant difference between groups in Vh/Cd or IELs (CD3+, +, γ +) 2. in the maltodextrins group, one patient, who was EmA+ve and tTG-ab –ve at baseline, was tTG-ab +ve at the end of the study. One additional patient became tTG-ab +ve during the trial. 3. no significant difference between groups in GSRS; 3%, 10% and 10%, respectively withdrew due to abdominal symptoms 4. 21 adverse events possibly related to study products (data not provided)
Catassi 2007	RCT, Italy	To establish the safety threshold of prolonged exposure to trace amounts of gluten.	**39 adults** (30 F, 9 M; median age, 30.6 years; range, 20-55 years) on GFD for ≥ 2 years (median duration 10 years, range 2-28 years)	90 days	- 50 mg gluten/day (n=13) - 10 mg gluten/day (n=13) - 50 mg placebo/day (n=13)	1. small-bowel histology (duodenum) 2. serology 3. clinical symptoms	1. change in Vh/Cd was -20% (95% CI: -22%, -13%) in the 50 mg group, -1% (-18%, 68%) in the 10 mg group and 9% (3%, 15%) in placebo group; the difference between the placebo group and the 50mg group was significant (p<0.05). No significant difference in IEL count. 2. no change in IgA anti tTG. IgG AGA was significantly decreased in the 50 mg group compared to placebo (p=0.04). All

[1] Gluten calculated by multiplying amount of product (200 g) by ppm provided and dividing by 1000
[2] Gluten calculated by multiplying amount of product (27.7 g and 2.7 g) by ppm as measured by high-performance liquid chromatography (4.3 and 1.7, respectively) and dividing by 1000

Study ID	Design	Aim	Participants	Duration	Exposure	Outcomes	Results
							within normal limits. 3. 1 patient relapsed in 10 mg group; no patients in the 50 mg group had clinical evidence of relapse
Peräaho 2003	RCT, Finland	To compare the response to a wheat-starch based or natural GFD.	**57 adults** (45 F, 12 M; median age, 44-47 years; range, 22-69 years) newly detected CD	1 year	- wheat-starch-based GFD (~41 mg gluten/day) (n=28) - natural GFD (authors assumed no gluten in this diet, however could have been as high as 39.5 mg gluten/ day)[3] (n=29) [dietary advice provided; assessed by means of an interview and 4-day food diary]	1. small-bowel histology (duodenum) 2. serology 3. clinical symptoms 4. quality of life	No significant differences between groups in: 1. mucosal morphology (Vh/Cd, enterocyte cell height), density of IELs (CD3+, + & Y +) or HLA-DR expression; complete recovery of mucosa was not seen in all patients 2. serum antibodies (IgA EMA, IgA AGA, IgA tTg-ab) 3. GSRS score 4. PGWB score
Catassi 1993	RCT, Italy	To investigate the effects of chronic ingestion of small amounts of gliadin.	**20 children** (12 F, 8 M; median age, 3.1 years; range, 1.6-9.6 years) on GFD for mean (SD) 14 (3) months	28 days	- 1 g gluten/day (n=10) - 200 mg gluten/day[4] (n=10)	1. small-bowel histology (peroral jejunal) 2. serology 3. clinical symptoms	1. significant decrease in Vh/Cd in both groups (p<0.01), significantly greater decrease in 1 g group (p,0.01); significant increase in mean IEL count in both groups (p<0.01) 2. IgA-AGA increased in 75% and 40%, respectively 3. clinical abnormalities in 30% and 0%, respectively
Non-randomised studies							
Ciclitira 1985	Crossover study, England	To investigate symptoms and enterotoxicity of a gliadin-contaminated GF product.	**10 adults** on GFD for ≥ 1 year	6 weeks	- GFD for six weeks then same GFD plus gliadin-containing GF bread (up to 4.8 mg gluten/day[4]) - [4 patients GFD first, gliadin second, wash-out period not stated]	1. small-bowel histology (jejunal) 2. clinical symptoms	1. no significant difference in Vh/Cd after the test periods 2. 6 patients (60%) had higher symptom scores on GF bread; 2 (20%) on GF diet; results not significantly different
Ciclitira 1984	Crossover study,	To evaluate gliadin-containing GF product.	**7 adults** on GFD for ≥ 1 year	1 week	- GFD for one week then same GFD plus	1. small-bowel histology	1. significant decrease in Vh/Cd (2.4, ±0.7 vs 2.0 ±0.5, p<0.05) after the test periods; no

[3] Gluten calculated by multiplying mean amount of product (79 g and 82 g) by CODEX ppm (<500) and dividing by 1000
[4] Calculated by doubling the amount of gliadin

Study ID	Design	Aim	Participants	Duration	Exposure	Outcomes	Results
	England				gliadin-containing GF bread (up to 4.8 mg gluten/day[4]), no washout period stated	(jejunal)	significant difference in IEL count or epithelial surface cells height
Laurin 2002	Before and after study, Sweden	To monitor gluten intake, clinical symptoms, and antibody kinetics to evaluate the influence of gluten exposure during the challenge.	**24 children** (16 F, 8 M; median age 3.8 years; range, 2.7-8.8 years) on GFD for ≥ 1 year	Median 13 (range, 5-51) weeks	- 1.7 g mean gluten/day (range, 0.2-4.3 g)	1. small-bowel histology (small intestine) 2. serology 3. clinical symptoms	1. 91% (21 of 23) patients had 'more severely effected mucosa after the challenge'; 96% (n=22) patients IEL count increased significantly after the challenge (26, ± 13 vs 78, ± 31, p<0.0001) 2. 96% (23 of 24) had elevated antibodies (at least one of: IgA-AGA, IgA-EmA, IgG-AGA, IgG-EmA) 3. symptoms occurred after median of 8 (range, 0-105) days; 79% (19 of 24) had clinical symptoms within 4 weeks (not significantly correlated to gluten intake)
Chartrand 1997	Before and after study, Canada	To evaluate tolerance of prolonged consumption of small amounts of gliadin contained in products containing wheat-starch.	**15 adults, 2 children** (12 F, 5 M; mean age 37 years; range, 8-54 years) on strict GFD for ≥ 1 year (mean 4 years) A control group was also involved in this study, but not considered relevant for this review.	Median 6 (range, 0.5-10) months	- 1.5 mg mean gluten/day[5] (range, 0.75-3.38 mg) in bread containing wheat-starch	1. serology 2. clinical symptoms	1. no significant change in IgA-AGA, IgG-AGA or EmA 2. 11 patients (64%) developed 'disturbing' symptoms and withdrew from the trial within 8 months; 4 patients (24%) had slight or occasional symptoms; 2 patients (12%) remained asymptomatic; 6 of 17 reported fatigue/irritability; 2 of 17 reported bone pain; 2 of 17 reported flare-up of dermatitis herpetiformis
Observational studies							
Dissanayake 1974	Cohort study, England	To study the response of the mucosa to a strict gluten-free diet or to a continued intake of gluten	**38 mostly adults** (27 F, mean age at diagnosis: 35 years, range, 15-71 years; 11 M, mean age 45 years, range, 12-70 years)	Mean 27.5 (range, 6-72) months	- strict GFD (n=18) - small amounts of gluten (est. < 500 mg gluten/day) (n=13) - large amounts of gluten (est. 500 mg gluten/day) (n=7) [based on dietary interviews covering a 4-week period]	1. small-bowel histology (jejunal) 2. serology 3. clinical symptoms	1. 90% on strict GFD had normal or near-normal mucosal histology; 80% on 'small amounts' showed persistent mucosal abnormalities; no improvement in 'large amounts' group 2. no significant difference in post-treatment serum immunoglobulins (IgA, IgG, IgM) between the groups 3. 95% were asymptomatic at follow-up; 1 patient in 'large amounts' group had

[5] Calculated by doubling the amount of gliadin

Study ID	Design	Aim	Participants	Duration	Exposure	Outcomes	Results
							severe symptoms and 1 in the 'small amounts' groups had abdominal symptoms
Collin 2004	Cross-sectional study, Finland	To estimate a reasonable limit for residual gluten in GF products.	**76 adults, 16 children**	Median 2 (range, 1-10) years	- naturally GFD (n=28) - wheat-starch-based GFD (n=64) (range 0.2-60 mg in both groups)[6] [estimated from 4-day food diary]	1. small-bowel histology (small intestine) 2. serology	1. no correlation between flours used and mucosal damage (VhCD, IEL (CD3+)) 2. 1 patient on natural GFD and 1 patient on wheat starch-based GFD were IgA EmA +ve after one year
Lohiniemi 2000	Cross-sectional study, Finland	To establish effect of wheat-starch based GF products on gastrointestinal symptoms.	**53 adults** (39 F, 16 M; median age 42 years; range, 30-76 years) Diagnosed 9-11 years previously	Duration unknown.	- 36 mg mean gluten/day (range, 0-180 mg) [estimated from 4-day food diary A control group was also involved, but not considered relevant for this review.	1. small-bowel histology (in 23 patients only) 2. clinical symptoms 3. quality of life	1. 2 of 23 patients had villous atrophy 2. no correlation between gluten intake and GI symptoms (GSRS) 3. the correlation between gluten intake and PGWB was not reported
Kaukinen 1999	Cross-sectional study, Finland	To investigate whether wheat-starch based gluten-free products are safe.	**25 adults**[7] (20 F, 5 M; median age 50 years; range, 43-67 years) on GFD for 9-12 years, **16 children** (11 F, 5 M; median age 12 years; range, 7-18 years) on GFD for 2-10 years	8 years on average	- strict wheat-starch-based GFD (34 mg gluten/day (range, 5-150 mg)) (n=40) - dietary lapse (wheat-starch-based GFD plus 1-2 g gluten/week or month) (n=6) - strict naturally GFD (n=6) [based on 4-day food diary]	1. small-bowel histology (duodenal for adults and jejunal for children) 2. serology	1. no significant difference in Vh/Cd, IELs ($CD3^+$, $^+$, γ^+) or HLA-DR expression between strict wheat-starch-based GFD and strict GFD; one of the dietary lapse group had subtotal and one had severe partial VA; the dietary lapse group had significantly enhanced HLA-DR expression than the wheat-starch based group (p=0.04); IEL count not reported for the dietary lapse group 2. In the strict wheat-starch-based GFD and the natural GFD, IgA EmA or ARA were not present, however 2 had +ve AGA titres. In the dietary lapse group, 2 had +ve ARA & one of these had +ve AGA
Selby 1999	Cross-sectional study, Australia	To determine whether persistent VA could be due to continued ingestion of the trace amounts of	**89 adults** (73 F, 16 M; mean age 47 years; range, 20-75 years) Diagnosed 8 years	Duration unknown.	- Codex GFD <500 ppm (n=39) - non-detectable gluten GFD < 30 ppm	1. small-bowel biopsy (duodenal) 2. serology	1. no significant differences in proportion of VA or in IELs between Codex and NDG GFDs 2. no association between serum

[6] Calculated from range of gluten-free product intake (grams) and range of ppm for gluten-free product

[7] 11 adults with dermatitis herpetiformis also included; results not reported separately

Study ID	Design	Aim	Participants	Duration	Exposure	Outcomes	Results
		gluten in 'GF' foods, as defined by the WHO/FAO Codex Alimentarius.	(range, 0.6-29 years) previously. About 50% symptomatic		(n=50) [based on dietary interviews and 1-week diary] A control group was also involved in this study, but not considered relevant for this review.	3. clinical symptoms	antibodies (IgA AGA, anti-EmA) and GFD (less than half patients tested) 3. there was an improvement in symptoms in 8 of 22 participants (36%) by removal of trace amount of gluten from the diet (changing from CODEX GF diet (<500 ppm) to NDG GF diet (< 30 ppm))[8]
Stuart 1997	Cross-sectional study, Australia	To examine the dietary patterns… and to relate any symptoms being experienced to the different levels of gluten restriction	**965 adults and children** (695 F, 270 M; mean age 50 years; range, 4-88 years; 6% of sample children) Diagnosed: mean age 39 years (range >1-83 years)	Duration unknown	- no-detectable gluten; (< 30 ppm) (n=176) - trace gluten; CODEX (< 500 ppm) (n=694) - overt gluten (n=57) - unclassified (n=38) [based on responses to diet and symptoms questionnaire]	1. clinical symptoms	1. only 2 of the 13 symptom categories were significantly different between the trace gluten and no detectable gluten diets; constipation was more severe (p=0.0051) in the trace gluten group, while diarrhoea occurred more frequently in those consuming no detectable gluten (p=0.0032)
Mayer 1991	Cross-sectional study, appears to be Italy	To relate compliance of gluten-free diet to general state of health, to serum antigliadin antibody concentrations, and to the morphoetric features of small intestinal mucosa	**123 adults and adolescents** (71 F, 52 M; mean age, 13.7; range 10.6-23 years; diagnosis when ≤ 3 years.	Duration unknown	- strict GF diet (n=80) - GF diet with small amounts of gluten (60 mg-2 g gluten/day) (n=14) - gluten containing diet (~15 g gluten/day) (n=29) [assessed by dietary questionnaire & interview]	1. small-bowel histology (jejunal) 2. serology 3. clinical symptoms	1. small amounts (60 mg-2 g gluten/day) compared with strict GF diet resulted in significantly increased crypt epithelial volume (p<0.05) and expanded crypt IEL count (p≤0.02)(results from only 33 patients) 2. concentrations of AGA (IgG, IgA) were raised in 19% on a strict GF diet, and 14% on a small amounts of gluten diet: these were not significantly different 3. there was no difference in clinical symptoms for patients on a small amounts of gluten diet (60 mg-2 g gluten/day) compared to a GF diet
Montgomery 1988	Cross-sectional study, England	To determine the effects of a LGD on jejunal morphology and serum antibody levels to gluten in a group of CD patients maintained on a controlled LGD	**25 adults** (12 F, 13 M; median age, 40 years; range, 17-74 years (GFD) and median age, 25 years; range, 18-70 years (LGD))	3-27 months	- a "90-100%" strict GFD for median 13 months (range, 6-27 months) (n=12) - LGD of 2.5-5 g gluten/day for median 6 months (range, 3-14	1. small-bowel histology (jejunal) 2. serology	1. no significant difference in Vh/Cd between GFD and LGD groups; significant increase in IEL count in LGD group compared to GFD group (p<0.05) 2. no significant difference in serum AG-IgA, -IgG and -IgM titres

[8] these data are from a subset of participants who were involved in an intervention study (reported in Faulkner-Hogg 1999)

Study ID	Design	Aim	Participants	Duration	Exposure	Outcomes	Results
					months) (n=13) [assessed by dietary interview and questionnaire]		
Baker 1975	Cross-sectional study, England	To assess the incidence and effects of continuing gluten ingestion in CD	**51 adults** (28 F, 23 M) on GFD for mean 5.25 years (range, 0.3-11 years)	Duration unknown	- large gluten-intake (≥ 2000 mg gluten/day) (n=9) - small gluten-intake (< 2000 mg gluten/day) (n=24) - no gluten-intake (n=18) [based on 4-week questionnaire]	1. small-bowel biopsy (jejunal) 2. serology	1. 73% for patients on no gluten diet and 67% of patients on small gluten intake had VA 2. gluten ingestion associated with the presence of gluten antibodies (significant difference $p<0.05$)

Abbreviations

AG anti-gliadin; **AGA** anti-gliadin antibody; **CD** coeliac disease; **CI** confidence interval; **EmA** endomysial antibody; **F** female; **GF** gluten-free; **GFD** gluten-free diet; **GI** gastrointestinal; **GSRS** Gastrointestinal Symptom Rating Scale; **HLA** human leucocyte antigen; **IEL** intraepithelial lymphocytes; **IgA** Immunoglobulin A; **IgG** Immunoglobulin G; **LGD** low gluten diet; **M** male; **NS** not specified; **PGWB** Psychological General Well-Being Questionnaire; **RCT** randomised controlled trial; **SD** standard deviation; **Tg2** tissue transglutaminase; **VA** villous atrophy; **Vh/Cd** villous height crypt depth ratio

4.4 Risk of bias in included studies

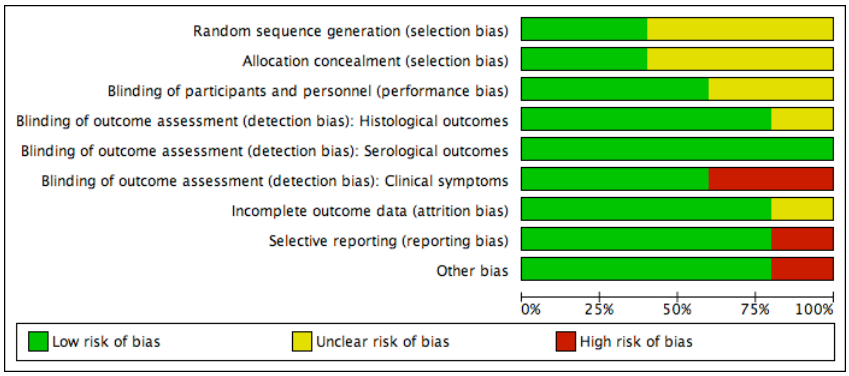
4.4.1 Randomised trials

Overall, the randomised trials were rated as being at low or unclear risk of bias. We were successful in obtaining additional information not available in the published reports from Catassi regarding the methods and design of the 1993 and 2007 trials.

Across the trials, information concerning randomisation (i.e. sequence generation and allocation concealment) was judged to be at low risk of bias in two trials and at unclear risk of bias in three trials (because the method was either missing or inadequately described). Three of the five studies (Catassi 1993; Catassi 2007; Kaukinen 2008) were judged to have blinded participants to the intervention. It was not always clear if outcome assessors were blinded. We considered this important for histologic readouts, but of less concern for the serology outcomes. For patient-reported outcomes and clinical symptoms a lack of blinding in two trials (Greco 2011; Peräaho 2003) meant these trials were judged at high risk of bias for this domain. Incomplete outcome data was generally well-reported (low risk of bias in four trials) with adequate descriptions of any missing data, including the numbers and reasons for withdrawal between the groups. In Peräaho 2003 our concern that the naturally gluten-free diet was contaminated with gluten and the difficulty in accurately estimating the levels of gluten tested led us classify this trial at high risk of 'other bias'.

The risk of bias of the included randomised trials is represented graphically in the two risk of bias summary graphs below (**Figure 2** and **Figure 3**). **Appendix 2** has the detailed risk of bias judgements for each study across the seven domains.

Figure 2 Review authors' judgements about each risk of bias item presented as percentages across all included studies



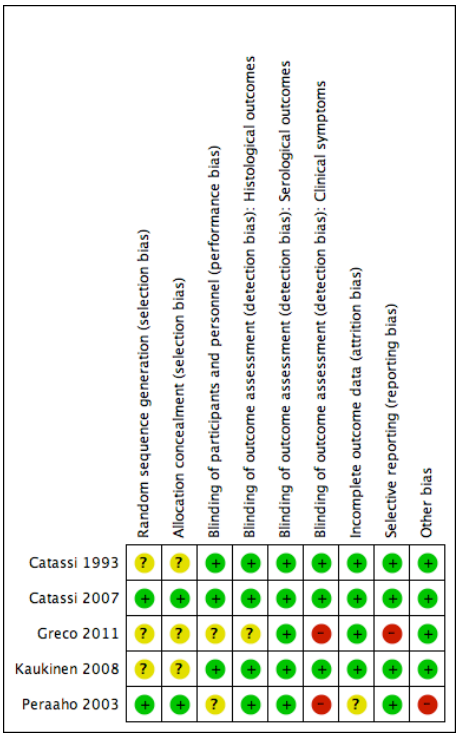


Figure 3: Review authors' judgements about each risk of bias item for each included study

4.4.2 Non-randomised studies and observational studies

The two small crossover studies by Ciclitira were non-randomised and rated at high risk of bias. It was unclear how the patients had been selected and neither study reported a washout period, making it difficult to rule out the possibility of a carry-over effect. This is of particular concern when switching from a gluten-containing diet to a gluten-free diet as the period of time required for bowel repair is unclear. Additionally, in the earlier study (Ciclitira 1984), all participants received the treatments in the same order (gluten-free diet followed by gluten-containing diet), which limits comparability. Despite these limitations, all relevant outcome data appeared to be reported, and in the case of the earlier one-week study of seven patients, the biopsy specimens were assessed blind.

The two before-and-after studies were also rated at high risk of bias. In Laurin 2002 children consumed different amounts of gluten for different periods of time. Parents of the children were not blinded and there was no indication that outcome assessors were blinded. Additionally, many participants did not have normal mucosal specimens prior to the challenge. Chartrand 1997 also challenged participants with different amounts of gluten for varying periods of time. This study was not blinded and had a very high dropout rate due to 'intolerable' clinical symptoms. No histological outcomes were measured.

The eight cross-sectional studies and one cohort study were considered to be at high risk of bias because of the significant limitations with attributing causation from such studies. Furthermore, the accuracy of exposure to gluten was dependent on estimates derived from questionnaires and dietary interviews. Interpreting the results was also problematic because individuals within the defined study groups had been exposed to varying amounts of gluten over different periods of time.

4.5 Effects of interventions

The heterogeneity with respect to the type, amount and duration of gluten exposure, as well as differences in how outcomes were measured, the populations studied and the types of study design included, precluded the possibility of conducting any quantitative synthesis of the data. As noted in section 4.2.1 above, the only study that set out to establish a safety threshold of gluten was the randomised trial by Catassi 2007. The results are thus presented narratively below and are summarised in **Table 3**. To help address the question of safety, **Table 4** presents the adverse changes associated with the lowest amount of gluten given in each study. The results have been grouped by the type of evidence (randomised versus non-randomised) and then by outcome (histology, serology and clinical symptoms).

4.5.1 Randomised evidence

Histology

The randomised trials assessed various amounts of gluten or gluten-containing products (ranging from 0.005 mg to 16 g gluten/day) over periods lasting from 60 days to 1 year. In Catassi 2007 patients receiving 50 mg gluten/day for 90 days had a significant decrease in the villous height/crypt depth (Vh/Cd) index (decline in 11 of 13 patients) whereas patients on placebo had a significant improvement (increase in 11 of 13 patients). Around half the patients receiving 10 mg gluten/day (6 of 13) showed an improvement in the Vh/Cd index. In the other trials, mucosal atrophy was present in the two patients who received 496 mg gluten/day in Greco 2011 and in 9 of 10 children who received 200 mg gluten/day in Catassi 1993. Mucosal atrophy was not present in patients exposed to smaller levels of gluten (1.6 mg in Greco 2011; 0.12 mg or 0.005 mg in Kaukinen 2008). In Peräaho 2003 both groups received similar amounts of gluten (estimated at ~40 mg gluten/day) and in neither group was there complete mucosal recovery among all participants.

Intraepithelial lymphocytes (IELs) increased in the two participants receiving 496 mg (Greco 2011) and all 10 participants receiving 200 mg (Catassi 1993) but not in those receiving 1.6 mg gluten/day (Greco 2011). In the two trials that included a placebo group the IEL count was not significantly different between the placebo and gluten groups (0.005 mg and 0.12 mg (Kaukinen 2008) and 10 mg and 50 mg (Catassi 2007)). Neither was there a difference in the IEL count between the two groups in Peräaho 2003.

Serology
Serum antibodies increased in 6 of 8 (75%) participants receiving 1 g (Catassi 1993), in 4 of 10 (40%) receiving 200 mg (Catassi 1993) and in 2 of 30 (7%) receiving 0.005 mg gluten/day (Kaukinen 2008), but remained within normal limits for participants consuming placebo, 10 mg or 50 mg (Catassi 2007) and 1.6 mg (Greco 2011). No difference was found in serum antibodies in the two groups (estimated at ~40 mg gluten/day) in Peräaho 2003.

Clinical symptoms
Clinical symptoms (as variously measured) were present in 1 of 13 (8%) participants receiving 10 mg (Catassi 2007), 3 of 10 (30%) receiving 1 g (Catassi 1993) and 2 of 6 (33%) receiving 16 g gluten/day (Greco 2011), but were not present in participants receiving 496 mg (Greco 2011), 200 mg (Catassi 1993), 50 mg (Catassi 2007) or 1.6 mg (Greco 2011). There was no significant difference in symptoms in groups receiving placebo, 0.005 mg or 0.12 mg gluten/day (Kaukinen 2008) or between groups receiving an estimated 40 mg gluten/day (Peräaho 2003).

4.5.2 Non-randomised evidence

Histology
The amount of gluten (actual or estimated) assessed in the non-randomised and observational studies ranged from 'gluten-free' to ~15 g of gluten/day over periods from four months to 11 years. There was considerable variation between the studies in the results for this outcome. Consistent with the randomised evidence, mucosal atrophy occurred in patients receiving 60 mg to 2 g (Mayer 1991), 200 mg to 4.3 g (Laurin 2002), < 500 mg (Dissanayake 1974) and 2 g gluten/day (Baker 1975). In the two crossover studies that evaluated the effects of 4.8 mg gluten/day, mucosal atrophy was present among children in the one-week study (Ciclitira 1984) but not among those in the six-week study (Ciclitira 1985). Montgomery 1988 reported no significant difference in mucosal atrophy between patients receiving a strict gluten-free diet compared to those estimated to ingest 2.5 to 5 g gluten/day, although the IEL count was higher in the 2.5 to 5 g group. IEL counts were higher in participants receiving 60 mg to 2 g (Mayer 1991) and 200 mg to 4.3 g gluten/day (Laurin 2002) compared to a gluten-free diet.

The studies of wheat-starch-based products found no difference between a gluten-free diet based on these products and a natural gluten-free diet, either in terms of mucosal atrophy or IEL count (Collin 2004; Kaukinen 1999; Lohiniemi 2000). Amounts of gluten were not reported for all the diets. Selby 1999 found no difference in mucosal atrophy and IEL count between consuming a 'CODEX' gluten-free diet (< 500 ppm) and a no-detectable gluten diet (< 30 ppm); mucosal atrophy was present in around half the participants in the two groups.

Serology
In three studies that compared a strict gluten-free diet to diets of differing levels of gluten (lowest intake estimated at 60 mg gluten/day) there was no significant difference between the participants in serological outcomes (Dissanayake 1974; Mayer 1991; Montgomery 1988). In the two before-and-after studies, serum antibodies were elevated in participants receiving 200 mg to 4.3 g gluten/day (Laurin 2002) but not in those receiving 0.75 to 3.38 mg gluten/day (Chartrand 1997). Baker 1975 reported a significant association between gluten ingestion and the presence of gluten antibodies. Serological outcomes were not measured in the two crossover studies by Ciclitira.

In Collin 2004 and Kaukinen 1999 there was no difference between a wheat-starch based gluten-free diet and a natural gluten-free diet in terms of serum antibodies. Similarly for this outcome, Selby 1999 found no association between consuming a 'CODEX' gluten-free diet (< 500 ppm) and a no-detectable gluten diet (< 30 ppm).

Table 4 Adverse changes associated with lowest amounts of gluten evaluated in each study

Study ID	Amount of gluten	Changes
Randomised trials		
Greco 2011	**496 mg/day** for 60 days	2 of 2 (100%) developed VA
Kaukinen 2008	**0.005 mg/day** for 24 weeks	2 of 90 (2%) tTG-ab –ve at baseline became tTG-ab +ve during the trial.
Catassi 2007	**10 mg/day** for 90 days	Vh/Cd decreased in 7 of 13 (54%) but not significantly different to placebo [p value not reported]. 1 of 13 (8%) showed signs of relapse.
	50 mg/day for 90 days	Vh/Cd was significantly worse than placebo (p<0.05)
Peräaho 2003	**~40 mg/day** (est.) for 1 year	Improvement in mucosa but complete recovery not seen in all patients
Catassi 1993†	**200 mg/day**‡ for 36 days	Significant decrease in Vh/Cd in 9 of 10 (90%) (p<0.01), significant increase in IEL count in 10 (p<0.01), IgA-AGA increased in 4 patients
Non-randomised studies		
Ciclitra 1985	**4.8 mg/day**‡ for 42 days	Mean symptom scores during each week of GF bread ingestion were greater than corresponding control weeks (not significant)
Ciclitra 1984	**4.8 mg/day**‡ for 7 days	Significant reduction in mean Vh/Cd (p<0.05) compared to usual gluten-free diet
Laurin 2002†	**200 mg to 4.3g /day**‡ for 5-51 weeks	21 of 23 (91%) had worsening mucosa, 22 of 23 (96%) had increased IEL count, 23 of 24 (96%) had elevated antibodies, 19 of 23 (79%) developed symptoms
Chartrand 1997	**1.5mg**/day‡ for ~6 months	15 of 17 (88%) developed symptoms and 11 withdrew from the study
Observational studies		
Dissanayake 1974	**< 500 mg/day** for ≥ 6 months	10 of 13 (77%) showed persistent mucosal abnormalities
Lohiniemi 2000	**~36 mg/day** for unspecified duration	2 of 23 (9%) had VA
Kaukinen 1999	**~34 mg/day** for > 6 months	No significant differences between 34 mg gluten/day (mean) and naturally GF group (mg gluten not specified) in all outcomes assessed
Selby 1999	**< 500 ppm**	There was an improvement in symptoms in 8 of 22 participants (36%) by removal of trace amount of gluten from the diet (changing from CODEX GF diet (< 500 ppm) to NDG GF diet (< 30 ppm))*
Stuart 1997	**< 30 ppm**	No significant differences in symptoms between NDG diet (< 30 ppm) and CODEX (< 500 ppm), except for more severe constipation in the CODEX group (p=0.0051) and more frequent diarrhoea in the NDG group (p=0.0032)
Mayer 1991†	**60 mg to 2 g/day** for > 10 years	Significantly increased crypt epithelial volume (p<0.05) and expanded crypt IEL count (p≤0.02 (results from only 33 patients) compared with GF diet
Montgomery 1988	**2.5 to 5 g/day** for ≥ 3 months	Significant increase in IEL compared with strict GF diet (p<0.05)
Baker 1975	**< 2 g/day** for > 4 months	16 of 24 (67%) had VA

‡ Calculated by doubling the amount of gliadin; † studies of children or adolescents; * in Faulkner-Hogg 1999

Abbreviations

AGA anti-gliadin antibodies; **GF** gluten-free; **IEL** intraepithelial lymphocyte; **IgA** immunoglobulin A; **NDG** no detectable gluten; **PVA** partial villous atrophy; **tTG-ab** tissue transglutaminase antibodies; **VA** villous atrophy; **Vh/Cd** villous height/crypt depth

Clinical symptoms

The frequency of symptoms varied between the studies and points to the lack of consistency across studies with respect to time periods and how symptoms were defined and measured. While clinical symptoms were present in 19 of 24 (79%) of children receiving 200 mg to 4.3 g (Laurin 2002) and in 2 of 20 (10%) participants receiving ~500 mg (Dissanayake 1974) gluten/day, there were also present in patients receiving much smaller amounts of gluten, e.g. in 15 of 17 (88%) participants receiving 1.5 mg (Chartrand 1997) and in 6 of 10 (60%) receiving 4.8 mg gluten/day (Ciclitra 1985). In Mayer 1991 there was no difference in clinical symptoms for patients on smaller amounts of gluten (range 6 to 2000 mg gluten/day) compared to a gluten-free diet. Lohiniemi 2000 found there was no correlation between gluten intake and symptoms.

Selby 1999 (reported in Faulkner-Hogg 1999) found clinically significant improvements in symptoms in a third of participants (8 of 22) who switched from a CODEX gluten-free diet (< 500 ppm) to a no-detectable

gluten diet (< 30 ppm). Stuart 1997 reported that only 2 of the 13 symptom categories (constipation and diarrhoea) were significantly different between the trace gluten (< 500 ppm) and no-detectable gluten (< 30 ppm) diets.

5 Discussion

5.1 Summary of main results

This systematic review aimed to establish whether there is a threshold level of gluten intake that is safe for people with coeliac disease to consume. The review's inclusion criteria were intentionally broad and consequently the 18 included studies were heterogeneous with respect to study design, aims, types of participants, gluten exposure and outcomes assessed. In the absence of any opportunity to conduct a quantitative synthesis, the results were summarised narratively. In discussing the results, primacy is given to the randomised evidence since the limitations inherent in observational studies to address the review question led to the evidence from these studies being rated as very low quality. The randomised evidence was also characterised by heterogeneity, particularly with respect to the amount, type and duration of gluten exposure.

Emphasis is given to the results of Catassi 2007 since this was the only included study that aimed to establish a safety threshold of gluten intake. The other studies provide information in relation to the effects of ingestion of small amounts of gluten. Determining what a current safe threshold might be relies on extrapolating data from studies not specifically designed to address this issue.

5.1.1 Tolerable levels of gluten

For the primary outcome of small bowel histology, important because persistent mucosal damage is associated with malignancy (Lebwohl 2013), there is considerable individual variation in the amount of gluten that causes mucosal damage. Only the randomised trial by Catassi 2007 directly addressed the objective of this review and therefore provided the most relevant information. This trial of 39 patients with coeliac disease found that consumption of 10 mg gluten/day for 90 days was tolerated by some patients but led to mucosal damage in others. It was only at the higher dose of 50 mg gluten/day that the difference between the placebo group with respect to Vh/Cd was significant. The study authors concluded that ingestion of gluten should be kept below 50 mg/day. This trial also highlighted the individual variation in sensitivity to gluten exposure as expressed by clinical symptoms. At the lower dose of 10 mg gluten/day, one participant withdrew from the study because of clinical symptoms, yet none of the patients receiving 50 mg had clinical evidence of relapse.

At the very low doses of gluten (in the form of wheat-starch based hydrolysate products) evaluated in Kaukinen 2008 (0.005 mg and 0.12 mg gluten/day), histological readouts were within normal range at the end of the 24-week trial. In Greco 2011 patients receiving 1.6 mg gluten/day for 60 days (in the form of hydrolysed wheat flour) did not suffer mucosal damage or report clinical complaints. Some of the earlier studies included in this review, such as the randomised trial by Catassi 1993 and the before-and-after study by Laurin 2002, assessed the effects of 200 mg gluten/day and found that such amounts caused mucosal damage in virtually all patients. In studies that estimated daily gluten intake at around 40 to 60 mg (Kaukinen 1999; Lohiniemi 2000; Peräaho 2003) histological changes were present in some patients.

From the limited randomised and other evidence available, mucosal deterioration was commonly present in patients receiving ≥ 50 mg gluten/day and was present in some patients receiving 10 mg gluten/day. However, the absence of randomised trials evaluating consumption of gluten in the range of 2-10 mg gluten/day limits our ability to establish a definitive amount of gluten that is safe to consume (whether for all patients or for an estimated proportion of patients). The data in relation to clinical symptoms is more difficult to interpret because of the uncertainty and variability over how this outcome is measured across studies, and the susceptibility of this outcome to recall bias and confounding, especially among the cross-

sectional studies. There was little consistency in the direction of effect across studies for clinical symptoms.

5.1.2 Thresholds

Coeliac disease patients, even those following a strict gluten-free diet, find it difficult to avoid at least some exposure to gluten. The two factors that affect the amount of gluten consumed are the gluten content of 'gluten-free' products (measured in ppm) and the amount of 'gluten-free' products consumed. The gluten-free threshold for product-labelling used by jurisdictions in the EU, US and Canada (the CODEX standard) has reduced over time, from 500 ppm in 1981 to 20 ppm in 2015.

As **Table 5** below shows (reproduced from Bruins Slot 2015), a patient with coeliac disease who consumes an average of 300-400 g of gluten-free products a day would be exposed to 6-8 mg gluten/day at a gluten-free threshold of 20 ppm. Under the current Australian threshold of no-detectable gluten, the gluten-free detectable limit is 3 ppm, equivalent to exposure to 0.9-1.2 mg gluten/day (based on the same average daily consumption of 300-400 g of gluten-free products). At 500 g of gluten-free products a day, the gluten exposure increases to 10 mg/day and 1.5 mg/day, respectively.

A further consideration is the time period covered by the included studies. Amounts of gluten able to be detected and included in products labelled as gluten-free has changed over time. It is likely that patients in low- or no-gluten groups in studies conducted many years ago would have been consuming higher background levels of gluten that at the time were undetectable.

Table 5 Estimated amount of daily gluten exposure (mg) [reproduced from Bruins Slot 2015]

Gluten content of gluten-free products (ppm)	**Amount of gluten-free products consumed (g)**					
	100g	**200g**	**300g**	**400g**	**500g**	**600g**
200	20	40	60	80	100	120
100	10	20	30	40	50	60
50	5	10	15	20	25	30
40	4	8	12	16	20	24
30	3	6	9	12	15	18
20	2	4	6	8	10	12
10	1	2	3	4	5	6
5	0.5	1	1.5	2	2.5	3
3*	0.3	0.6	0.9	1.2	1.5	1.8

* considered the lowest limit of detection for gluten currently

5.2 Quality and completeness of the evidence

In considering the overall quality of the evidence we have used the GRADE domains (imprecision, inconsistency, directness and risk of bias) as a framework for those studies that most influence the results of the review (Balshem 2011). In relation to the directness of the evidence, only one study directly addressed the review question (Catassi 2007). This was a well-designed randomised trial, judged to be at low risk of bias, conducted in a population whose results can be considered applicable to other populations of coeliac patients. However, the trial's small size resulted in imprecision around both the direction and magnitude of the effect at the important lower amount of gluten assessed (10 mg gluten/day). Consequently, we rated this study as moderate quality evidence (downgraded for imprecision).

For the remaining randomised trials the major limitation concerns the directness of the evidence. Since none of the trials aimed to establish a safe threshold of gluten intake they effectively provide indirect or supplementary evidence. However, even their contribution as indirect evidence is limited by the type and amounts of gluten assessed, and the different populations in which gluten was evaluated. Other than confirming the correlation between gluten intake and histological change, they provide limited

information in relation to establishing a safe threshold. Because of the serious reservations over directness we rated these studies as low quality evidence.

The non-randomised and observational studies provide very low quality evidence with respect to establishing of a safe threshold. The findings of these studies need to be interpreted very cautiously, and at best, only provide additional information about the association between intake and histological changes. Many of the studies included were cross-sectional, a design that is useful for finding associations but not causality. There are significant limitations in inferring relevant evidence from these studies to the review question. The reliance on food diaries, questionnaires and interviews to estimate gluten intake and establish the comparative groups makes these studies highly susceptible to recall bias and other confounding factors. Furthermore, the wide range of gluten amounts consumed by individuals within the groups is problematic in identifying a single gluten amount that can be associated with the changes observed in each study.

In terms of completeness of evidence, there is a need for a larger double blind placebo-controlled randomised trial that compares smaller amounts of gluten intake (< 10 mg gluten/day).

5.3 Potential biases in the review process

An important source of bias in any systematic review is the potential to miss relevant studies. We minimised this potential bias by comprehensively searching the literature and included several supplementary steps, including searches of trial registers and conducting forward citation searches of existing systematic reviews and the studies included in these reviews. We are confident that our approach is unlikely to have resulted in missing randomised trials directly addressing the review question. We excluded studies reported in languages other than English, and although we didn't identify any such relevant studies, it is possible that such studies may contribute additional evidence.

5.4 Agreement with existing reviews

We identified three reviews that considered safe gluten thresholds for patients with coeliac disease (Akobeng 2008, Bruins Slot 2015 and Hischenhuber 2006). We deemed the reviews by Akobeng and Bruins Slot to be systematic reviews and carried out a risk of bias assessment using the new ROBIS tool (Whiting 2015). The review by Hischenhuber, although not identifiable as a systematic review, is included here because it aims to assess safe limits of gluten in food and includes most of the studies included in the Akobeng review plus a few others (see **Appendix 3**). **Table 6** briefly summarises the main characteristics and conclusions of the reviews.

5.4.1 ROBIS assessment of Bruins Slot and Akobeng

Both reviews shared the same major concern, namely the lack of appropriate integration of information about study quality with the reporting of the results. This is particularly important when different types of study design are included in reviews. In Akobeng there was a description of the risk of bias of the intervention studies but no such appraisals were included or mentioned in Bruins Slot. Because we have conducted our own review of primary studies rather than an overview of existing reviews, the limitations identified have little bearing on the results of our review. **Appendix 4** provides the full ROBIS assessments.

5.4.2 Main findings

The recent systematic review by Bruins-Slot 2015 assessed whether the currently applied gluten thresholds are suitable to protect people with coeliac disease. The review included 11 studies conducted since the mid-1990s and found that while the 20 ppm 'gluten-free' threshold is safe for most of the coeliac population, it does not protect sensitive and recovering patients. The authors propose that a 3 ppm gluten-free threshold would be safe for all coeliac disease patients and, furthermore, would genuinely indicate if a product is gluten-free, since 3 ppm is currently the lowest detectable level of gluten.

The systematic review by Akobeng 2008 investigated the threshold amount of gluten and the threshold concentration of gluten in food products that can be tolerated by people with coeliac disease. The review included 13 studies. The review found that the daily amount of tolerable gluten varied among people with coeliac disease, making it impossible to set a single definitive threshold. The Codex standard prevailing at the time of the review (200 ppm) was deemed not sufficiently protective for all people with coeliac disease. The authors argued that if the concentration was set at 20 ppm (equivalent to consumption of around 6 mg gluten/day) this would be less likely to induce mucosal changes.

The earliest of the three reviews by Hischenhuber was published in 2006 (before the Catassi 2007 randomised trial). The review evaluated whether an upper limit for gluten content in food could be identified which would be safe for people with coeliac disease and for people with wheat allergy. The review included 15 studies and concluded that for people with coeliac disease the safe daily limit is between 10 mg and 100 mg of gluten. The authors state that current data indicate that wheat-starch-based food is safe provided it contains < 100 ppm, and they conclude that there is no evidence to support a definition of naturally gluten-free requiring no detectable gluten.

Table 6 Summary of existing reviews

Review ID	Purpose	Included studies and participants	Quality	Review authors' conclusions
Bruins Slot 2015	"This paper discusses the different thresholds that are currently used to label products gluten-free and compares tolerable gluten levels to gluten levels CD patients can be exposed to with these thresholds in place."	**11 studies** (3 RCTs; 6 cross-sectional; 1 cohort; 1 case report) **349 participants** (229 adults; 120 children/adolescents)	High risk of bias	"Currently, the most applied gluten threshold to label products gluten-free does not protect the most vulnerable patients. Therefore, we propose to lower the threshold for products with a gluten-free label to 3 ppm gluten."
Akobeng 2008	"To investigate the threshold amount of gluten and threshold concentration of gluten in food products that can be tolerated by people with coeliac disease."	**13 studies** (3 RCTs; 2 crossover; 7 cross-sectional; 1 cohort) **568 participants**	High risk of bias	"The amount of tolerable gluten varies among people with CD. Although there is no evidence to suggest a single distinctive threshold, a daily gluten intake of <10 mg is unlikely to cause significant histological abnormalities." "…the current Codex standard of 200 ppm is not sufficiently protective for *all* people with CD…"
Hischenhuber 2006	"Literature was reviewed to evaluate whether an upper limit for gluten content in food,… could be identified."	**15 studies** (incl. 2 RCTs) **292 participants**	Not assessed, as results superseded by more recent reviews.	"For coeliac disease sufferers [the gluten] limit should lie between 10 and 100 mg daily intake." "Based on a consideration of diet of coeliac patients, current data indicate that wheat-starch based food is safe, provided it contains <100 mg gluten/kg."

5.4.3 Agreements or disagreements with our review

The conclusions of the earlier review by Hischenhuber and the stipulation of specific safe limits of gluten have been superseded by more recent evidence. The findings of the reviews by Akobeng 2008 and Bruins Slot 2015 are broadly concordant with our review. These two reviews share a similar message although their perspective is different. By advocating for a reduction in the threshold to 3 ppm for gluten-free products, Bruins Slot is seeking to protect the most vulnerable coeliac patients. Akobeng, on the other hand, recognises that consumption of up to 10 mg gluten/day can cause histological abnormalities in some patients, but that these are unlikely to be significant.

6 Conclusions

Despite a comprehensive search of the literature, few studies have been published that directly address the issue of a safe level of gluten intake for people with coeliac disease. There is moderate evidence based on one randomised trial that patients with coeliac disease develop mucosal damage following consumption of 50 mg gluten/day, but the evidence of effect is uncertain for 10 mg gluten/day (equivalent to consuming 500 g a day of 'gluten-free' products with a gluten content of 20 ppm). The overall evidence included in the review is characterised by a high level of heterogeneity with respect to the type, amount and duration of exposure to gluten, the populations studied and range of study designs. Most studies included in the review, because of the limitations above, provide low quality indirect evidence of an association between gluten consumption and mucosal change. The studies reflect the variability in the amount of gluten that can be tolerated by individuals with coeliac disease and show a progression over time towards assessing the safety and tolerability of ever smaller amounts of gluten. In the absence of a large adequately powered randomised trial that compares low amounts of gluten intake (0 to 10 mg), the current evidence precludes establishing a definitive threshold level of gluten that is safe for all people with coeliac disease to consume.

7. References

Balshem 2011
Balshem H, Helfand M, Schunemann HJ, Oxman AD, Kunz R, Brozek J, Vist GE, Falck-Ytter Y, Meerpohl J, Norris S, Guyatt GH. GRADE guidelines 3: rating the quality of evidence - introduction. Journal of Clinical Epidemiology 2011;64(4):401-6. DOI: http://dx.doi.org/10.1016/j.jclinepi.2010.07.015

Colgrave 2012
Colgrave ML, Goswami H, Howitt CA, Tanner GJ. What is in a beer? Proteomic characterization and relative quantification of hordein (gluten) in beer. *Journal of Proteome Research*. 2012 Jan 1;11(1):386-96. doi: 10.1021/pr2008434.

Hall 2009
Hall N, Rubin G, Charnock A. Systematic review: adherence to a gluten-free diet in adult patients with coeliac disease. *Alimentary Pharmacology & Therapeutics* 2009;30(4):315-30.

Lebwohl 2015
Lebwohl, B., Ludvigsson, J. F., & Green, P. H. (2015). Celiac disease and non-celiac gluten sensitivity. *BMJ* 2015;351, h4347. doi:10.1136/bmj.h4347.

Whiting 2015
Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, Davies P, Kleijnen J, Churchill R; ROBIS group. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology*. 2016 Jan;69:225-34. doi:10.1016/j.jclinepi.2015.06.005. Epub 2015 Jun 16.

7.1 References to included studies

* = primary publication used for the review

Baker 1975
Baker PG, Barry RE, Read AE. Detection of continuing gluten ingestion in treated coeliac patients. *BMJ* 1975;1(5956):486-8.

Catassi 1993
Catassi C, Rossini M, Ratsch IM, et al. Dose dependent effects of protracted ingestion of small amounts of gliadin in coeliac disease children: a clinical and jejunal morphometric study. *Gut* 1993;34(11):1515-9.

Catassi 2007
* Catassi C, Fabiani E, Iacono G, et al. A prospective, double-blind, placebo-controlled trial to establish a safe gluten threshold for patients with celiac disease. *American Journal of Clinical Nutrition* 2007;85(1):160-6.

Catassi C, Fabiani E, Mandolesi A, et al. Toxicity of gluten traces: the Italian study on gluten microchallenge. American Gastroenterological Association Digestive Disease Week 2005.

NCT00250146. Catassi C. Multicenter study on the toxicity of gluten traces in the treatment of celiac disease. https://clinicaltrials.gov/ct2/show/NCT00250146 (accessed 15 December 2015).

Chartrand 1997
Chartrand LJ, Russo PA, Duhaime AG, et al. Wheat starch intolerance in patients with celiac disease. *Journal of the American Dietetic Association* 1997;97(6):612-8.

Ciclitira 1984
Ciclitira PJ, Ellis HJ, Fagg NL. Evaluation of a gluten free product containing wheat gliadin in patients with coeliac disease. *BMJ* 1984;289(6437):83.

Ciclitira 1985
Ciclitira PJ, Cerio R, Ellis HJ, et al. Evaluation of a gliadin-containing gluten-free product in coeliac patients. *Human Nutrition. Clinical Nutrition* 1985;39(4):303-8.

Collin 2004
Collin P, Thorell L, Kaukinen K, et al. The safe threshold for gluten contamination in gluten-free products. Can trace amounts be accepted in the treatment of coeliac disease? *Alimentary Pharmacology and Therapeutics* 2004;19(12):1277-83.

Dissanayake 1974
Dissanayake AS, Truelove SC, Whitehead R. Jejunal mucosal recovery in coeliac disease in relation to the degree of adherence to a gluten-free diet. *Quarterly Journal of Medicine* 1974;43(170):161-85.

Greco 2011
Greco L, Gobbetti M, Auricchio R, et al. Safety for patients with celiac disease of baked goods made of wheat flour hydrolyzed during food processing. *Clinical Gastroenterology and Hepatology* 2011;9(1):24-9.

Kaukinen 1999
Kaukinen K, Collin P, Holm K, et al. Wheat starch-containing gluten-free flour products in the treatment of coeliac disease and dermatitis herpetiformis. A long-term follow-up study. *Scandinavian Journal of Gastroenterology* 1999;34(2):163-9.

Kaukinen 2008
Kaukinen K, Salmi T, Collin P, et al. Clinical trial: gluten microchallenge with wheat-based starch hydrolysates in coeliac disease patients - a randomized, double-blind, placebo-controlled study to evaluate safety. *Alimentary Pharmacology and Therapeutics* 2008;28(10):1240-8.

Laurin 2002
Laurin P, Wolving M, Fälth-Magnusson K. Even small amounts of gluten cause relapse in children with celiac disease. *Journal of Pediatric Gastroenterology and Nutrition* 2002;34(1):26-30.

Lohiniemi 2000
Lohiniemi S, Maki M, Kaukinen K, et al. Gastrointestinal symptoms rating scale in coeliac disease patients on wheat starch-based gluten-free diets. *Scandinavian Journal of Gastroenterology* 2000;35(9):947-9.

Mayer 1991
* Mayer M, Greco L, Troncone R, et al. Compliance of adolescents with coeliac disease with a gluten free diet. *Gut* 1991;32(8):881-5.

Troncone R, Mayer M, Spagnuolo F, et al. Endomysial antibodies as unreliable markers for slight dietary transgressions in adolescents with celiac-disease. *Journal of Pediatric Gastroenterology & Nutrition* 1995;21:69-72.

Montgomery 1988
Montgomery AM, Goka AK, Kumar PJ, et al. Low gluten diet in the treatment of adult coeliac disease: effect on jejunal morphology and serum anti-gluten antibodies. *Gut* 1988;29(11):1564-8.

Peräaho 2003
Peräaho M, Kaukinen K, Paasikivi K, et al. Wheat-starch-based gluten-free products in the treatment of newly detected coeliac disease: prospective and randomized study. *Alimentary Pharmacology and Therapeutics* 2003;17(4):587-94.

Selby 1999
* Selby WS, Painter D, Collins A, et al. Persistent mucosal abnormalities in coeliac disease are not related to the ingestion of trace amounts of gluten. *Scandinavian Journal of Gastroenterology* 1999;34(9):909-14.

Faulkner-Hogg K. Assessment of gluten-free diets in an Australian population with coeliac disease and their impact on symptoms, mucosal, nutritional and metabolic parameters [PhD thesis]. Sydney, Australia: University of Sydney, 2004.

Faulkner-Hogg KB, Selby WS, Loblay RH. Dietary analysis in symptomatic patients with coeliac disease on a gluten-free diet: the role of trace amounts of gluten and non-gluten food intolerances. *Scandinavian Journal of Gastroenterology* 1999;34(8):784-9.

Stuart 1997
Stuart M. Different dietary gluten restrictions and symptom expression in a large population of people with coeliac disease [Essay]. Sydney, Australia: Royal Prince Alfred Hospital, 1997.

7.2 References to excluded studies

Biagi 2004
Biagi F, Campanella J, Martucci S, et al. A milligram of gluten a day keeps the mucosal recovery away: a case report. *Nutrition Reviews* 2004;62(9):360-3.

Buchanan 2008
Buchanan R, Dennis S, Gendel S, et al. Approaches to establish thresholds for major food allergens and for gluten in food. *Journal of Food Protection* 2008;71(5):1043-88.

Ejderhamn 1988
Ejderhamn JV B, Strandvik B. The long-term effect of continual ingestion of wheat starch-containing gluten-free products in coeliac patients. In: Kumar PJ ed. Coeliac Disease: One Hundred Years. 1988. Leeds, UK: Leeds University: 294-7.

Gibert 2006
Gibert A, Espadaler M, Angel Canela M, et al. Consumption of gluten-free products: should the threshold value for trace amounts of gluten be at 20, 100 or 200 p.p.m.? *European Journal of Gastroenterology & Hepatology* 2006;18(11):1187-95.

Gibert 2013
Gibert A, Kruizinga AG, Neuhold S, et al. Might gluten traces in wheat substitutes pose a risk in patients with celiac disease? A population-based probabilistic approach to risk estimation. *American Journal of Clinical Nutrition* 2013; 97(1):109-16.

Hamilton 1972
Hamilton JR, McNeill LK. Childhood celiac disease: response of treated patients to a small uniform daily dose of wheat gluten. *Journal of Pediatrics* 1972; 81(5):885-93.

Jansson 2001
Jansson UH, Gudjónsdóttir AH, Ryd W, et al. Two different doses of gluten show a dose-dependent response of enteropathy but not of serological markers during gluten challenge in children with coeliac disease. *Acta paediatrica* 2001;90:255-9.

Kumar 1988
Kumar PJ, Walker-Smith J, Milla P, et al. The teenage coeliac: follow up study of 102 patients. *Archives of Disease in Childhood* 1988;63(8):916-20.

Lahdeaho 2011
Lahdeaho ML, Maki M, Laurila K, et al. Small- bowel mucosal changes and antibody responses after low- and moderate-dose gluten challenge in celiac disease. *BMC Gastroenterology* 2011;11:129.

Mazzarella 2012
Mazzarella G, Salvati VM, Iaquinto G, et al. Reintroduction of gluten following flour transamidation in adult celiac patients: a randomized, controlled clinical study. *Clinical & Developmental Immunology* 2012, 329150 DOI: 10.1155/2012/329150.

Nayar 2009
CTRI/2009/091/000346. Nayar S. A clinical trial to study the effects of very minute quantity of wheat in the pediatric patients with celiac disease. http://apps.who.int/trialsearch/Trial2.aspx?TrialID=CTRI/2009/091/000346 (accessed 15 December 2015).

Pyle 2005
Pyle GG, Paaso B, Anderson BE, et al. Low-dose gluten challenge in celiac sprue: malabsorptive and antibody responses. *Clinical Gastroenterology & Hepatology* 2005;3(7):679-86.

Srinivasan 1996
Srinivasan U, Leonard N, Jones E, et al. Absence of oats toxicity in adult coeliac disease. *BMJ* 1996;313:13001. doi: 10.1136/bmj.313.7068.1300.

7.3 References to reviews

Akobeng 2008

* Akobeng AK, Thomas AG. Systematic review: tolerable amount of gluten for people with coeliac disease. *Alimentary Pharmacology and Therapeutics* 2008;27(11):1044-52.

Akobeng AK, Thomas AG. Tolerable amount of gluten for people with coeliac disease: a systematic review conducted on behalf of the Food Standards Agency (Project T07048). Manchester, UK: Booth Hall Children's Hospital and Manchester Children's University Hospitals NHS Trust, 2006.

Bruins Slot 2015

Bruins Slot ID, Bremer MGEG, Hamer RJ, et al. Part of celiac population still at risk despite current gluten thresholds. *Trends in Food Science & Technology* 2015;43(2):219-26.

Hischenhuber 2006

Hischenhuber C, Crevel R, Jarry B, et al. Review article: safe amounts of gluten for patients with wheat allergy or coeliac disease. *Alimentary Pharmacology and Therapeutics* 2006;23(5):559-75.

Appendix 1 Search strategies

MEDLINE via Ovid

Ovid MEDLINE In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily, Ovid MEDLINE and Ovid OLDMEDLINE <1946 to Present> [Searched 11 Nov 2015]		
#	Search Statement	Results
1	Celiac Disease/	17741
2	c?eliac$.ti,ab.	22742
3	1 or 2	28183
4	exp Glutens/	7171
5	Diet, Gluten-Free/	1166
6	(gluten$ or gliadin$).ti,ab.	10849
7	4 or 5 or 6	12876
8	exp Toxicity Tests/	95055
9	Intestinal Mucosa/pa [Pathology]	19571
10	adverse effects.fs.	1469516
11	(threshold$ or safe$ or dose$ or dosage$ or adverse$).ti,ab.	2056807
12	8 or 9 or 10 or 11	3251085
13	3 and 7 and 12	2179
14	limit 13 to (english language and humans)	**1717**

Embase via Ovid

Embase <1974 to 2015 November 09> [Searched 11 Nov 2015]		
#	Search Statement	Results
1	celiac disease/	24144
2	c?eliac$.ti,ab.	29848
3	1 or 2	35665
4	gluten/	6808
5	gliadin/	2967
6	gluten free diet/	5954
7	(gluten$ or gliadin$).ti,ab.	13572
8	4 or 5 or 6 or 7	16832
9	exp toxicity testing/	60200
10	intestine mucosa/	32955
11	si.fs.	741590
12	(threshold$ or safe$ or dose$ or dosage$ or adverse$).ti,ab.	2723306
13	9 or 10 or 11 or 12	3243603
14	3 and 8 and 13	1658
15	limit 14 to (human and english language)	**1174**

PubMed

PubMed [Searched 11 November 2015]		
#	Search Statement	Results
1	((((Celiac Disease[MeSH Terms]) OR ((celiac*[TIAB] OR coeliac*[TIAB])))) AND (((Glutens[MeSH Terms]) OR Diet, Gluten Free[MeSH Terms]) OR ((gluten*[TIAB] OR gliadin*[TIAB])))) AND (((Toxicity Tests[MeSH Terms]) OR ((Intestinal Mucosa[mh] AND pathology[sh]))) OR ((adverse effects[sh] OR threshold*[TIAB] OR safe*[TIAB] OR dose*[TIAB] OR dosage*[TIAB] OR adverse*[TIAB])))	2275
2	((((((Celiac Disease[MeSH Terms]) OR ((celiac*[TIAB] OR coeliac*[TIAB])))) AND (((Glutens[MeSH Terms]) OR Diet, Gluten Free[MeSH Terms]) OR ((gluten*[TIAB] OR gliadin*[TIAB])))) AND (((Toxicity Tests[MeSH Terms]) OR ((Intestinal Mucosa[mh] AND pathology[sh]))) OR ((adverse effects[sh] OR threshold*[TIAB] OR safe*[TIAB] OR dose*[TIAB] OR dosage*[TIAB] OR adverse*[TIAB])))) AND pubmednotmedline[sb]	**22**

Cochrane Library

Cochrane Library [Searched 11 November 2015]		
#	Search Statement	Results
1	MeSH descriptor: [Celiac Disease] explode all trees	225
2	celiac*:ti,ab,kw	460
3	coeliac*:ti,ab,kw	133
4	#1 or #2 or #3	490
5	MeSH descriptor: [Glutens] explode all trees	114
6	gluten*:ti,ab,kw	246
7	gliadin*:ti,ab,kw	76
8	#5 or #6 or #7	278
9	#4 and #8	**180***

* Other reviews 4, Trials 170, Technology assessments 1, Economic evaluations 5

Scopus

Scopus [Searched 16 November 2015]		
#	Search Statement	Results
1	((TITLE-ABS-KEY (celiac*)) OR (TITLE-ABS-KEY (coeliac*))) AND ((TITLE-ABS-KEY (gluten*)) OR (TITLE-ABS-KEY (gliadin*))) AND ((TITLE-ABS-KEY (threshold*)) OR (TITLE-ABS-KEY (safe*)) OR (TITLE-ABS-KEY (dose*)) OR (TITLE-ABS-KEY (dosage*)) OR (TITLE-ABS-KEY (adverse*))) AND (LIMIT-TO (LANGUAGE , "English"))	**795**

Web of Science

Core collection (Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years)

Web of Science [Searched 16 November 2015]		
#	Search Statement	Results
1	TOPIC: (celiac*)	21,916
2	TOPIC: (coeliac*)	5,690
3	#2 OR #1	25,770
4	TOPIC: (gluten*)	18,002
5	TOPIC: (gliadin*)	5,593
6	#5 OR #4	20,416
7	TOPIC: (threshold*)	415,811

8	TOPIC: (safe*)	778,524
9	TOPIC: (dose*)	986,262
10	TOPIC: (dosage*)	122,627
11	TOPIC: (adverse*)	363,650
12	#11 OR #10 OR #9 OR #8 OR #7	2,387,230
13	#12 AND #6 AND #3	499
14	#12 AND #6 AND #3 Refined by: LANGUAGES**:** (ENGLISH)	**480**

Google Scholar

Google Scholar [Searched 16 November 2015]		
#	Search Statement	Results
1	(celiac OR coeliac) AND (gluten OR gliadin) AND (threshold OR safe OR dose OR dosage OR adverse)	Retrieved 25,700 results. Downloaded first **260** citations

Open Grey

Open Grey http://www.opengrey.eu/ [Searched 20 November 2015]		
#	Search Statement	Results
1	celiac*	39
2	coeliac*	46
3	celiac* OR coeliac*	81
4	gluten*	148
5	gliadin*	34
6	gluten* or gliadin*	166
7	(celiac* OR coeliac*) AND (Gluten* OR gliadin*)	40

All 40 records were theses, of which 20 were in French and five in Czech.

Grey Literature Report in Public Health

Searched http://www.greylit.org/ on 20 November 2015 using the terms: coeliac, celiac, coeliacs, celiacs, gluten, glutens, gliadin and gliadins. Searched within title, summary and full-text. Boolean searching not supported. No records retrieved.

Clinical trials registers

WHO International Clinical Trials Registry Platform (ICTRP) www.who.int/ictrp/ [Searched 15 December 2015]		
#	Search Statement	Results
1	((coeliac or celiac) in condition) AND ((gluten* or gliadin*) in intervention)	71

ClinicalTrials.gov www.clinicaltrials.gov [Searched 15 December 2015]		
#	Search Statement	Results
1	Celiac AND gluten[1] and Celiac and gliadin [[1] retrieves coeliac and celiac]	87

Appendix 2 Risk of bias assessments for RCTs

Catassi 1993

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "… patients were randomly assigned to either group A or B." (p1516) Comment: Method not stated.
Allocation concealment (selection bias)	Unclear risk	Comment: Method not stated.
Blinding of participants (performance bias)	Low risk	Quote: 100 mg or 500 mg of "daily gliadin was administered with some sugar in a single dose." (p1516) Comment: Gliadin was delivered in capsules, and both patients and investigators were blinded to the dose (C. Catassi, personal communication).
Blinding of outcome assessment (detection bias) 1. histology 2. serology 3. clinical symptoms	 Low risk Low risk Low risk	Quote: "The morphometric analysis of the sections was performed on at least 10 villi by a computerised image analyser IBAS-AT Kontron (Munich)." (p1516) Comment: Additionally, the staff examining the biopsies were unaware of group assignment (C. Catassi, personal communication). Serology tests were performed by standard laboratory methods and clinical symptoms were assessed by investigators who were unaware of group assignment.
Incomplete outcome data (attrition bias)	Low risk	Quote: "All patients completed the study protocol." (p1516) Comment: Additionally, Table 1 shows the results for all patients.
Selective reporting (reporting bias)	Low risk	Comment: All outcomes declared are reported.
Other bias	Low risk	Comment: No other bias detected.

Catassi 2007

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "This was a prospective, multicentre, placebo-controlled, double-blind, randomized trial … Patients were randomly assigned (by the coordinating centre) …" (p161) Comment: Block randomisation was performed by the coordinating centre (C. Catassi, personal communication).
Allocation concealment (selection bias)	Low risk	Comment: Randomisation was performed centrally by the coordinating centre (C. Catassi, personal communication).
Blinding of participants (performance bias)	Low risk	Quote: "… a capsule containing either 10 mg purified gluten, 50 mg purified gluten, or 50 mg cornstarch as a placebo (double-blind microchallenge)." (p161) Comment: Capsule preparation was performed by the coordinating centre. All capsules looked the same (C. Catassi, personal communication).
Blinding of outcome assessment (detection bias) 1. histology 2. serology 3. clinical symptoms	 Low risk Low risk Low risk	Quote: Biopsy specimens were "examined in batches by 2 pathologists … who were blinded to subject assignment." Analysis was performed using a "computerized image analyzer" (p 161) Comment: Patients and investigators were blinded. (Confirmed in personal communication from C. Catassi.)
Incomplete outcome data (attrition bias)	Low risk for histology Unclear risk for symptoms	Comment: Reasons for exclusion from the study prior to the challenge (n=7) and for not completing the challenge (n=3) were provided. One of the drop-outs challenged with 10 mg gluten withdrew because of symptoms. Although this has little impact on the primary outcome (small bowel histology) it is potentially relevant to clinical outcomes.
Selective reporting (reporting bias)	Low risk for histology Unclear risk for symptoms	Comment: Results of small bowel mucosal morphology, clinical and serological examinations are reported. Clinical symptoms are only reported narratively as showing no significant change between the three groups.
Other bias	Low risk	Comment: No other bias detected.

Greco 2011

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Twelve patients were randomized to treatments by random number. The first 6 patients consumed NFBG, the following 2 patients consumed S1BG, and the last 5 patients consumed S2BG." (p25) Comment: Despite the phrase "randomized by random number", it is unclear how this was performed when patients were allocated consecutively to the three treatment groups. Note, although it states that 12 patients were randomised, the results are provided for 13 patients (6+2+5).
Allocation concealment (selection bias)	Unclear risk	Comment: Method not stated.
Blinding of participants (performance bias)	Unclear risk	Comment: No information provided.
Blinding of outcome assessment (detection bias) 1. histology 2. serology 3. clinical symptoms	 Unclear risk Low risk Unclear risk	Comment: The serological, morphometric and immunohistochemistry analyses were conducted by standard measures. There was no information regarding blinding of outcome assessors for the histologic readouts or dietary interviews.
Incomplete outcome data (attrition bias)	Low risk	Comment: Two patients in the natural flour baked goods (NFBG) group ceased the challenge after four weeks due to symptoms. These patients still provided outcome data, though it is unclear if this was taken at 4 weeks or 60 days. All patients in the sourdough 1 baked goods (S1BG) and sourdough 2 baked goods (S2BG) groups completed the 60-day challenge.
Selective reporting (reporting bias)	High risk	Comment: The declared outcomes of small morphology and coeliac serology are reported. Clinical reasons for the two dropouts in the NFBG group were reported. However, this study states that 12 patients were randomised but presents results for 13 patients. It also refers to a table of results for the S1BG group which is not present in the paper and reports EMA results for only one of the three groups. We therefore judge this study to be at high risk of reporting bias.
Other bias	Low risk	Comment: No other bias detected.

Kaukinen 2008

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Ninety eligible patients were randomized either to glucose syrups (A), maltodextrins (B) or to the placebo group (C) in a ratio of 1:1:1 using a double dummy technique." (p1241) Comment: Unclear if the phrase "double dummy technique" refers to the double blind nature of the study or to some aspect of how randomisation was implemented.
Allocation concealment (selection bias)	Unclear risk	Comment: Method not stated.
Blinding of participants (performance bias)	Low risk	Quote: "The products were powder packed in single-use sachets to be diluted in 300 mL water. … Placebo sachets contained no gluten." (p1242) Comment: Judged to be low risk on the basis that the sachets were assumed to be identical in look and taste.
Blinding of outcome assessment (detection bias) 1. histology 2. serology 3. clinical symptoms	 Low risk Low risk Low risk	Quote: "…the specimens were evaluated by the same investigator without prior knowledge of the history or findings." (p1242) Comment: For small bowel biopsies we judge this to mean the investigator was unaware of group assignment. The serology and other laboratory values unlikely to be affected. Self-reported gastrointestinal symptoms deemed low risk because of participant blinding.
Incomplete outcome data (attrition bias)	Low risk	Comment: Seven participants withdrew due to abdominal symptoms and one due to non-compliance. Withdrawals were spread between the groups and withdrawal does not appear to be linked to the level of consumption of sachets.

		Quote: "Only four of the 90 patients did not consent to complete the last study visit and follow-up small-bowel biopsy. However, when we hypothesized that if these four patients would have developed abnormal small bowel mucosal morphology with total villous atrophy and massive inflammation, the main results of the study did not change" (p1246). Comment: Authors have performed a worst-case scenario based on missing data.
Selective reporting (reporting bias)	Low risk	Comment: All outcomes declared (small bowel mucosal morphology, serological and chemical analysis and gastrointestinal symptoms) are reported.
Other bias	Low risk	Comment: No other bias detected.

Peräaho 2003

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: "The randomization was carried out using random-number tables with permuted blocks." (p588)
Allocation concealment (selection bias)	Low risk	Quote: "Each patient entered the trial before random treatment assignment was revealed." (p588) Comment: Seen as confirming that the permuted blocks successfully concealed allocation.
Blinding of participants (performance bias)	Unclear risk	Comment: Although patients were aware of their group assignment and given advice from the dietician on the two diets before the start of the study, it is unclear what impact this has on patients' disposition towards the intervention.
Blinding of outcome assessment (detection bias) 1. histology 2. serology 3. clinical symptoms	 Low risk Low risk High risk	Quote: "… all specimens were evaluated by the same investigator, who had no previous knowledge of the disease history or laboratory findings." (p589) Comment: For small bowel biopsies we judge this to mean the investigator was unaware of group assignment. Serology, chemical analysis and bone mineral density unlikely to be affected. Comment: Unclear if the gastrointestinal symptom rating scale (GSRS) was self-administered or via interview but knowledge of the intervention may affect responses.
Incomplete outcome data (attrition bias)	Unclear risk	Quote: "Two patients in group I decided to discontinue the study after 2 months, as they found the diet too convoluted. In addition, four patients in group I and two in group II did not follow a strict gluten-free diet. Thus, 23 in group I and 26 in group II completed the study with a proper diet." (p590) Comment: There is disparity between the groups in terms of numbers who completed the study, and this may impact on the results.
Selective reporting (reporting bias)	Low risk	Small bowel morphology, clinical symptoms, serology and patient-reported gastrointestinal symptoms and quality of life are all reported.
Other bias	High risk	The authors assumed that the naturally gluten-free diet contained no gluten, however, they acknowledge that these products may be contaminated. Although in principle testing the difference between a no-gluten diet and a very low level gluten diet, it is very difficult to be sure of the gluten levels consumed by the two groups.

Appendix 3 Matrix comparing overlap of studies in reviews

	Hischenhuber 2006	Akobeng 2008	Bruins Slot 2015	Cochrane 2016
Dissanayake 1974	✓	✓		✓
Baker 1975	✓	✓		✓
Ciclitira 1984	✓	✓		✓
Ciclitira 1985	✓	✓	*Only included studies published in last 20 years*	✓
Ejderhamn 1988	✓			
Kumar 1988	✓			
Montgomery 1988	✓	✓		✓
Mayer 1991	✓			✓
Catassi 1993	✓	✓	✓	✓
Troncone 1995	✓		✓	
Srinivasan 1996			✓	
Chartrand 1997	✓	✓	✓	✓
Stuart 1997				✓
Kaukinen 1999	✓	✓	✓	✓
Selby 1999	✓	✓		✓
Lohiniemi 2000	✓	✓	✓	✓
Laurin 2002			✓	✓
Peräaho 2003	✓	✓		✓
Collin 2004		✓	✓	✓
Biagi 2004			✓	
Catassi 2007		✓	✓	✓
Kaukinen 2008				✓
Greco 2011			✓	✓

Appendix 4 ROBIS assessment of systematic reviews

Akobeng 2008

Summary of assessment for each domain

1. Concerns regarding specification of study eligibility criteria	Low risk	There is no reference to or mention of a protocol or pre-specified plan, so can't tell if objectives and eligibility criteria were pre-specified. Overall no major concerns.
2. Concerns regarding methods used to identify and/or select studies	Low risk	No important concerns.
3. Concerns regarding methods used to collect data and appraise studies	High risk	The authors note missing information about study characteristics that may impact on the interpretation of findings by both reviewers and readers. There was no numeric reporting of results, and thus it is impossible to tell whether all relevant data were collected from each study (i.e. quantitative data were missing from studies) or whether the reviewers have chosen not to report quantitative results/effect estimates.
4. Concerns regarding synthesis and findings	High risk	The authors do not adequately address consistency of findings across studies or the quality of the evidence in reporting the results/key findings (e.g. reporting a 'clear association' based on results from one study of 20 children p1050).

Describe whether conclusions were supported by the evidence:

While there are some overall remarks about the small number of good quality studies, there isn't appropriate integration of information about study quality with the reporting of results. For this reason, concerns relating to domain 4 were judged not to be adequately addressed. The absence of any quantitative data makes it impossible to tell whether the authors avoided emphasising results based on statistical significance. While the overall conclusions about the body of evidence are cautious, some of the findings are worded in a such a way that implies greater confidence in the findings from individual studies than is appropriate given the quality of the evidence and the data presented.

Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?	Probably not
Was the relevance of identified studies to the review's research question appropriately considered?	Yes
Did the review authors avoid emphasising results on the basis of their statistical significance?	No information
Overall risk of bias	High

Bruins Slot 2015

Summary of assessment for each domain

1. Concerns regarding specification of study eligibility criteria	Low risk	No indication that the methods of the review were pre-specified and the eligibility criteria are incompletely specified. There is potential for some relevant studies to be excluded.
2. Concerns regarding methods used to identify and/or select studies	High risk	No information reported about the screening of studies or inclusions/exclusion decisions. Can't tell if potentially relevant studies may have been missed in the screening process.
3. Concerns regarding methods used to collect data and appraise studies	High risk	No information reported about data collection (what data, who collected, how, data checks) and no risk of bias assessment.
4. Concerns regarding synthesis and findings	High risk	There is no risk of bias assessment and, while the authors mention that trials are needed, there is no mention of potential biases arising from the design/conduct/reporting of individual studies. Nor is there any mention of uncertainty in the findings arising from the quality of the evidence overall.

Describe whether conclusions were supported by the evidence:

The authors do not mention any of the limitations or potential biases in their review. The description of the methods of the review is so limited that it is not possible to assess potential biases introduced through the design/conduct/reporting of the review. Perhaps most importantly, the authors don't address the implications of not assessing the risk of bias (or any aspect of quality) of the included studies and how potential biases in the included studies might influence their findings. Almost all of the studies are small, especially two of the three trials that appear to be contributing to the overall conclusions. This means the findings of the review are likely to be uncertain due to imprecision, yet this is not reflected in the author's conclusions. For these reason the review is rated as at high risk of bias.

The relevance of the identified studies seems adequately addressed, although two of the three trials are in children/adolescents and there is no consideration of whether these results are applicable to adults. Although the authors do not report or mention 'statistical significance' or p values, there is no indication that findings from any particular study/outcome/comparison are over-emphasised.

Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?	No
Was the relevance of identified studies to the review's research question appropriately considered?	Probably yes
Did the review authors avoid emphasising results on the basis of their statistical significance?	Probably yes
Overall risk of bias	High