Article:
de Graaf, M.C.G., Lawton, C.L., Croden, F. et al. (12 more authors) (2024) The effect of expectancy versus actual gluten intake on gastrointestinal and extra-intestinal symptoms in non-coeliac gluten sensitivity: a randomised, double-blind, placebo-controlled, international, multicentre study. The Lancet Gastroenterology & Hepatology, 9 (2). pp. 110-123. ISSN 2468-1253

https://doi.org/10.1016/s2468-1253(23)00317-5

The effect of expectancy versus actual gluten intake on gastrointestinal symptoms in non-coeliac gluten sensitivity: a randomised, double-blind, placebo-controlled, international multicentre study

Authors

Marlijne C.G. de Graaf MSc[1,2], Clare L. Lawton PhD[3], Fiona Croden BSc[3], Agnieszka Smolinska PhD[2,4], Bjorn Winkens PhD[5], Martine A.M. Hesselink MSc[1,2], Gonny van Rooy MD[1,2], Peter L. Weegels PhD[6,7], Prof. Peter R. Shewry PhD[8], Prof. Lesley A. Houghton PhD[9,10], Prof. Ben J.M. Witteman PhD[11,12], Prof. Daniel Keszthelyi PhD[1,2], Prof. Fred J.P.H. Brouns PhD[2,13], Prof. Louise Dye PhD[3,14]*, Prof. Daisy M.A.E. Jonkers PhD[1,2]*

* Shared last author.

Affiliations

[1] Division Gastroenterology-Hepatology, Department of Internal Medicine, Maastricht University Medical Center+, Maastricht, the Netherlands

[2] NUTRIM School of Nutrition and Translational Research in Metabolism, Faculty of Health, Medicine and Life Sciences, Maastricht University, the Netherlands

[3] School of Psychology, University of Leeds, Leeds, United Kingdom

[4] Department of Pharmacology and Toxicology, Maastricht University, Maastricht, the Netherlands

[5] Department of Methodology and Statistics, CAPHRI, Care and Public Health Research Institute, Maastricht University, Maastricht, The Netherlands

[6] Laboratory of Food Chemistry, Wageningen University and Research, Wageningen, the Netherlands

[7] European Bakery Innovation Centre, Sonneveld Group BV, Papendrecht, the Netherlands

[8] Rothamsted Research, Harpenden, United Kingdom

[9] Division of Gastroenterology and Surgical Sciences, Leeds Institute of Medical Research, University of Leeds, Leeds, United Kingdom

[10] Division of Gastroenterology and Hepatology, Mayo Clinic, Jacksonville, Florida, USA

[11] Division Gastroenterology-Hepatology, Gelderse Vallei Hospital, Ede, the Netherlands
[12] Division of Human Nutrition, Wageningen University and Research, Wageningen, the Netherlands
[13] Department of Human Biology, Maastricht University, Maastricht, the Netherlands
[14] School of Food Science and Nutrition, University of Leeds, Leeds, United Kingdom

Corresponding author:
Prof. Daisy M.A.E. Jonkers
Maastricht University Medical Center+
PO Box 5800
6202 AZ Maastricht
The Netherlands
E-mail: d.jonkers@maastrichtuniversity.nl
Tel: +31 43 38 84 266

Summary

Background: Many individuals reduce their gluten intake based on their belief that this reduces gastrointestinal (GI) symptoms. Symptoms may be affected by negative expectancy. Therefore, we investigated the effects of expectancy versus actual gluten intake on symptoms in non-coeliac gluten sensitivity (NCGS).

Methods: This randomised, double-blind, placebo-controlled trial was conducted at the University of Leeds (United Kingdom), Maastricht University, and Wageningen University (the Netherlands). Eighty-four participants (18-70 years), with self-reported NCGS, exclusion of coeliac disease and wheat allergy, and reporting symptoms within 8 hours of gluten consumption, were recruited. Participants were randomised (1:1, with blocks of 8, stratified by study site and sex) to one of four groups based on the expectation to consume "gluten-containing" (E+) or "gluten-free" (E-) oat bread for breakfast and lunch (two slices each), and actual intake of gluten-containing (G+) or gluten-free (G-) oat bread. Apart from the expectancy, participants, investigators, and those assessing outcomes were blinded to the actual gluten assignment. The primary outcome was overall GI symptoms evaluated by per-protocol analysis of visual analogue scale (VAS) ratings at baseline (before breakfast) and hourly for 8 hours, with a lunch served after 4 hours. The study was registered at ClinicalTrials.gov, NCT05779358, and has ended.

Findings: Between October 19, 2018, and February 14, 2022, participants were randomised into E+G+ (n=21 including one exclusion (female) due to failure to understand the test day instructions, 16 female and 5 male), E+G- (n=21, 19 female and 2 male), E-G+ (n=20, 17 female and 3 male), and E-G- (n=22, 19 female and 3 male). Mean overall GI symptoms were significantly higher in E+G+ (VAS 16·6mm [95%CI 13·1-20·0mm]) compared with E-G+ (VAS 6·9mm [95%CI 3·5-10·4mm], p=0·0010) and E-G- (VAS 7·4mm [95%CI 4·2-10·7mm], p=0·0016), but not E+G- (VAS 11·7mm [95%CI 8·3-15·1mm], p=0·28), nor between E+G- versus E-G+ (p>0·99), E+G- versus E-G- (p>0·99), and E-G+ versus E-G- (p>0·99). Adverse events were reported by two participants in E+G- (itching jaw; feeling lightheaded and stomach rumble), and one participant in E-G+ (vomiting).

Interpretation: The combined effect of expectancy and actual gluten intake had the largest effect on GI symptoms, reflecting a nocebo effect, although an additional effect of gluten could not be ruled out. The results of this study necessitate further research into possible involvement of gut-brain interaction in NCGS.

Funding: The Dutch Government “TKI - Top Knowledge Institute”, AB-Mauri Bakery Ingredients, Made, the Netherlands; Baking Industry Research Trust, Wellington, New Zealand; Borgesius Holding BV-Albert Heijn, Stadskanaal, the Netherlands; CSM Innovation Bakery Center, Bingen, Germany; CYMMIT, Texcoco, Mexico; DSM Food Specialties, Delft, the Netherlands; Fazer Bakeries Oy, Helsinki, Finland; Health Grain Forum, Vienna, Austria; ICC-Intl., Vienna, Austria; IWGA, Kansas 66210, USA; Lantmännen EK, Stockholm, Sweden; Mondelez, Saclay, France; Nederlands Bakkerij Centrum, Wageningen, the Netherlands; Nutrition et Sante, Revel, France; Puratos BV, Groot Bijgaarden, Belgium; Rademaker BV-Bakery equipment, Culemborg, the Netherlands; Sonneveld Group BV, Papendrecht, the Netherlands; Zeelandia-HJ Doeleman BV, Zierikzee, the Netherlands.

Research in context

Evidence before this study

We searched PubMed for randomised controlled trials, systematic reviews and meta-analyses published in English from database inception up to May 31, 2023, using the search terms (“non-celiac gluten sensitivity” OR “non-coeliac gluten sensitivity” OR “nonceliac gluten sensitivity” OR “noncoeliac gluten sensitivity” OR “NCGS”) AND (“nocebo” OR “expectancy” OR “expectation” OR “perception”) AND (“randomized controlled trials” OR “systematic review” OR “meta-analysis”). This search yielded one narrative review from 2019, which concluded that a large nocebo effect had been found in some studies. Central to this conclusion was the pooled analysis of all DBPC gluten challenge studies done in NCGS subjects up to March 31, 2016 which showed a nocebo response in 40% of participants.

Additionally, we searched for recent systematic reviews and meta-analyses on NCGS in general. The most recent systematic review available in English, including all articles until June 1, 2020, concluded that the vast majority of studies reported a predominant nocebo effect, which the authors considered intrinsically related to the DBPC design. Moreover, the authors asserted that the carry-over and order effects found in previous studies are strictly connected to the psychological background of the study participants, and these characteristics should be considered in all DBPC studies. We found no studies specifically designed to investigate the effect of the nocebo effect in NCGS.

Added value of this study

To our knowledge, this is the first study designed to investigate the role of the nocebo effect in NCGS. Our randomised, double-blind, placebo-controlled, international multicentre study showed that the combination of expectancy to receive gluten and actual gluten intake resulted in the highest scores for overall GI symptoms, abdominal discomfort, and bloating. Repeated exposure further accentuated symptom scores and differences between the intervention groups. We found no significant effect of actual gluten intake within each expectancy group. Although an additional effect of gluten could not be ruled out, our findings indicate that the nocebo effect plays a significant role in symptom occurrence in NCGS. The nocebo effect and the possible involvement of gut-brain interaction warrant further research.

Implications of all the available evidence

Our study is consistent with previous research showing a nocebo effect. This study is the first to explicitly manipulate expectancy and confirm the nocebo effect in NCGS. The fact that the expectancy to consume gluten was associated with symptoms points towards possible involvement of gut-brain interaction in symptom occurrence in NCGS. Furthermore, there was a discernible effect of gluten. Although this was only found in the presence of expectancy to receive gluten, these findings also highlight the need to elicit possible biological mechanisms underlying gluten-related symptoms.

Introduction

Wheat is the most important staple food consumed in the Western world. Whole grain wheat products are an important source of carbohydrates, dietary fibres, proteins, vitamins, minerals, and phytochemicals, and globally provide a major contribution to daily energy intake and a healthy diet.[1] Based on epidemiological evidence, the consumption of whole grain cereal foods has been associated with several beneficial health effects, showing a reduced risk of for example obesity, type 2 diabetes, cardiovascular disease, cancer, overall and cause-specific mortality.[2-5]

However, wheat products can also elicit adverse (immune-mediated) effects, such as in coeliac disease (CD) and wheat allergy (WA). In addition, a substantial proportion of the general population is avoiding or reducing their consumption of wheat products due to self-reported symptoms following wheat intake, without having CD or WA. Gluten proteins (gliadins and glutenins) are often attributed to be the wheat component responsible for inducing negative reactions, referred to as non-coeliac gluten sensitivity (NCGS).

Individuals with NCGS mostly report gastrointestinal (GI) symptoms such as abdominal pain or discomfort, bloating, and altered stool patterns and, to a lesser extent, extra-intestinal symptoms like tiredness and headache.[6] The estimated prevalence of NCGS ranges from 0·6-13%.[7] Due to a lack of biomarkers, diagnosis is defined by the Salerno Experts' Criteria[6] including a double-blind, placebo-controlled (DBPC) gluten challenge, which is not always feasible in clinical practice.

Furthermore, previous studies have reported the presence of NCGS in 6·8-46·1% of irritable bowel syndrome (IBS) patients, indicating substantial overlap between these conditions.[8] IBS is a disorder of gut-brain interaction (DGBI), characterised by recurrent abdominal pain and altered bowel habits that affects 5-10% of the population.[9] Wheat is among the top five of foods reported by IBS patients to trigger their symptoms.[8] Gluten-free diets (GFD) are becoming more popular because of perceived symptom alleviation as well as negative media attention about gluten.[10] However, a GFD is associated with an increased risk of nutritional deficiencies.[11]

To date, little evidence is available on the role of gluten in symptom occurrence in NCGS, nor on the underlying mechanisms. Previous studies suggest involvement of the immune system, intestinal inflammation, dysbiosis and/or altered barrier function, but the exact mechanism remains unclear.[12] Furthermore, the role of psychological

factors cannot be ruled out. Anxiety and depression are more prevalent in NCGS patients compared to the general population.[13]

Additionally, Biesiekierski and colleagues'[14] DBPC cross-over study showed significant worsening of overall GI symptoms and abdominal pain irrespective of dietary intervention (whether placebo, low-gluten or high-gluten diet). Remarkably, the symptom scores were highest with the first treatment participants received, irrespective of the actual treatment, suggesting a nocebo effect. The significance of a nocebo effect was further highlighted in a pooled analysis of 10 DBPC gluten-challenge trials which found 40% of subjects showed similar or increased symptoms in response to placebo.[15]

These findings indicate that expectation may mediate a nocebo effect, for example by influencing GI sensory and motor functions.[16] The relevance of the nocebo effect has previously been demonstrated in IBS patients, with a pooled nocebo response rate of 32% to clinical drug trials.[17] However, to our knowledge, the contribution of negative expectation about gluten consumption in NCGS symptom occurrence has never been investigated. Exploring this may contribute to understanding of the pathophysiology of NCGS, as well as improving diagnostic procedures and dietary and/or psychological treatment of these individuals.

Therefore, we aimed to investigate the effects of expectancy about gluten intake versus actual gluten intake on GI and extra-intestinal symptoms in individuals with self-reported NCGS. In addition, we aimed to investigate the role of psychological factors in these symptoms, and the effect on mood. We hypothesised that expected gluten intake, but not actual gluten intake, would increase symptom severity. As an expectancy effect would reflect a psychological process, we further hypothesised that symptom ratings would be higher in individuals who scored higher on psychological measures of anxiety, depression and somatic symptoms, and that negative affect would be higher in those who reported higher levels of symptoms.

Methods

Study design

This randomised, double blind, placebo-controlled, international multicentre study was conducted at the University of Leeds (United Kingdom (UK)), Maastricht University and Wageningen University and Research (the Netherlands (NL)), between October 19, 2018 and February 14, 2022 (see appendix p3). A cross-over design was not deemed

feasible as this could have undermined or revealed the expectancy part of the study. The study protocol was written in close collaboration between the University of Leeds and Maastricht University, and was approved by the Faculty Research Ethics Committee of the University of Leeds, and by the Medical Ethics Committee of the Academic Hospital Maastricht/Maastricht University, which was also accepted by the Board of Directors of Wageningen University and Research. Although the study design was identical in each country, the study was planned to be analysed separately by country in order to compare the results between countries. However, because of recruitment delays due to the COVID-19 pandemic, the study was terminated early and, after performing a blinded interim analysis to confirm there were no significant differences between countries, data were pooled. The study protocol is available online (https://mdl.mumc.nl/sites/mdl/files/2023-08/Onderzoeksprotocol%20WoW%20studie.pdf). The study was conducted in compliance with Good Clinical Practice, the Declaration of Helsinki (2013), the Food and Drug Administration (UK), and the Medical Research Involving Human Subject Act (NL).

Participants

Participants were recruited via advertisements on social media, patient association websites, notice boards on the university campuses and in local public areas, and local newspapers.

After receiving written and verbal information, interested participants were invited for a screening visit to assess eligibility. Males and females (by self-reported sex) aged 18-70 years, with self-reported GI symptoms within 8 hours after a single intake of gluten-containing products were included. Participants had to be asymptomatic or only mildly symptomatic (overall GI symptom score $\leq$ 30mm on visual analogue scale (VAS)), while following a GFD or gluten-restricted diet as defined by a Biagi and colleagues[18] score of 2-4, for at least one week prior to and throughout study participation. Concurrent medication had to be stable for at least 6 weeks prior to and during the study. Participants were excluded if they had been diagnosed with CD, WA, other organic GI diseases, any malignancies or any other disease which could interfere with GI function, or if they previously had major abdominal surgery or radiotherapy which could interfere with GI function (participants with uncomplicated appendectomy, cholecystectomy and hysterectomy were considered eligible if more than six months

ago). If CD was not excluded by previous serology or upper GI endoscopy, and participants still consumed some gluten or were willing to re-introduce gluten into their diet for at least six weeks, an additional visit was scheduled for serological testing (total immunoglobulin A (IgA) and anti-tissue transglutaminase IgA) to exclude CD. Furthermore, use of antibiotics, probiotics or prebiotics, participation in other studies 14 days prior to the study, excessive use of alcohol or drugs, and intentional weight-loss during the study period were not allowed. Females could not be pregnant or lactating. Current smokers were included but asked not to smoke during the test day. Participants had to have sufficient knowledge of Dutch or English to understand the nature of the study, give consent and complete the measures.
All volunteers gave their written informed consent prior to participation.

Randomisation and masking

The intervention was based on a combination of expectancy and actual gluten intake (see appendix p4). On the test day, participants were informed by the researcher that they had been allocated to either consume gluten-containing (E+) or gluten-free (E-) bread. They then received two slices of gluten-containing (G+) or gluten-free (G-) oat bread. This resulted in four groups: E+G+ (expectancy to consume gluten-containing bread, combined with actual intake of gluten-containing bread); E+G- (expectancy to consume gluten-containing bread, combined with actual intake of gluten-free bread); E-G+ (expectancy to consume gluten-free bread, combined with actual intake of gluten-containing bread); and E-G- (expectancy to consume gluten-free bread, combined with actual intake of gluten-free bread). Randomisation was done by a colleague unconnected with the trial. The randomisation list was generated using a publicly available procedure (http://randomizer.org), allocating participants (1:1) to the four treatment conditions using block randomisation (block sizes of eight) and stratified by study site and sex. The investigator (involved in the trial) enrolled participants. The unconnected colleague provided the investigator with a unique randomisation number (WoW 701-744 or WoW 401-444), which indicated the expectancy (7xx = E+ and 4xx = E-), and corresponded to the subject code on the study bread label. The study breads were identical in appearance, and the actual intervention (G+ or G-) could not be identified from this code. The participant and the investigator were both blind to the actual intervention, and participants were not aware of the expectancy part of the study. Data analysis was executed before unblinding. Therefore, the colleague

responsible for the randomisation labelled the intervention groups as A-D for further analysis. They provided the investigator with the blinded intervention groups, labelled A, B, C, and D. After completing the analysis, the code was revealed (A = E+G+, B = E+G-, C = E-G+, D = E-G-). To maintain secrecy about the study design, special approval was granted by the Dutch Central Committee on Research Involving Human Subjects (CCMO) (reference number CCMO18.0344/lvV/ek) for the expectancy part of the study and for delayed registration on ClinicalTrials.gov.

Procedures

At the screening visit, the researcher assessed the in- and exclusion criteria by checking demographic characteristics (including self-reported sex (male or female)), medical history, comorbidities, the Biagi questionnaire,[18] usual symptom(s) experienced after gluten consumption, and VAS for overall GI symptom score during the preceding week (*i.e.* while on a GFD). Additionally, Rome IV criteria for IBS and functional dyspepsia (FD) were assessed.

When CD had not been excluded already, an additional visit was scheduled prior the screening visit (*i.e.* prior to starting the GFD) for serological testing (total IgA and anti-tTG IgA) to exclude CD.

Participants completed the Generalized Anxiety Disorder assessment (GAD-7), Patient Health Questionnaire-9 (PHQ-9), and the Patient Health Questionnaire-15 (PHQ-15) to assess anxiety, depression, and somatic symptoms, respectively, at home between the screening visit and the test day.

A 100% gluten-free oat-based bread mix (SonFit Gluten Free Original/SGFO, Sonneveld Group B.V., Papendrecht, the Netherlands) was used as the base material for the production of both the gluten-free and gluten-containing breads. The gluten-free oat bread was baked under gluten-free conditions and confirmed to be gluten-free by the R5 Ridascreen Gliadin test. Vital wheat gluten (Kröner Stärke, Ibbenbüren, Germany; 8.6% of the total dough weight was added to the gluten-free oat-based bread mix to generate gluten-containing bread, amounting to 3.35 g of gluten per slice. The amount of gluten to add was determined on the basis of average daily gluten intake as described in previous studies.[19-21] The recipes were the same except for the addition of gluten, and both were similar in texture, taste, and appearance, as also confirmed a blind test in healthy volunteers. Both breads were baked for this study by the European

Bakery Innovation Centre, Papendrecht, the Netherlands. Further details about the study breads can be found in the appendix (p5-6).

Participants were instructed to adhere to a gluten-free diet from 1 week prior to test day 1, through days 2 and 3 of follow-up. They were allowed to consume gluten-free bread on days 2 and 3 as this would not interfere with the intervention.

On the test day (day 1, see appendix p4), participants were asked to come to the study site in a fasted state at 8.00AM. The test day started with a baseline questionnaire (t = 0 h) before breakfast. The questionnaire consisted of a symptom diary with 100mm VAS to assess overall GI symptoms, individual GI symptoms, and extra-intestinal symptoms, the Bristol Stool Scale (BSS, only after bowel movement), and the Positive And Negative Affect Schedule (PANAS) to assess mood.

After completion of the baseline questionnaire, *e.g.* at 8.15AM unless finished sooner or later, participants were informed about the group they were assigned to (E+ or E-), and then received breakfast with two slices of bread (G+ or G-) with a gluten-free topping of their choice (margarine with one standardised portion of cheese, cooked ham, or jam, which was noted) for breakfast. Throughout the test day participants completed the same hourly questionnaires, starting directly after breakfast, *e.g.* at 8.30AM unless finished sooner, for 8 hours (t = 1-8 hours, see Supplementary Figure 1).

After t = 4 hours, participants received lunch (*e.g.* at 12.30PM) with the same expectancy information repeated and the same bread type (two slices) as they had consumed for breakfast. Participants were allowed to drink coffee, tea, or water (ad libitum, but quantity was noted) during the test day, but no other foods or drinks were allowed during the test day (*e.g.* until 4.30PM). Between measurements, participants were requested to remain in the research unit and were free to watch television, read, or work.

After t = 8 hours, participants could go home. The test day questionnaires were repeated on the evening of day 1 (the test day) and on the two consecutive days (t = 2 days and t = 3 days) before going to bed (between 8PM and 2AM). Participants also completed a food record to assess adherence to the GFD, and medication use over the three days of the study.

For females, test days (including follow-up) were not scheduled during menstruation.

Participants could leave the study at any time if they wished to do so, and the investigator could decide to remove a participant for urgent medical reasons.

Safety was evaluated by reporting (serious) adverse events, *i.e.* any undesirable experience occurring to a participant, whether or not considered related to the food intervention, as reported spontaneously by the participant or observed by the investigator during the study.

Outcomes

The primary outcome was the effect of expectancy related to gluten intake on the overall GI symptom score, measured on a 100mm VAS as part of the symptom diary. The primary outcome was assessed centrally, *i.e.* data for all study sites was combined, comparing mean VAS over t = 1-8 hours, corrected for the baseline value for each individual, among all four groups. Thereby, the effect of actual gluten intake on overall GI symptoms was also analysed. Further secondary outcomes included the effects of expectancy and actual gluten intake on individual GI symptoms (*i.e.* abdominal discomfort, abdominal pain, belching, bloating, constipation, diarrhoea, flatulence, fullness, nausea, and urge to empty bowel), extra-intestinal symptoms (*i.e.* confusion/foggy mind, headache, and tiredness), and changes in mood (PANAS) throughout the test day (again using mean VAS over t = 1-8 hours, corrected for the baseline value). Furthermore, average stool frequency and consistency (BSS), and the impact of psychological state/emotional well-being *i.e.* anxiety, depression and somatic symptoms, on symptoms were of interest. A substantial proportion of the participants did not defecate at baseline and/or during the test day, and the remainder mostly had a single defecation at varying time points. Therefore, insufficient data was available for a reliable analysis of the BSS, and these data were not analysed or reported.

Statistical analysis

Sample size was calculated using G*power version 3·1. The sample size calculation was based on increase in overall GI symptoms scores as reported by Biesiekierski and colleagues after gluten consumption in irritable bowel syndrome (IBS) patients.[14] We assumed a difference of 15mm based on clinical relevance, standard deviation of 12·8mm, power of 80%, and a Bonferroni-corrected alpha of 0·0083, correcting for six pairwise comparisons. Based on this calculation, 20 participants were required per group, resulting in 80 participants in total. We aimed to include 84 participants because of an estimated drop-out rate of 5%. Although this sample size provided sufficient power to examine the primary research question, initially we aimed to obtain this

sample size in each country (UK and NL), so that any country differences could be compared. Because of recruitment delays due to the COVID-19 pandemic, an interim analysis was done in July 2021. This analysis was not pre-specified in the study protocol as the COVID-19 pandemic was unforeseen. The interim analysis compared E+ (n=37, with n=20 UK and n=17 NL) and E- (n=38, with n=19 UK and n=17 NL) without de-blinding the gluten intervention. This showed symptom profiles were comparable between the countries. Based on this interim analysis, we decided to recruit until a combined sample size of 84 was reached, as obtained from the power calculation. Thereafter the data was pooled for final analyses.

Statistical analyses were conducted using IBM SPSS statistics version 26.0. Normality of data was evaluated using histograms and QQ-plots. Baseline characteristics were presented as mean with corresponding standard deviation (SD) or medians with interquartile ranges (25^{th} – 75^{th} percentile), and as frequencies with percentages for categorical variables.

We planned for an intention-to-treat analysis including all randomised participants. However, we did not foresee that one participant, after completing the screening visit and being randomised, would fail to understand the test day instructions, resulting in no data being available for this participant. Therefore, we had to exclude this person and performed a per-protocol analysis.

The primary and secondary outcomes between the four groups were analysed per-protocol using repeated measures analysis of covariance (RM ANCOVA) with intervention group as the between-subject factor, baseline (t = 0 hours) as a covariate, and time (1-8 hours) as the repeated measures factor. For the primary outcome, we first checked the expectancy effect on overall GI symptoms by assessing pairwise comparison E+G- vs E-G-, and thereafter also assessed the other pairwise comparisons. For the secondary outcomes, we first did an overall comparison of all four groups and only if that showed significant differences were post-hoc pairwise comparisons performed, with post-hoc Bonferroni correction applied as appropriate (per symptom the alpha was corrected for six pairwise comparisons). Only Bonferroni-corrected p-values are reported.

Similarly, three post-hoc sensitivity analyses were done separately for the morning (t = 1-4 hours), afternoon (t = 5-8 hours), and follow-up (t = 1-3 days) using RM ANCOVA with baseline (t = 0 hours) as a covariate and t = 1-4 hours, t = 5-8 hours, or t = 1-3 days as repeated measures, respectively.

Additionally, post-hoc sensitivity analyses were performed for the test day, morning, afternoon, and follow-up analyses, in which the following variables were added sequentially to each model as single covariates to assess their impact: study site, sex, age, BMI, education level (university educated or not), smoking behaviour (current smoker, former smoker, or never smoked), alcohol consumption, IBS according to Rome IV criteria, FD according to Rome IV criteria, GAD-7 score, PHQ-9 score, and PHQ-15 score.

Missing values for the primary outcome measures were imputed using the median of the repeated measures from that participant for that symptom. This is a straightforward imputation method, as the median is robust to non-normal data distributions and the overall central tendency of the variable is preserved, and was considered reliable as only three participants had single missing values out of nine time points (*i.e.* t = 0-8 hours). The follow-up measurements included three time points (t = 1-3 days) and had more missing data (15·7% of measurements). Therefore, insufficient information was available to impute missing values using the median. Instead, for the follow-up measurements multiple imputation (generating 20 imputed data sets, each subjected to 20 iterations, utilising Fully Conditional Specifications and the Predictive Mean Matching technique), was used. A two-sided p-value of less than 0·05 was considered statistically significant.

As we noted substantial variation in individual responses within the groups, we explored symptom patterns post-hoc using an explorative random forest (URF) analysis. The URF was performed with overall GI symptoms and all individual GI symptoms at time points 0-8 hours. Results were visualised using a principal coordinate analysis plot (PCoA) plot.

In the UK, all data collection and entry was monitored and checked by the principal investigator and coordinating investigator. Additionally, as part of local regulations in the Netherlands, the study (both in Maastricht and in Wageningen) was monitored by a clinical study monitor. They checked *e.g.* informed consent forms, data collection and entry, compliance to protocols, and reporting of (serious) adverse events.

The study was registered at ClinicalTrials.gov (NCT05779358).

Role of the funding source

Representatives from the funding sources were permitted to ask questions and provide suggestions to the academic research consortium team, but were not involved in final

decisions regarding the study design, data collection, data analysis, data interpretation, and writing of this paper.

Results

Between October 19, 2018 and February 14, 2022, 683 individuals received the full study information. Of these, 301 (44·7%) individuals were pre-screened by phone and thereafter, 165 (24·2%) completed full screening, with 49 (7·2%) also undergoing a blood test to exclude CD. Main reasons for ineligibility were that CD could not be ruled out (n=43, 6·3%); individuals linked their symptoms to bread, wheat or other food products rather than to gluten (n=42, 6·1%); comorbidities or medication use (n=24, 3·5%); high GI symptom scores despite following a GFD or gluten-restricted diet (n=20, 2·9%); and symptoms reported to occur later than 8 hours after gluten consumption (n=5). Furthermore, 25 (3·7%) eligible participants dropped-out prior to randomisation, mainly due to delays to test day booking due to COVID-19 restrictions. As described in the Methods section, initially it was planned to include 84 (12·3%) participants in both the United Kingdom and the Netherlands, but because of the recruitment delays resulting from the COVID-19 pandemic, recruitment was halted early and data was pooled. Finally, 84 participants were randomised as follows: 21 E+G+, 21 E+G-, 20 E-G+, and 22 E-G-. One of the participants in the E+G+ group had to be excluded due to failure to understand the test day instructions, leaving data from 83 participants available for per-protocol analysis. See Figure 1 for the complete trial profile.

Figure 1. Trial profile.

Most participants (n=71, 85·5%) were female. The median age was 27·0 [21·0-45·0] years old, and mean BMI 23·8±3·9 kg/m^2. The majority had a university education (n=50, 60·2%), never smoked (n=66, 79·5%) or had quitted smoking (n=11, 13·3%), and overall alcohol intake was modest. In total, 29 (34·9%) of the participants met the Rome IV criteria for IBS with IBS-D (n=14, 16·9%) being the most common subtype, and 19 (22·9%) fulfilled the Rome IV criteria for functional dyspepsia. For full details, see Table 1, and appendix p7. At the screening visit, participants reported bloating (n=72, 86·7%) and abdominal pain (n=68, 81·9x%) as predominant symptoms after gluten exposure, see appendix p8.

Mean overall GI symptom score (Figure 2A), as measured on the VAS throughout the test day (*i.e.* t = 1-8 hours), corrected for baseline (t = 0 hours) in the RM ANCOVA, was not significantly different between E+G- (VAS 11·7mm [95% CI 8·3-15·1mm]), and E-G- (VAS 7·4mm [95% CI 4·2-10·7mm], p=0·47). The score in the E+G+ group (16·6mm [95% CI 13·1-20·0mm]) was significantly higher than E-G+ (VAS 6·9mm [95% CI 3·5-10·4mm], p=0·0010) and E-G- (p=0·0016), but not E+G- (p=0·28). Also no significant differences in mean estimated VAS were found between E-G+ and E-G- (VAS difference -0·5mm [95% CI -7·0-5·9mm], p>0·99).

When analysed separately in a post-hoc sensitivity analysis, mean estimated VAS differences between groups were more pronounced in the afternoon (t = 5-8 hours; E+G+ versus E-G+: VAS difference 11·9mm [95% CI 3·7-20·1mm], p=0·011, and E+G+ versus E-G-: VAS difference 11·7mm [95% CI 3·7-19·8mm], both p=0·0010) than in the morning (t = 1-4 hours; E+G+ versus E-G+: VAS difference 7·4mm [95% CI 0·4-14·3mm], p=0·031, E+G+ versus E-G-: VAS difference 6·5mm [95% CI -0·3-13·3mm], p=0·068). There was no significant effect of gluten within each expectancy group (*i.e.* E+G+ vs. E+G-, and E-G+ vs. E-G-). The other pairwise comparisons showed no significant differences between groups (appendix p9-12). These observed differences in overall GI symptom score for the test day, morning, and afternoon were still significant after correction for covariates (see appendix p17).

Figure 2. Test day scores for [A-D] gastrointestinal and [E-F] extra-intestinal symptoms, assessed by visual analogue scale (0-100mm), with significant differences between groups, and [G-H] test day scores for positive and negative affect scores. Participants consumed breakfast directly after t = 0 hours, and lunch directly after t = 4 hours. Differences between groups were analysed using repeated measures analysis of covariance (RM ANCOVA) with intervention group as the between-subject factor, baseline (t = 0 hours) as a covariate, and time (t = 1-8 hours) as the repeated measures factor. Post-hoc sensitivity analyses were done for the morning (t = 1-4 hours) and the afternoon (t = 5-8 hours). E+ = expectancy of getting gluten-containing bread; E- = expectancy of getting gluten-free bread; G+ = actual gluten-containing bread; G- = actual gluten-free bread; * p<0·05; ** p<0·01; *** p<0·001; comparisons with p>0.05 are not marked in the figure.

Compared to the test day, missing values were slightly more frequent for the follow-up measurements (t = 1 day (n=2), t = 2 days (n=6), and t = 3 days (n=8)) and therefore imputed by multiple imputations (n=20) using the repeated measures available for that participant. Observed differences between groups for overall GI symptoms persisted throughout the follow-up measurements (appendix p14-15), except for E+G+ vs. E-G- after corrections for covariates. For full details on the effect of the covariates, see appendix p17.

Evaluation of individual GI symptoms showed that mean estimated abdominal discomfort (Figure 2B) was significantly higher throughout the test day (t = 1-8 hours) in the E+G+ group (VAS 19·1mm [95% CI 14·5-23·7mm]) compared to E-G+ (VAS 6·7mm [95% CI 2·1-11·4mm], p=0·002) and E-G- (VAS 8·6mm [95% CI 4·2-13·0mm], p=0·010) (see appendix p9-10), again with differences more pronounced in the afternoon (t = 5-8 hours, appendix p12) than in the morning (t = 1-4 hours, appendix p11). Mean estimated bloating (Figure 2C) was significantly higher throughout the test day in E+G+ (VAS 14·4mm [95% CI 10·3-18·5mm]) compared to E-G+ (VAS 4·7mm [95% CI 0·6-8·8mm], p=0·008) (see appendix p9-10), but when analysed separately only significant in the afternoon (appendix p12). Within each expectancy group, gluten had no significant effect on abdominal discomfort and bloating (appendix p9-12). Observed differences for these symptoms were still significant during follow-up (appendix p14-15) and after inclusion of covariates.

Mean fullness (Figure 2D and appendix p9-12) was significantly higher in E+G+ compared to E-G+ and E-G- in the afternoon only. However, the difference between E+G+ and E-G+ was no longer significant after adding covariates (appendix p18), nor during follow-up (appendix p14 and p16).

The other GI symptoms, *i.e.* abdominal pain, belching, constipation, diarrhoea, flatulence, nausea, and urge to empty bowel, did not differ significantly between the groups, apart for abdominal pain between E+G+ and E-G+ during follow-up (see appendix p9-13 (test day), and appendix p14-16 (follow-up)). Covariates were also checked for symptoms that were initially not significantly different between groups, which resulted in some of the symptoms occasionally being affected (see appendix p17-18).

For the extra-intestinal symptoms that were assessed, mean confusion/foggy mind (Figure 2E and appendix p9-12) was significantly higher in the E+G+ group compared to E-G+ throughout the test day, and remained so after inclusion of covariates except smoking and alcohol intake (appendix p18). Headache (Figure 2F and appendix p9-12) was significantly higher in E+G+ compared to E-G+. After correction for sex, BMI or GAD-7 score, headache was also significantly higher in E+G+ compared to E+G- (appendix p18). When analysed separately, these differences between groups for both confusion/foggy mind and headache were only significant in the morning. The differences also did not persist through follow-up (appendix p14 and p16), nor after corrections for covariates (appendix p18). Tiredness was not significantly different between groups (appendix p9-14 and p16).

Overall participants scored low on the screening questionnaires for anxiety (2·0 [0·0-4·3]), depression (2·0 [0·0-4·0]) and somatisation (6·2±3·6), with only few participants meeting the cut-off point of ≥10 (see Table 1). When added as covariates to the RM ANCOVA model, these psychological factors affected differences between intervention groups for headache and tiredness during the test day, for fullness and tiredness in the afternoon, and for overall GI symptoms and abdominal pain during follow-up. See appendix p17-19 for full details. Furthermore, throughout the test day and follow-up, both positive affect and negative affect, scored using the PANAS, did not differ significantly between the four groups (see Figures 2G-2H and appendix p9-12, p14 and p16).

To further explore heterogeneity in symptom response, an URF analysis was done. This identified some patterns in the data, as illustrated by the score plot shown in Figure 3A based on PCo5 and Pco7, which showed partial data separation by intervention groups. As can be seen within the specific intervention groups, a group of individuals (see appendix p20) showed clear separation with respect to the measured symptoms, especially in groups E+G+ and E+G- in comparison to groups E-G- and E-G+. All symptoms play a role in this separation, with diarrhoea and constipation having the lowest importance (Figure 3B). Taking into account individuals with the highest responsiveness with respect to symptoms, *i.e.* groups E+G+ and E+G-, overall GI symptoms, abdominal pain, abdominal discomfort, urge to empty bowel and fullness

could be defined as top symptoms. The observed data separation could not be explained by other demographic and clinical parameters, such as IBS (appendix p20).

Figure 3. [A] Principal coordinate analysis (PcoA) score plot based on unsupervised random forest (URF) analysis and [B] relative contribution in URF model of overall gastrointestinal (GI) symptoms (GISymp) and all individual GI symptoms (*i.e.* abdominal discomfort (AbDis), abdominal pain (AbPain), belching (Belch), bloating (Bloat), constipation (Const), diarrhoea (Diarr), flatulence (Flat), fullness (Full), nausea (Naus), and urge to empty bowel (Urge)) at t = 0-8 hours. This plot is colour-coded with respect to the intervention type. E+ = expectancy of getting gluten-containing bread; E- = expectancy of getting gluten-free bread; G+ = actual gluten-containing bread.

Three (3·6%) of 83 participants reported adverse events on the test day. In the E+G- group, one (4·7%) of 21 participant reported an itching sensation in their jaw between t = 0-1 hours, and one (4·7%) of 21 participant reported a lightheaded feeling and rumbling stomach between t = 7-8 hours. In the E-G+ group, one (5·0%) of 20 participant vomited twice between t = 6-8 hours.

Discussion

This randomised, double-blind, placebo-controlled, international multicentre study was, to our knowledge, the first study designed to investigate the role of the nocebo effect in NCGS. Our findings showed that the combined effect of expectancy and actual gluten intake had the largest effect on overall GI symptoms. Repeated exposure compounded this effect, evidenced by the more pronounced effect in the afternoon (after lunch) compared to the morning (after breakfast). Similar patterns were found for predominant GI symptoms, abdominal discomfort and bloating. Furthermore, expectancy had a significant effect on the extra-intestinal symptoms confusion/foggy mind and headache. Most differences between intervention groups persisted throughout follow-up. These findings confirm our hypothesis that a nocebo effect is involved in symptom occurrence in NCGS. We found no significant differences based on actual gluten intake within each expectancy group, but our data also indicate that a concurrent biological effect of gluten cannot be excluded. Additionally, contrary to our hypothesis, we found that emotional well-being, *i.e.* anxiety, depression, or

somatisation, did not significantly affect differences between groups for overall and predominant symptom scores during the test day.

This study showed that the nocebo effect plays a significant role in symptom occurrence in NCGS. Hereby, we add to the findings from a study by Biesiekierski and colleagues, which also indicated that the expectancy to receive gluten played a greater role than the actual consumption of gluten, as demonstrated by the order effect they reported.[14] A more recent study by Ponzo and colleagues also found an order effect when comparing gluten to placebo in individuals with self-reported NCGS.[22] So far, previous studies have considered occurrence of the nocebo effect a limitation of the DBPC study design, rather than the nocebo effect itself being an important causal factor.[12,23,24] Expectancy, typically induced via verbal suggestions, and learning are the two best-characterised mechanisms that mediate the nocebo effect. These processes are mediated centrally, involving multiple brain regions and influencing gastrointestinal sensory and motor functions along the bidirectional gut-brain axis.[16] The gut-brain axis is a bidirectional interaction between the GI tract and the central nervous system including the brain and spinal cord. It involves multiple pathways, such as the autonomic and enteric nervous system, the endocrine system, the hypothalamic-pituitary-adrenal axis, the immune system, and the gut microbiota and its metabolites.[25] The nocebo effect is also an important feature in patients with IBS, in whom the gut-brain interaction plays a clear role.[17] We consider the role of the nocebo effect in NCGS symptom occurrence as a new lead for possible involvement of interaction between the gut and the brain that warrants further study.

This is further supported by the substantial overlap between NCGS and IBS, currently characterised as a DGBI.[9] In our study population, 34·9% met the Rome IV criteria for IBS. This is higher than in the general population and comparable to the prevalence reported by previous studies, which ranged from 20-44%.[24] IBS-D was most prevalent in our study, but numbers were too small for further analyses on subtypes. Furthermore, in the present study, the number of IBS patients was comparable between intervention groups, and symptom response was not different between those with or without IBS.

Remarkably, we found no significant effect of actual gluten intake within each of the expectancy groups. Nevertheless, the combined effect of expectancy and gluten had the largest effect, pointing to an additive or synergistic effect of gluten exposure. Previous studies have shown conflicting evidence for the role of gluten in NCGS.[24]

Although several studies found a gluten challenge induced higher symptom scores compared to placebo,[26-28] others reported no significant effects,[29-31] or improvement of symptoms scores on a GFD, [32] or even a higher response after placebo.[33] Furthermore, several studies indicate that other wheat components, including FODMAPs (fermentable oligosaccharides, disaccharides, monosaccharides and polyols) such as fructans and amylase trypsin inhibitors (ATIs), may be more important triggers than gluten.[14,24,34-36] It is important to establish whether a strict GFD is needed to manage symptoms. On one hand, following a strict GFD without adequate guidance and replacement may lead to unbalanced dietary intake and nutrient deficiency.[11] On the other hand a strict GFD may not be necessary in the absence of CD.[12] Either way, in clinical practice it remains important that these NCGS individuals receive adequate dietetic guidance, to identify and replace potential trigger foods while maintaining a balanced diet.

The demand for an individualised dietary approach for NCGS individuals was further supported by our explorative URF analysis. We were able to identify responders and non-responders within each intervention group, but could not fully explain the variation in symptom response by predominant symptoms or patient characteristics such as IBS. Thus, these results confirm that symptom occurrence in NCGS is heterogeneous and cannot be explained by one clear mechanism. Therefore, further research should also focus on determining the biological mechanisms by which gluten and other wheat components can lead to GI symptoms in NCGS, the cause for inter-individual differences in symptom responses, and on investigating the need for a strict GFD.

In line with some previous studies,[22,26,27,37] we found that expectancy had a significant effect on the extra-intestinal symptoms, confusion/foggy mind and headache, during the test day. However, contrary to our hypothesis, anxiety, depression, and somatic symptoms did not significantly affect observed differences between groups for overall and predominant symptom scores during the test day. Furthermore, they had only a minor effect on differences between intervention groups for extra-intestinal symptoms during the test day, and for GI symptoms during follow-up. Mood was also not significantly affected by the intervention. Although previous studies found a higher prevalence of psychological comorbidities in NCGS patients and that psychological well-being was affected by gluten intake,[13,29,35] our study did not confirm these findings. This may be due to a selection bias, because it is plausible that more anxious or

symptomatic patients were less willing to participate. The impact of psychological factors should be considered in future studies.

The main strength of our study was that it was the first well-designed study to investigate the role of the nocebo effect in NCGS, using a physiologically relevant dose of gluten administered in a clinically controlled environment. The breads used in this study differed only in gluten content, and had equal levels of fibres, including FODMAPs. Strict inclusion criteria were used and CD and WA were ruled out in our study population, although the latter was based on medical history only.

Another strength was the hourly measurements during the 8 hour test day, with a repeated exposure to expectancy and actual gluten intake. Subsequently, we noted a higher response rate in the afternoon as compared to the morning. Although the time-course of gluten-evoked symptoms may be a plausible explanation in some individuals, our figures for overall GI symptoms and predominant GI symptoms abdominal discomfort and bloating show a first peak in symptom score after 1-2 hours, which decreases before lunch, and again increases after lunch. Therefore, we hypothesise this was mainly due to the repeated exposure to the same condition.

However, it should be noted that overall GI symptom scores, although significantly different between groups, were rather low. We cannot exclude selection bias, as those with high symptoms or more anxious individuals may be less willing to participate. A limitation of the present study is that stress was not measured, although this is known to affect GI symptoms. Furthermore, due to delays in recruitment due to COVID-19, the study was terminated early, resulting in pooling of British and Dutch data. However, as the same pattern of effects was found in each country this was not an issue. Although our analyses would have had more power with twice as many participants, lending more confidence to the generalisability of our results between the countries, the effects are clear and consistent. As our effect sizes are similar to previous studies,[14,28,31,34] we consider generalisability among Western (European) countries to be adequate. Furthermore, baseline differences did not differ between the countries. Although the majority of our study population was female, this is in line with other studies,[14,22,26-37] and indicates that being female can be considered a population characteristic or risk factor for NCGS.

Based on these findings, future research efforts should aim to identify biomarkers which distinguish heterogeneous symptom patterns of NCGS. Furthermore, the role of the gut-brain axis and psychological factors should be investigated, alongside the

potential pathophysiological effects of gluten and other wheat components. For clinical management, both adequate dietary guidance, including proper identification of trigger foods and adequate replacement of these products guided by a dietitian, as well as addressing potential psychological or behavioural factors should be considered.

To conclude, we found that the combined effect of expectancy and actual gluten intake had the largest effect on overall and individual GI symptoms, reflecting a significant nocebo effect, although an additional effect of gluten could not be ruled out. This combined effect was accentuated by the repeated exposure following a lunch bolus. The results of this study support the importance of further research into the possible involvement of gut-brain interaction in NCGS.

Contributors

Conceptualization: FJPHB, LD, DMAEJ, DK, PRS, GvR; Data curation: MCGdG, FC, AS, MAMH, GvR; Formal analysis: MCGdG, AS, LD, BW, DMAEJ; Funding acquisition: FJPHB, PRS; Investigation: MCGdG, FC, GvR; Methodology: FJPHB, LD, DMAEJ, PLW, GvR, LAH, MCGdG; Project administration: MCGdG, FC, CLL, GvR, FJPHB, DMAEJ, LD, BJMW; Resources: FJPHB, LD, DMAEJ, PLW, BJMW; Supervision: FJPHB, LD, DMAEJ, CLL; Validation: MCGdG, LD, BW; Visualisation: MCGdG, AS; Writing – original draft: MCGdG; Writing – review & editing: DMAEJ, LD, FJPHB, CLL, MAMH, GvR, PLW, LAH, BJMW, DK.

All authors had full access to all the data in the study and had final responsibility of the decision to submit for publication. MCGdG, CLL, FC, MAMH, and LD have accessed and verified the data.

Declarations of interest

All authors have declared their conflicts of interest according to the rules of the International Committee of Medical Journal Editors (ICMJE). MCGdG was partially financed by Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr. TKI16060BMNU. CLL was partially financed by Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr. TKI16060BMNU. FC was partially financed by Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr. TKI16060BMNU. AS received consulting fees and financial support for conference attendance by Owlstone Medical

(all outside submitted work). BW has nothing to declare. MAMH was partially financed by Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr. TKI16060BMNU. GvR was partially financed by Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr. TKI16060BMNU. PLW is in part employed by European Bakery Innovation Centre, Sonneveld Group BV, who provided the study bread. Sonneveld Group BV received payment from Public-private partnership grant of TKI (“Well on Wheat?”, grant nr. TKI16060BMNU) for providing the study breads. PRS received Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr. TKI16060BMNU, AHDB Studentship Project Ref 061, FIBRAXFUN; the Biotechnology and Biological Sciences Research Council (BBSRC) grant nr BB/S013954/1, BBSRC BB/S020101/1, BBSRC BB/T013923/1 and BBSRC BB/W01792X/1, and EPSRC EP/S01679/1 GCRF Hub SANC South Asian Nitrogen Hub; received honorarium for membership of Editorial Board of Journal of Cereal Science (Elsevier); received payment of flight, hotel costs and/or registration by conferences; is a member of DAFM (Ireland) grant board (paid to self), and works at Rothamsted Research. Rothamsted Research receives strategic funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and the work forms part of the Designing Future Wheat (BB/P016855/1) and Developing Sustainable Wheat (BB/X011003/1) strategic programmes. LAH acts as consultant to GSK, Ironwood and non-paid to Pfizer, received financial support for attendance of the Rome V foundation meeting in Chicago May 2023, and acts as member of the British Society of Gastroenterology (BSG) Neurogastroenterology and Motility Committee and BSG Food and Function Committee (all outside submitted work). BJMW has nothing to declare. DK has received research funding from Grunenthal, Allergan, ZonMw, MLDS, Rome Foundation, Horizon 2020, Horizon Europe, and UEG, has received speaker’s fee (paid to host institution) from Dr Falk, participates on Data Safety Monitoring/Advisory Board for the NUTRECOVER study (no financial reimbursement), board member of the Dutch Society of Gastroenterology (no financial reimbursement) (all outside submitted work). FJPHB received Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr. TKI16060BMNU, consulting fees from Lesaffre France, personal speaker fee and travel cost reimbursement from Zeelandia, Netherlands, and payment to institution for travel or hotel costs for invited scientific meetings. LD was partially financed by Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr.

TKI16060BMNU. Additionally, outside the submitted paper, she is a past President of ILSI Europe and current Co-Chair of ILSI Global, has received grant funding from Sanofi CHC, and sits on the Hass Avocado Scientific Advisory Board. DMAEJ received Public-Private partnership grant of Top Knowledge Institute (TKI) “Well on Wheat?” grant nr. TKI16060BMNU. Additionally, part of her work, outside the submitted paper, has been financed by other the public-private partnership grants of Top Knowledge Institute (TKI) Agri&Food and Health Holland, by the Carbokinetics program as part of the NWO-CCC Partnership Program, by Organic A2BV/Mothersfinest BV and by H2020 DISCOvERIE/848228.

Data sharing

On request, de-identified participant data can be available to researchers who provide a methodologically sound proposal, in line with aims in the approved proposal and in line with EU regulations. Data will be available immediately following publication, within 15 years after publication. Proposals and requests should be directed to d.jonkers@maastrichtuniversity.nl to gain access, data requestors will need to sign a data transfer access agreement.

Acknowledgements

This trial was part of the “Well on Wheat?” project, funded by a Public-Private Partnership Grant from the Top Knowledge Institute, grant nr. TKI16060BMNU and a wide range of additional partners* from the Agri-Food industry who donated unrestricted research grants. **These partners were: AB-Mauri Bakery Ingredients, Made, the Netherlands; Baking Industry Research Trust, Wellington, New Zealand; Borgesius Holding BV-Albert Heijn, Stadskanaal, the Netherlands; CSM Innovation Bakery Center, Bingen, Germany; CYMMIT, Texcoco, Mexico; DSM Food Specialties, Delft, the Netherlands; Fazer Bakeries Oy, Helsinki, Finland; Health Grain Forum, Vienna, Austria; ICC-Intl., Vienna, Austria; IWGA, Kansas 66210, USA; Lantmännen EK, Stockholm, Sweden; Mondelez, Saclay, France; Nederlands Bakkerij Centrum, Wageningen, the Netherlands; Nutrition et Sante, Revel, France; Puratos BV, Groot Bijgaarden, Belgium; Rademaker BV-Bakery equipment, Culemborg, the Netherlands; Sonneveld Group BV, Papendrecht, the Netherlands; Zeelandia-HJ Doeleman BV, Zierikzee, the Netherlands.* The project was coordinated and executed by an academic

research consortium team (ARCT). Research specialists from funding partners were entitled to ask questions and give suggestions to ARCT for consideration during biannual progress meetings. Decisions on study set-up, execution and data interpretation of data were exclusively taken by ARCT. Scientific output communications were exclusively organized by ARCT without involvement of funding partners.

None of the authors are employed by NIH or are in receipt of an NIH grant.

We thank all volunteers who participated in this study, and the staff and students at all participating sites (University of Leeds: Neil Boyle, Alice Kininmonth, Baptiste Goubel, Diana Goldberg, Cassy Dingena, Anastasia Aivatoglou, Isabella Ockleton, Amy Murgatroyd, Rebecca Spencer, Cheryl Low, Mikaela Gerber, Daniel Slatter; Maastricht University: Ellen Wilms, Yala Stevens, Ashkan Rezazadeh Ardabili, Wendy Brouwer-van den Oever, Greta Lennertz; and Wageningen University: Henriëtte Fick-Brinkhof) for their cooperation.

References

1. Poutanen KS, Kårlund AO, Gómez-Gallego C, et al. Grains – a major source of sustainable protein for health. *Nutr Rev* 2022; **80**(6): 1648-63.
2. Gaesser GA. Whole Grains, Refined Grains, and Cancer Risk: A Systematic Review of Meta-Analyses of Observational Studies. *Nutrients* 2020; **12**(12).
3. Hu H, Zhao Y, Feng Y, et al. Consumption of whole grains and refined grains and associated risk of cardiovascular disease events and all-cause mortality: a systematic review and dose-response meta-analysis of prospective cohort studies. *Am J Clin Nutr* 2023; **117**(1): 149-59.
4. Jayedi A, Ge L, Johnston BC, et al. Comparative effectiveness of single foods and food groups on body weight: a systematic review and network meta-analysis of 152 randomized controlled trials. *Eur J Nutr* 2023; **62**(3): 1153-64.
5. Sanders LM, Zhu Y, Wilcox ML, Koecher K, Maki KC. Whole grain intake, compared to refined grain, improves postprandial glycemia and insulinemia: a systematic review and meta-analysis of randomized controlled trials. *Crit Rev Food Sci Nutr* 2023; **63**(21): 5339-57.
6. Catassi C, Elli L, Bonaz B, et al. Diagnosis of Non-Celiac Gluten Sensitivity (NCGS): The Salerno Experts' Criteria. *Nutrients* 2015; **7**(6): 4966-77.
7. Sabença C, Ribeiro M, Sousa T, Poeta P, Bagulho AS, Igrejas G. Wheat/Gluten-Related Disorders and Gluten-Free Diet Misconceptions: A Review. *Foods* 2021; **10**(8).
8. Mansueto P, D'Alcamo A, Seidita A, Carroccio A. Food allergy in irritable bowel syndrome: The case of non-celiac wheat sensitivity. *World J Gastroenterol* 2015; **21**(23): 7089-109.
9. Oka P, Parr H, Barberio B, Black CJ, Savarino EV, Ford AC. Global prevalence of irritable bowel syndrome according to Rome III or IV criteria: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* 2020; **5**(10): 908-17.

10. Jansson-Knodell CL, Rubio-Tapia A. The fashionable gluten-free diet—wear with caution. *Am J Clin Nutr* 2021; **113**(3): 491-2.
11. Kreutz JM, Heynen L, Vreugdenhil ACE. Nutrient deficiencies in children with celiac disease during long term follow-up. *Clin Nutr* 2023; **42**(7): 1175-80.
12. Khan A, Suarez MG, Murray JA. Nonceliac Gluten and Wheat Sensitivity. *Clin Gastroenterol Hepatol* 2019.
13. Croall ID, Trott N, Rej A, et al. A Population Survey of Dietary Attitudes towards Gluten. *Nutrients* 2019; **11**(6).
14. Biesiekierski JR, Peters SL, Newnham ED, Rosella O, Muir JG, Gibson PR. No effects of gluten in patients with self-reported non-celiac gluten sensitivity after dietary reduction of fermentable, poorly absorbed, short-chain carbohydrates. *Gastroenterology* 2013; **145**(2): 320-8 e1-3.
15. Molina-Infante J, Carroccio A. Suspected Nonceliac Gluten Sensitivity Confirmed in Few Patients After Gluten Challenge in Double-Blind, Placebo-Controlled Trials. *Clin Gastroenterol Hepatol* 2017; **15**(3): 339-48.
16. Elsenbruch S, Enck P. Placebo effects and their determinants in gastrointestinal disorders. *Nat Rev Gastroenterol Hepatol* 2015; **12**(8): 472-85.
17. Li R, Chen F, He X, et al. Nocebo response intensity and influencing factors in the randomized clinical trials of irritable bowel syndrome: A systematic review and meta-analysis. *Front Med (Lausanne)* 2022; **9**: 1018713.
18. Biagi F, Andrealli A, Bianchi PI, Marchese A, Klersy C, Corazza GR. A gluten-free diet score to evaluate dietary compliance in patients with coeliac disease. *Br J Nutr* 2009; **102**(6): 882-7.
19. Biesiekierski JR. What is gluten? *J Gastroenterol Hepatol* 2017; **32**: 78-81.
20. van Overbeek FM, Uil-Dieterman IGA, Mol IW, Köhler-Brands L, Heymans HSA, Mulder CJJ. The daily gluten intake in relatives of patients with coeliac disease compared with that of the general Dutch population. *Eur J Gastroenterol Hepatol* 1997; **9**(11): 1097-9.
21. Rijksinstituut voor Volksgezondheid en Milieu. Brood, granen, rijst en pasta. https://www.wateetnederland.nl/resultaten/voedingsmiddelen/brood-en-granen (accessed 7 August 2023).
22. Ponzo V, Ferrocino I, Goitre I, et al. Non-Celiac Gluten/Wheat Sensitivity: Clinical Characteristics and Microbiota and Mycobiota Composition by Response to the Gluten Challenge Test. *Nutrients* 2021; **13**(4).
23. Borrelli DEAF, Schiepatti A, Gibiino G, Fabbri C, Baiardi P, Biagi F. Is it time to rethink the burden of non-coeliac gluten sensitivity? A systematic review. *Minerva Gastroenterol* 2021.
24. Barbaro MR, Cremon C, Wrona D, et al. Non-Celiac Gluten Sensitivity in the Context of Functional Gastrointestinal Disorders. *Nutrients* 2020; **12**(12).
25. Vanner S, Greenwood-Van Meerveld B, Mawe G, et al. Fundamentals of Neurogastroenterology: Basic Science. *Gastroenterology* 2016.
26. Biesiekierski JR, Newnham ED, Irving PM, et al. Gluten causes gastrointestinal symptoms in subjects without celiac disease: a double-blind randomized placebo-controlled trial. *Am J Gastroenterol* 2011; **106**(3): 508-14; quiz 15.
27. Di Sabatino A, Volta U, Salvatore C, et al. Small Amounts of Gluten in Subjects With Suspected Nonceliac Gluten Sensitivity: A Randomized, Double-Blind, Placebo-Controlled, Cross-Over Trial. *Clin Gastroenterol Hepatol* 2015; **13**(9): 1604-12 e3.
28. Elli L, Tomba C, Branchi F, et al. Evidence for the Presence of Non-Celiac Gluten Sensitivity in Patients with Functional Gastrointestinal Symptoms: Results from

a Multicenter Randomized Double-Blind Placebo-Controlled Gluten Challenge. *Nutrients* 2016; **8**(2): 84.
29. Peters SL, Biesiekierski JR, Yelland GW, Muir JG, Gibson PR. Randomised clinical trial: gluten may cause depression in subjects with non-coeliac gluten sensitivity - an exploratory clinical study. *Aliment Pharmacol Ther* 2014; **39**(10): 1104-12.
30. Moleski SM, Shah A, Durney P, et al. Symptoms of gluten ingestion in patients with non-celiac gluten sensitivity: A randomized clinical trial. *Nutrition* 2021; **81**: 110944.
31. Crawley C, Savino N, Halby C, et al. The effect of gluten in adolescents and young adults with gastrointestinal symptoms: a blinded randomised cross-over trial. *Aliment Pharmacol Ther* 2022; **55**(9): 1116-27.
32. Algera JP, Magnusson MK, Öhman L, Störsrud S, Simrén M, Törnblom H. Randomised controlled trial: effects of gluten-free diet on symptoms and the gut microenvironment in irritable bowel syndrome. *Aliment Pharmacol Ther* 2022; **56**(9): 1318-27.
33. Dale HF, Hatlebakk JG, Hovdenak N, Ystad SO, Lied GA. The effect of a controlled gluten challenge in a group of patients with suspected non-coeliac gluten sensitivity: A randomized, double-blind placebo-controlled challenge. *Neurogastroenterol Motil* 2018.
34. Skodje GI, Sarna VK, Minelle IH, et al. Fructan, Rather Than Gluten, Induces Symptoms in Patients With Self-Reported Non-Celiac Gluten Sensitivity. *Gastroenterology* 2018; **154**(3): 529-39 e2.
35. Dieterich W, Schuppan D, Schink M, et al. Influence of low FODMAP and gluten-free diets on disease activity and intestinal microbiota in patients with non-celiac gluten sensitivity. *Clin Nutr* 2019; **38**(2): 697-707.
36. Nordin E, Brunius C, Landberg R, Hellström PM. Fermentable oligo-, di-, monosaccharides, and polyols (FODMAPs), but not gluten, elicit modest symptoms of irritable bowel syndrome: a double-blind, placebo-controlled, randomized three-way crossover trial. *Am J Clin Nutr* 2022; **115**(2): 344-52.
37. Cartee AK, Choung RS, King KS, et al. Plasma IL-2 and Symptoms Response after Acute Gluten Exposure in Subjects With Celiac Disease or Nonceliac Gluten Sensitivity. *Am J Gastroenterol* 2022; **117**(2): 319-26.