

Safety and Acceptance of a Relaxed Gluten-Free Diet in Celiac Disease: Protocol for a Mixed Methods Pilot Randomized Controlled Trial

Diana Studerus

diana.studerus@diets-hin.ch

GastroZentrum Hirslanden Zürich

Jonas Zeitz

GastroZentrum Hirslanden Zürich

Vanessa Dürr

GastroZentrum Hirslanden Zürich

Nadine Bosbach

Clinical Trial Unit Hirslanden Clinic

Ewerton Marques Maggio

Medica Pathology, Enge

Lauren Clack

University of Zurich

Michael Scharl

University Hospital Zurich

Study protocol

Keywords: celiac disease, gluten-free diet, cross-contact, mixed methods, pilot randomized controlled trial, de-implementation, practice theory

Posted Date: October 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10874600/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

The threshold below which gluten is harmless for an individual with celiac disease has not been established. In its absence, guidelines recommend complete avoidance of gluten from all sources, including cross-contact with gluten, which in everyday life translates into continuous label-checking, control over food preparation, and heightened vigilance in social settings. This vigilance carries a substantial burden for dietary quality of life and can foster rigidity, avoidance, and food preoccupation, features recognized as risk factors for disordered eating. Existing threshold evidence derives from fixed-dose gluten capsule challenges, acute biomarker experiments, and retrospective syntheses, none of which corresponds to distributed, chronic, food-borne exposure with gluten coming from cross-contact evaluated against a histological endpoint. Whether the strictness currently demanded is proportionate to the physiological risk of low-level exposure therefore remains unanswered.

Methods

This single-center, single-blind, parallel-group pilot randomized controlled trial is embedded within a convergent parallel mixed methods design in which the qualitative strand carries equal methodological weight. Thirty adults with celiac disease in clinical remission are randomized 1:1 to a relaxed gluten-free diet permitting cross-contact up to 200 ppm, or to an optimized strict gluten-free diet, for 24 weeks. The primary endpoint is the villus height to crypt depth ratio from duodenal biopsies at baseline and week 24, assessed by a blinded histopathologist, with a margin of 0.4 units for a clinically meaningful deterioration. Secondary endpoints comprise serum tissue transglutaminase IgA, fecal gluten immunogenic peptides, symptom and acceptability measures, and estimated gluten exposure. Semi-structured interviews at five timepoints are analyzed by inductive–deductive content analysis. Integration proceeds through connected sampling, joint displays at weeks 12, 18, and 24, and meta-inference. Analyses are estimation-focused; an individual safety switch is triggered by serum tissue transglutaminase IgA above 10 U/mL (reference: < 7 U/mL).

Discussion

The trial provides the first prospective histological evaluation of deliberately relaxed gluten avoidance from cross-contact. It is designed to inform a definitive multicenter trial.

Trial registration

German Clinical Trials Register, DRKS00039577. Registered 3rd of June 2026

Background

Celiac disease (CeD) is a chronic, immune-mediated small-intestinal enteropathy precipitated by dietary gluten in genetically predisposed individuals (Ludvigsson et al., 2013), affecting 1–2% of the global population (Singh et al., 2018). Its only established treatment is lifelong adherence to a gluten-free diet (GFD) (Lebwohl et al., 2018). The threshold below which gluten becomes harmless for the individual patient with CeD has not been established. In the absence of such a threshold, current guidelines recommend the complete avoidance of gluten from all sources (often referred to a „strict“ GFD), including food, beverages, medications, and supplements, as the standard of care (Ludvigsson et al., 2014; Rubio-Tapia et al., 2023) a rationale articulated by Wolf et al. (2018). In everyday life this translates into a demand for maximal strictness: people with CeD must also manage cross-contact with gluten, which refers to the *«unintentional transfer of gluten proteins from gluten-containing sources into inherently gluten-free (GF) foods, resulting in gluten content above legally established thresholds»* (Studerus et al., 2025). Cross-contact with gluten can occur at multiple stages, from agricultural harvesting and industrial processing through shared manufacturing equipment to food preparation in restaurants, bakeries, and home kitchens (Vargas et al., 2024). The complexity is further compounded by inconsistent precautionary allergen labelling practices such as "may contain traces of gluten", leaving people to navigate uncertainty with limited evidence-based guidance (Rasmussen et al., 2022): following a strict GFD requires continuous label-checking, control over food preparation, and heightened vigilance in social eating settings (Gholmie et al., 2023; Satherley et al., 2022; Simon et al., 2023). Beyond this burden, the same vigilance can foster features that are themselves recognized risk factors for disordered eating, namely rigidity, avoidance, controlling behavior around food, and food preoccupation; both extreme dietary adherence and inadvertent transgressions have been associated with disordered eating patterns (Cadenhead et al., 2019; Gholmie et al., 2023; Satherley et al., 2017). Whether the „strictness“ currently demanded is proportionate to the actual risk of the assumed low-level gluten exposure coming from cross-contact is therefore both a clinical and a methodological question.

The amount of gluten that can be safely tolerated remains contested, and the existing evidence does not so much resolve this question as fracture it across different endpoints. The most frequently cited threshold trial administered fixed daily doses of 0, 10, or 50 mg of gluten over three months and found that 50 mg/day induced mucosal injury in most participants, with intermediate and partly unsafe findings at 10 mg/day (Catassi et al., 2007). A subsequent dose–response meta-analysis, which incorporated this trial alongside further challenge and observational data, situated such doses within a graduated risk curve in which the relative risk (RR) of relapse remains close to baseline up to approximately 150 mg/day ($RR \approx 1.07$) and reaches a clinically relevant magnitude only near 880 mg/day ($RR \approx 1.50$) (Rostami-Nejad et al., 2023). The apparent tension between these findings is one of endpoint rather than of contradiction: a measurable histomorphometric shift under controlled challenge is not the same event as clinical or histological relapse, and the two need not coincide at the same dose. On this clinical and histological axis, the daily gluten exposure attributable to realistic cross-contact appears modest: drawing on the most comprehensive synthesis of cross-contact data available to date (Studerus et al., 2025) together with our own targeted gluten analysis (manuscripts in preparation), an exposure ceiling

defined by cross-contact of up to 200 parts per million (mg/kg) (ppm) corresponds, across typical food portions, to estimated daily intakes that lie well below the steep portion of this curve.

This apparently reassuring picture has, however, been complicated by evidence drawn from a different axis. A recent randomized, double-blind, placebo-controlled dose–response trial, using the serum interleukin-2 response as a sensitive biomarker of gluten-specific T-cell activation, found that acute immune activation occurs at gluten doses far below current food-labeling thresholds, with an estimated eliciting dose of 2.4 mg for 10% of patients and 0.8 mg for 5% of patients, and with patient-reported symptoms providing an unreliable indicator of exposure below 1000 mg (Daveson et al., 2026). Critically, the same authors emphasize that the long-term clinical significance of such low-dose immune activation is unknown, and that their acute design was not intended to evaluate cumulative effects such as sustained mucosal damage. The immune-activation threshold and the histological-injury threshold thus appear to lie at very different dose scales, and the relationship between them remains unestablished.

These divergent axes have recently acquired regulatory expression. In November 2025 an ad hoc joint FAO/WHO expert consultation recommended a reference dose of 4 mg gluten for the risk assessment of unintended gluten presence in foods, replacing the earlier reference dose of 5 mg wheat protein (Food & Agriculture Organization of the United Nations & World Health Organization, 2025). The corresponding risk-based framework for precautionary allergen labeling was adopted by the Codex Alimentarius Commission in July 2026; the 20 ppm limit for gluten-free claims remained unchanged (Codex Alimentarius, 2026). The consultation explicitly noted that a reference dose for CeD must be derived on a different basis than for IgE-mediated allergy, since the relevant risk arises from chronic rather than acute exposure, and that the data available for this purpose are limited. The normative direction of this development converges with the present trial: precautionary labeling is to follow documented risk assessment rather than blanket application, and the risk-relevant quantity is an absolute dose rather than a concentration. The reference dose postdates the development of the present intervention and could not inform its design, whose exposure ceiling was derived from empirical cross-contact data and dose–response evidence rather than from a protective regulatory threshold. The two address different questions : when a warning statement is warranted, and what level of exposure is in fact tolerated, and it is the second that remains empirically unresolved.

What the existing body of evidence has in common is that it was generated under exposure conditions that differ systematically from the one that actually determines whether cross-contact within a GFD is safe. Threshold estimates have been derived from fixed-dose capsule challenges (Catassi et al., 2007), from acute single boluses (Daveson et al., 2026), or from the retrospective synthesis of heterogeneously defined relapse outcomes (Rostami-Nejad et al., 2023). Each of these designs answers a legitimate question, and together they delineate the dose range within which the risk of low-level exposure is likely to be small. None of them, however, corresponds to gluten exposure as it arises in everyday life: distributed across meals, food-borne rather than encapsulated, sustained over months, and occurring under a diet in which cross-contact is deliberately accepted rather than accidentally incurred. Nor has

such an exposure been evaluated prospectively against a histological endpoint. It is this gap that the present study addresses.

Dietary interventions in chronic disease are increasingly recognized as complex interventions in the sense defined by the Medical Research Council framework: they involve multiple interacting components, operate through behavioral and psychosocial mechanisms, and are highly sensitive to context, including social environment, individual illness beliefs, and the material conditions of everyday food practice (Moore et al., 2015; Skivington et al., 2021). For people with CeD, the GFD exemplifies this complexity. Adherence is not simply a function of knowledge or motivation but is shaped by social norms, household composition, food availability, and the negotiation of risk and identity in daily life (Simon et al., 2023; Wolf et al., 2018). Yet the methodological response to this complexity in medical nutritional intervention research has remained limited. Randomized controlled trials in this field have predominantly captured clinical and biomarker outcomes while leaving the mechanisms, meanings, and lived experience of dietary interventions unexplored. Mixed methods designs that integrate qualitative and quantitative strands within a single trial are rare in nutritional research and, to our knowledge, have not yet been reported in the CeD literature: a systematic PubMed search using the terms "celiac disease", "mixed methods", and "randomized controlled trial" yielded no results (searched March 2026).

Furthermore, patient and public involvement, increasingly recognized as a quality criterion in clinical trial design, remains underdeveloped in investigator-initiated dietary trials, where participant perspectives are seldom systematically incorporated into protocol development or outcome selection. The present study was designed to address these methodological gaps simultaneously.

We therefore designed a trial comparing a „relaxed GFD“ (rGFD) permitting cross-contact of up to 200 ppm with an optimized strict GFD over 24 weeks (n = 30; 15 per group). The villus height to crypt depth ratio (Vh:Cd) serves as the primary safety endpoint, with the pre-specified margin for a clinically meaningful deterioration set at a reduction of 0.4 units which corresponds to the error margin of standardized morphometric assessment and is therefore the smallest change that can reliably be distinguished from measurement variability (Taavela et al., 2013).

Methods

Study design and framing

The study is an investigator-initiated, single-center, single-blind, parallel-group, randomized controlled trial embedded within a convergent parallel mixed methods design (Creswell & Plano Clark, 2017); a concurrent qualitative strand of equal weight is integrated into the experimental structure to address questions the trial alone cannot answer.

Philosophical and theoretical foundation

The study adopts a pragmatist worldview (Cook, 1993; Johnson & Onwuegbuzie, 2004; Morgan, 2007), which licenses the combination of quantitative and qualitative inquiry by treating the research question, rather than paradigmatic allegiance, as the arbiter of method. Within this worldview, dialectical pragmatism (Johnson, 2009, 2017) serves as the methodological bridge between the strands, holding quantitative and qualitative findings as complementary and equally weighted knowledge sources whose convergences, complementarities, and divergences are themselves informative.

The study's substantive social-theoretical lens is social practice theory (SPT) (Schatzki, 2010; Schatzki et al., 2001; Warde, 2005; Warde, 2016). On this view, dietary conduct in CeD is not an attitude or a series of rational choices but a social practice: a routinized, embodied bundle of doings and sayings, held together by practical understanding, rules, and a teleoaffective structure, and inseparable from the material arrangements (foods, kitchens, food labeling, restaurants) in and through which it is enacted. The knowledge that carries such a practice is largely tacit and held in the body, a *knowing-how* as well as a *knowing-that*, and the practice is sustained through repetition. Crucially, repetition secures not only the reproduction of a practice but also the possibility of its transformation, since no enactment is identical to the one before.

This matters for the present intervention because, following diagnosis, gluten-free eating does not merely function as a treatment but consolidates into a norm. Norms of this kind are not adopted as propositions and subsequently applied; they are generated and sustained in the doing, emerging from the interaction between practices and the food environments in which they are performed (Mayes et al., 2026). What is repeatedly enacted — the second toaster, the question asked of the server, the product returned to the shelf — becomes the normal, and normality then exerts a force of its own, independently of the evidence that first motivated it. On this account, eating practices performed in the context of CeD generate new norms and, with them, new ways of doing and saying. The same logic indicates how such a norm can be moved: if norms are made by doing, then altering what is done is the mechanism through which a relaxed norm becomes available, which is precisely what the intervention attempts (see Intervention and comparator).

This lens is load-bearing rather than decorative. It reframes strict gluten avoidance as an entrenched, materially and bodily anchored practice rather than a modifiable belief, and it specifies the mechanism by which such a practice can be moved: if norms are made in the doing, they are altered by altering what is done. This is why the intervention is designed as a transformation of practice across its material, embodied, affective, and social dimensions rather than as the transmission of new information (see Intervention and comparator); why acceptance is understood as the reorganization of established practice bundles rather than a change of opinion (see Acceptance as an implementation outcome); and why the qualitative strand analyzes participants' accounts through the constitutive dimensions of practice (see Qualitative strand).

Objectives, research questions and endpoints

The primary question of the quantitative strand is one of non-inferiority: whether a rGFD permitting cross-contact of up to 200 ppm preserves mucosal integrity relative to an optimized strict gluten-free diet, within a margin for clinically meaningful deterioration of 0.4 Vh:Cd units (Taavela et al., 2013).

Let μ_{rGFD} and $\mu_{comparator}$ denote the mean change in the Vh:Cd ratio from baseline to week 24 in the intervention and comparator arm, respectively. The null hypothesis states that cross-contact up to 200 ppm leads to a reduction in Vh:Cd that reaches or exceeds this margin ($H_0: \mu_{rGFD} - \mu_{comparator} \leq -0.4$); the alternative hypothesis states that it leads to no change, or to a reduction not greater than 0.4 units ($H_1: \mu_{rGFD} - \mu_{comparator} > -0.4$). The corresponding estimand is the between-group difference in the change in Vh:Cd from baseline to week 24 (rGFD minus optimized strict GFD) in all randomized participants. Given the pilot sample size, this difference is estimated and interpreted descriptively in relation to the margin rather than subjected to confirmatory testing (see Statistical analysis); formal non-inferiority testing is deferred to the definitive trial that this pilot is designed to inform.

Secondary objectives address the immune response to the rGFD, its acceptability relative to the intervention, gastrointestinal symptoms, estimated gluten exposure from cross-contact, psychosocial wellbeing, and how adults with CeD experience and negotiate dietary flexibility in daily life. The qualitative strand addresses the following research question: How do patients experience and shape the de-implementation of the strict diet in various practices of their everyday lives?

The corresponding endpoints, their instruments, and their timepoints are specified under Outcomes and instruments.

Acceptance as an implementation outcome

Acceptance is examined as an implementation outcome rather than a clinical efficacy endpoint. Following Proctor et al. (2011), acceptability is defined as the perception that the intervention is agreeable and satisfactory, given the recipient's knowledge and experience at the time. Throughout this protocol, we use acceptance as the broader term, encompassing both this perceptual dimension and its behavioral expression : the willingness to be randomized to a relaxed diet, and to adopt and continue it once allocated. Acceptability is thus the measured attitudinal component of acceptance, while acceptance additionally comprises enactment.

This distinction matters here because the rGFD is not the adoption of a new behavior but the deliberate de-implementation of an entrenched one: the habitual strict avoidance that adults with CeD commonly maintain beyond what the evidence supports. De-implementation inverts the usual acceptance problem. What is asked of participants is not to take up something unfamiliar, but to relinquish a practice that has come to feel protective and that is reinforced by clinical advice, patient organizations, and everyday vigilance. Willingness to relax may therefore be constrained less by the perceived merits of the intervention than by the perceived risk of abandoning the established one — which is why acceptance in this trial must be characterized through enacted behavior and not through attitudinal report alone. The intervention is accordingly grounded in the Framework for Taking Action on Overuse (Parchman et al.,

2017) and structured through an Implementation Research Logic Model (Smith et al., 2020) (see Additional file 1), which specifies, a priori, the links between implementation determinants, strategies, mechanisms, and outcomes. The implementation strategies (a stakeholder-co-designed educational video, explicit risk-philosophy framing, and iterative biomarker-informed evaluation) are hypothesized to act through defined mechanisms: recalibration of risk perception, reduction of cross-contact overestimation, and the reorganization of established dietary practice bundles in the sense of Schatzki's Social Practice Theory (Schatzki, 2010; Schatzki et al., 2001).

Acceptability is measured with the Acceptability of Intervention Measure (AIM); (Kien et al., 2021; Weiner et al., 2017), capturing anticipated acceptability at baseline and experienced acceptability at weeks 6–24 ((Sekhon et al., 2017); see Outcomes and instruments), and is contextualized by the qualitative strand's analysis of how acceptance is negotiated in everyday practice. Whether adults with CeD find this relaxation acceptable is a precondition both for a definitive trial and for any eventual change in dietary practice; acceptance is accordingly treated as a piloted signal that the design is built to characterize.

Setting and registration

The trial is conducted at the Swiss Celiac Center at GastroZentrum Hirslanden, Zurich, Switzerland, which provides specialist gastroenterological and dietetic care for adults with CeD. This manuscript reports protocol version 1.2, dated 3rd of August 2026. Recruitment is passive and inbound: potential participants self-refer in response to study information, and no individuals are identified from internal records. Enrollment continues until 30 participants have been randomized or until 31 October 2026, whichever occurs first; the pace of enrollment is constrained primarily by endoscopic capacity rather than by candidate supply. Individual participation lasts 24 weeks from randomization, with on-site visits every six weeks and duodenal biopsies at baseline and week 24, so that the final participant's closing visit is expected around April 2027.

The trial is registered in the German Clinical Trials Register (DRKS00039577) and was approved by the responsible cantonal ethics committee through the Swiss Business Administration System for Ethics Committees (BASEC) portal (BASEC 2026-00721). Under the Swiss Clinical Trials Ordinance it is categorized as a category B clinical trial and conducted as a low-risk study under a risk-based monitoring plan, in accordance with the Human Research Act and International Council for Harmonisation (ICH) Good Clinical Practice. Reporting follows the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT 2025) (Chan et al., 2025), extended by the Consolidated Standards of Reporting Trials (CONSORT) guidance for pilot and feasibility trials (Eldridge et al., 2016) and the Good Reporting of A Mixed Methods Study (GRAMMS) reporting standards for mixed methods studies (O'Cathain et al., 2008); the completed SPIRIT and GRAMMS checklists are provided as additional file 3 and 4, and the schedule of enrolment, interventions, and assessments appears as Figure 1.

Recruitment opened at the end of June 2026. At the time of writing (July 2026), the trial is in active recruitment and data collection. No outcome data are available, and this manuscript reports the protocol

as prospectively registered and approved.

Participants and eligibility

Eligible participants are adults aged 18 to 80 years with a documented initial diagnosis of seropositive CeD, confirmed by biopsy or, where histologic documentation at diagnosis is unavailable, by a tissue transglutaminase immunoglobulin A (tTG-IgA) level exceeding ten times the upper limit of normal at diagnosis, made at least twelve months before baseline. They must have adhered to a GFD for at least twelve months, no villus atrophy at baseline (Vh:Cd $\geq 2.0:1$ on distal duodenal biopsies (Lebwohl et al., 2024)), and be willing to maintain the assigned dietary approach and to keep a weekly record of cross-contact-related eating. Written informed consent is required.

Participants are excluded for seronegative celiac disease, since celiac-specific antibodies serve as safety monitoring; active symptoms, defined as a Celiac Symptom Index (CSI) (Leffler et al., 2009) score above 36 within four weeks before screening; concomitant intestinal disease that could confound symptom assessment (including Crohn's disease, ulcerative colitis, other inflammatory bowel disease, severe irritable bowel syndrome, microscopic colitis, small intestinal bacterial overgrowth, or exocrine pancreatic insufficiency); relevant metabolic disorders (obesity with a body mass index above 35, type 1 diabetes, or insulin-treated type 2 diabetes); mental disorders liable to interfere with eating habits (major depression, eating disorders, or anxiety disorders); existing or intended pregnancy or breastfeeding; continuous use of drugs known or suspected to affect villous architecture (detailed in additional file 2); use of oral gluten-digesting enzyme preparations (for example *Aspergillus niger*-Prolyl-Endopeptidase) within four weeks before baseline, or unwillingness to discontinue them; participation in the study's stakeholder engagement process or in another modified-GFD or celiac-related trial within the preceding twelve months; and inability to follow study procedures for reasons of language, psychological disorder, or dementia. The full eligibility criteria are provided in Additional file 2.

Patient and public involvement

Patient and public involvement is embedded throughout the trial lifecycle through a structured stakeholder engagement process, conducted as a separately registered sub-project (BASEC Req-2025-01479) and informed by the 7P framework (Concannon et al., 2019; Concannon et al., 2012), which distinguishes patients and the public, providers, purchasers, payers, policymakers, product makers, and principal investigators. Approximately 10 to 15 stakeholder representatives, with patient and public representatives at its center, take part in structured consultations at three points in the trial lifecycle: a pre-study consultation to refine the research questions, intervention design, educational materials, and outcome selection; an optional mid-study consultation to address implementation challenges; and a post-study consultation to co-interpret the findings and co-develop dissemination. Stakeholder input during design confirmed that both dietary conditions are experienced as a meaningful reduction in dietary burden, supporting the choice of an active rather than a sham comparator, and informed the framing of the educational materials. All inputs are systematically documented, and their influence on study procedures is reported transparently in accordance with the GRIPP2 reporting checklist for patient

and public involvement (Staniszewska et al., 2017). Trial participants are distinct from the stakeholder representatives; participation in the engagement process is an exclusion criterion for the trial. Trial results will be disseminated through publication in a peer-reviewed journal, posted in the trial registry (DRKS), and communicated to participants and the wider celiac community after consulting with the stakeholder representatives.

Intervention and comparator

Overview and rationale

Both arms receive the identical evidence base on gluten cross-contact, derived from the study team's preparatory work: a scoping review of worldwide cross-contact data ((Studerus et al., 2025); 59 studies, 28 countries, 27,806 food samples), a targeted laboratory analysis closing the evidence gaps it identified (products containing wheat starch or barley malt, precautionary-allergen-labeled gluten-free grains, oat drink, and soy sauce; two manuscripts in preparation, project registered on the Open Science Framework (<https://osf.io/3esxc>)), and tolerance data (Rostami-Nejad et al., 2023). The experimental contrast lies not in the information provided but in its interpretive framing and in the depth of practice change each arm is invited to undertake. Social Practice Theory (Schatzki, 2010; Schatzki et al., 2001; Warde, 2016) informed not only this framing but the construction of the rGFD itself, which was designed to address the material, embodied, affective, and social dimensions of established avoidance practices rather than cognitive knowledge alone.

Relaxed gluten-free diet (intervention arm)

The rGFD permits foods carrying a precautionary allergen label with a gluten content up to 200 ppm (median), foods containing wheat starch or barley malt, and practices such as using shared kitchen equipment (utensils, toaster, deep fryer), under the guiding heuristic "when in doubt, eat it." Based on typical consumption of 300–500 g of gluten-free products per day, this corresponds to an estimated exposure of 60–95 mg gluten/day, which remains below the 150 mg/day threshold associated with only minimally increased relapse risk (relative risk 1.07; (Rostami-Nejad et al., 2023)). The intervention alters the decision rule applied under uncertainty, not the quantity consumed. Participants are given no instruction on how much or how little of the permitted foods to eat; the extent to which the permitted practices are taken up is left to everyday circumstance and is captured, not directed, by the weekly food diary. Whether a relaxed decision rule translates into measurably higher gluten exposure is therefore an empirical question of this pilot rather than a design assumption, and is examined through the achieved-exposure analysis specified under Statistical analysis. A product-category risk stratification (low, moderate, high) structures the messaging so that relaxation is evidence-based rather than indiscriminate. Conceptually, the rGFD is framed as a deliberate de-implementation of habitual strict gluten-avoidance behavior (Parchman et al., 2017): it replaces reflexive avoidance („doing less“) with

informed trust, and is defined not by the absence of caution but by the redirection of vigilance toward the situations where cross-contact is clinically relevant („doing more“).

Optimized strict gluten-free diet (comparator arm)

The comparator maintains the conventional strict threshold (≤ 20 ppm) under the heuristic "when in doubt, avoid it," but is an active, genuinely beneficial comparator rather than usual care or a sham: the same evidence base is used to prioritize vigilance, directing attention to situations where cross-contact is clinically relevant and reassuring participants where it is negligible. Both arms therefore relax somewhat relative to the maximal avoidance historically assumed by patients; they differ in how far and under which interpretive frame. This evidence-based relaxation of the comparator was endorsed by stakeholders during the study's stakeholder engagement process and constitutes a genuine, clinically defensible relaxation of independent benefit; the comparator is thus an active comparator. Its genuineness is precisely what sustains participant blinding: because both arms undergo a real and credible relaxation of previously avoided practices, participants cannot readily infer their allocation. This non-disclosure of allocation and comparator content was approved by the responsible ethics committee.

Delivery and materials

Each arm receives a group-specific, evidence-based educational screencast video and an accompanying written handout, built from identical data with differentiated risk-interpretation frameworks. Materials are equally resourced across arms to control for attention and educational exposure. The video is hosted on the Swiss Celiac Center website; after the histology results have been received and definitive inclusion confirmed, the investigator (DS) sends the group-specific personal video link and QR code together with the handout by secure messaging channel (Threema Work; Threema GmbH, Switzerland) .

Timing

The intervention begins on the day following delivery of the educational materials but is counted as formally initiated only once the participant has viewed the video; viewing is verified through the hosting platform, and non-viewing within two days prompts a proactive reminder. Standardized screencast delivery ensures identical intervention content within each arm. Ongoing contact is maintained through a secure messaging channel (Threema Work; Threema GmbH, Switzerland). Further the focused weekly food-diary functions as a structured retention measure.

Randomization, blinding, and safety switch

Randomization and allocation concealment

Eligible participants are randomly assigned in a 1:1 ratio to the rGFD or the optimized strict gluten-free diet using a computer-generated sequence with permuted blocks of variable length (block sizes of two and four), yielding a balanced target of 15 participants per arm. The allocation sequence is generated in R Studio with a documented seed for auditability, implemented in the electronic data capture system (REDCap; Vanderbilt University), and its correct application is verified by the clinical trials unit (CTU) as part of monitoring. The variable, concealed block structure ensures that the person performing randomization cannot anticipate the next assignment, minimizing selection bias.

Timing of randomization

Randomization is performed only after histopathological confirmation of baseline eligibility (Vh:Cd $\geq 2.0:1$), that is, after the baseline biopsy has been obtained and assessed. Because the primary safety measurement is therefore taken before an allocation exists, baseline outcome assessment is structurally independent of group assignment. Participants whose baseline histology does not confirm the absence of atrophy are recorded as screening failures.

The trial is single-blind. Participants are blinded to their allocation and to the content of the comparator arm; because both arms receive equally resourced educational materials that differ only in interpretive framing, allocation cannot be inferred from the materials. Full double-blinding is not feasible for a behavioral dietary intervention. Outcome assessment is protected where it matters most: the histopathologist assessing the primary endpoint (Medica, Zurich), the sponsor-investigator and investigator performing the endoscopies, and the laboratory conducting the blood and stool analyses (Analytica, Zurich) remain blinded to allocation throughout. Study visits and interviews are conducted by two members of the study team. The investigator (DS) is necessarily unblinded, as her role comprises randomization and the distribution of the group-specific educational materials. The research assistant (VD) is blinded to allocation and has no access to the randomization records. Because both interview guides are group-blind by design, using experiential rather than rule-based questioning, visits and interviews can be conducted without disclosing group assignment and remain comparable across arms irrespective of who conducts them.

Because participants are recruited to a single center from a patient population in which mutual contact is common, exchange of information between arms cannot be structurally excluded. Contact with the study team is therefore conducted individually through the secure messaging channel, and no group sessions, group materials, or shared discussion formats are used at any point in the trial; the educational video link is issued personally to each participant.

Individual safety switch

From week 6 onward, an elevation of tTG-IgA above 10 U/mL triggers an individual switch of that participant from the rGFD to the comparator diet. Switched participants remain in the study and continue to be assessed and analyzed within their originally allocated group, with follow-up continuing until week

24. The frequency of safety switches is itself a predefined feasibility and safety signal and feeds the progression criteria.

Study-level stop criterion and notification

The occurrence of more than two individual safety switches (approximately 20 % of the intervention arm) constitutes a study-level stop criterion. Should immediate safety or protective measures become necessary during conduct, such as emergency switches or temporary suspension, the sponsor-investigator notifies the ethics committee within seven days, in accordance with Article 62b of the Swiss Clinical Trials Ordinance.

Data monitoring

Given the low-risk category B classification, the short duration, and the pre-specified individual safety switch and study-level stop criterion, no independent data monitoring committee was established; safety oversight is exercised by the sponsor-investigator under a risk-based monitoring plan overseen by the CTU Hirslanden, with reporting to the cantonal ethics committee as required.

Outcomes and instruments

All measurement timepoints are shown in Figure 1.

Primary endpoint. The villus height to crypt depth ratio (Vh:Cd) from distal duodenal biopsies, measured by histomorphometry on calibrated digital microscopy by a blinded histopathologist at baseline and week 24.

Immune response

- tTG-IgA at baseline and weeks 6, 12, 18, and 24 (remission reference <7 U/mL; an elevation >10 U/mL from week 6 triggers the safety switch);
- the modified Marsh-Oberhuber grade (types 0–3c) with intraepithelial lymphocyte (IEL) count at baseline and week 24;
- and fecal anti-human tissue transglutaminase secretory IgA (anti-htTG sIgA) at baseline and weeks 6, 12, 18, and 24, quantified with the IDK® Anti-htTG sIgA enzyme-linked immunosorbent assay (ELISA) (Immundiagnostik AG, Bensheim, Germany; Cat. No. K9393; CE-marked in vitro diagnostic; cut-off 100 U/L).

Fecal anti-htTG sIgA is used exploratively and directionally and not as a surrogate for mucosal activation. A value above the manufacturer's cut-off is interpreted as evidence of probable mucosal immune activation attributable to gluten exposure. A negative result neither excludes exposure nor confirms adherence, given the well-documented insensitivity of transglutaminase antibodies to low-dose dietary transgression. The diagnostic performance of the fecal assay for this purpose has not been

established, and the longitudinal trajectory across the five timepoints is therefore more informative than any single measurement.

Acceptance

- The AIM; (Kien et al., 2021; Weiner et al., 2017), with anticipated acceptability at baseline and experienced acceptability at weeks 6, 12, 18, and 24, reported as distinct constructs on a common scale.

Consent for using the AIM was given by the authors of the validated German version.

Symptoms

- The CSI; (Leffler et al., 2009), 16 items, total 16–80, higher scores indicating greater burden) at baseline and weeks 6, 12, 18, and 24.

As no validated German version existed at study design, a forward–backward translation was performed following established cross-cultural adaptation guidance with consent by the authors. A CSI score exceeding 36 points within four weeks before screening constitutes an exclusion criterion.

Estimated gluten exposure

- A focused weekly food diary completed at home (estimated intake in mg/week)

with intake derived by combining product-category contamination values from the study team's preparatory analyses (third-quartile values as a conservative floor) with standard portion sizes from the Swiss national nutrition survey (menuCH);

- fecal gluten immunogenic peptides (GIP) at baseline and weeks 6, 12, 18, and 24, quantified with the IDK® Gluten Fecal ELISA (Immundiagnostik AG; Cat. No. KR9309), targeting the immunodominant gliadin 33-mer.

Given its limit of quantification (approximately 1.16 µg/g stool), GIP is used as a threshold detector for substantial gluten exposure rather than as a sensitive marker of trace-level exposure; this ceiling-effect avoidance is an intentional design choice, since a positive result specifically flags intake well above the intended range within the intervention.

Psychosocial wellbeing

- Food-related quality of life
- anxiety about gluten ingestion
- willingness to take risks each measured on a 100 mm visual analogue scale (VAS) at baseline and weeks 6, 12, 18, and 24.

Visual analogue scales were chosen for these constructs after consideration of existing instrumentation. The Celiac Disease Food Attitudes and Behaviours scale (CD-FAB) is the only disease-specific instrument addressing an adjacent domain, but it captures maladaptive food attitudes as a clinical construct and no cut-offs for clinical relevance have yet been established (Gholmie et al., 2023) which limits the interpretability of within-participant change over the trial period. Single-item measurement of conceptually narrow and clearly defined subjective states is methodologically defensible where the construct does not require multiple facets to be represented (Bowling, 2005), and it limits respondent burden across five assessment timepoints in a battery that already comprises the CSI, the AIM, and a prospective food diary (Streiner et al., 2014). The concurrent qualitative strand contributes validity evidence by allowing convergence or divergence between the scaled ratings and participants' accounts of their lived experience to be examined at the point of integration.

Lived experience. Participants' lived experience of dietary change is captured through semi-structured interviews analyzed by qualitative content analysis (see Qualitative strand).

Safety endpoints. tTG-IgA elevation above 10 U/mL (reference: < 7 U/mL) triggering the individual safety switch from week 6; adverse events related to the dietary intervention; and withdrawal rates with reasons.

Statistical analysis

The primary analysis estimates the between-group difference in Vh:Cd change over 24 weeks from a linear mixed-effects model of the form $\text{outcome} \sim \text{group} \times \text{time} + (1 \mid \text{participant})$, fitted in R Studio (lme4, lmerTest); this difference is interpreted descriptively in relation to the -0.4 margin (Taavela et al., 2013) rather than tested for non-inferiority. Group-by-time estimated marginal means and their contrasts are obtained with the emmeans package and reported with 95% confidence intervals, and the same model form is applied to the secondary clinical and patient-reported outcomes. A trend analysis relates ordered gluten-exposure categories to Vh:Cd change (and IEL change) (emtrends), and the modified Marsh-Oberhuber grade is treated as ordinal. Analyses are estimation-focused and exploratory, emphasizing effect sizes and precision; no multiplicity adjustment is applied.

Because the intervention prescribes a decision rule rather than a fixed dose, achieved exposure is treated as a measured quantity rather than an assumption. The between-arm difference in estimated weekly gluten exposure from the food diary and the proportion of participants with fecal GIP above the limit of quantification are reported descriptively as a check of differential uptake, and condition the interpretation of the primary estimate: in the absence of an exposure differential between arms, a null finding on Vh:Cd would be uninformative with respect to safety. The magnitude of any observed differential is itself a design finding relevant to the definitive trial.

Analyses follow the intention-to-treat principle. Missing data are assumed missing at random and addressed by multiple imputation respecting the longitudinal structure, with results compared with and without imputation as a sensitivity analysis. Model assumptions are examined through residual

diagnostics; as ratio outcomes and summed Likert scores are not expected to be normally distributed, robust estimators, transformations, or generalized mixed models are used where indicated.

Sample size

No formal power calculation was performed, as hypothesis-driven power calculations do not determine sample size in pilot trials (Eldridge et al., 2016; Lancaster & Thabane, 2019; Lee et al., 2014). The sample of 30 participants (15 per arm) is determined by feasibility and by the available capacity for 60 endoscopic procedures (baseline and week 24), while consistent with established recommendations for pilot trial sample sizes (Julious, 2005; Teare et al., 2014). The sample size is characterized in terms of the precision it affords rather than the effects it would allow to be detected: with 15 participants per arm, the half-width of the 95 % confidence interval for a between-group difference in Vh:Cd change corresponds to approximately 0.74 standard deviations. A central purpose of the pilot is to generate the variance estimates required to inform the design of a subsequent definitive trial.

Progression criteria

The safety-based stop criterion and the recruitment criterion are pre-specified in the study protocol; the remaining progression criteria were specified for the feasibility reporting of this pilot as shown in Table 1. Progression to a definitive trial is judged against these pre-specified feasibility criteria, reported in line with the CONSORT extension for pilot and feasibility trials (Eldridge et al., 2016). Two domains carry binary stop-or-proceed thresholds; the remaining domains are assessed descriptively, their observed values informing the design and assumptions of a definitive trial rather than a formal pass/fail rule, as is appropriate to a pilot of this size.

Qualitative strand

Building on the study's pragmatist worldview and its social practice theory lens, the qualitative strand examines how participants experience and reconfigure everyday dietary practices as gluten avoidance from cross-contact is deliberately relaxed. It uses inductive–deductive qualitative content analysis (Elo & Kyngas, 2008); a separate, theory-generating analysis of the complete qualitative dataset using constructivist grounded theory (Charmaz, 2014) is planned as a follow-up study in 2027 / 2028.

Interviews and data collection

Semi-structured interviews are conducted at five timepoints (baseline and weeks 6, 12, 18, and 24) using two theory-informed guides developed for the study. The baseline guide establishes individual profiles across five domains: biographical relationship with the diagnosis, everyday dietary practices, bodily experience, social dimensions of eating, and the management of uncertainty and risk. The follow-up guide traces practice-transformation trajectories across seven thematic blocks, with a dedicated block on meaning and identity introduced from week 18, once sufficient time has elapsed for reflection on identity work. Both guides are group-blind by design, using experiential rather than rule-based questioning, so that interviews can be conducted without revealing allocation and remain comparable

across arms and across timepoints. Food-diary entries and inter-visit chat messages are treated as supplementary qualitative data alongside the transcripts.

Interviews are conducted by two members of the study team, the investigator (DS) and the research assistant (VD). Both use the same theory-informed guides and were jointly trained in their application, including a shared calibration of probing strategies and of the experiential rather than rule-based questioning style on which the group-blind construction of the guides depends. Interviewer allocation is documented for each interview so that it can be examined as a potential source of variation during analysis. Interview conduct is reviewed at regular intervals against the guides, and interpretive decisions arising during data collection are discussed jointly and recorded in the sampling journal alongside the sampling decisions.

Sampling

All thirty participants are interviewed at baseline and week 6, establishing longitudinal reference profiles. From week 12, a purposive maximum-variation strategy selects participants to represent expected diverse quantitative response trajectories from the parallel trial strand (tTG-IgA, fecal GIP, symptom scores, VAS), with deliberate balance between arms (approximately five to eight per group per timepoint) to protect the credibility of subsequent meta-inference. Detection of fecal GIP above the limit of quantification in the intervention arm constitutes an additional sampling trigger, as it indicates that a participant's dietary practice has departed from the planned cross-contact framework and thus represents a case of particular interpretive interest. From week 18, selection is additionally guided by the concepts emerging from ongoing content analysis of earlier interviews; week 24 interviews continue until thematic saturation, the point at which new interviews yield no substantially new categories (Elo et al., 2014), which is distinct from theoretical saturation in the grounded-theory sense (Saunders et al., 2018). This iterative, concept- and indicator-informed selection is understood abductively, as a disciplined movement between emerging observations and theory (Timmermans & Tavory, 2012). All sampling decisions, including the quantitative indicators considered and group balance, are recorded in a sampling journal; approximately 90–108 interviews are anticipated across all timepoints.

Analysis

Interviews are imported into MaxQDA, transcribed verbatim, and analyzed following the three phases of qualitative content analysis: preparation, organization, and reporting (Elo & Kyngas, 2008). The unit of analysis is the meaning unit, a coherent segment ranging from a phrase to several sentences, with the whole transcript retained as context (Graneheim & Lundman, 2004); both manifest and latent content are considered. Organization proceeds inductively and deductively: a deductive, unconstrained categorization matrix derived from the dimensions of social practice in Schatzki's Social Practice Theory (embodied knowledge, material arrangements, social meanings, and competences; (Schatzki, 2010; Schatzki et al., 2001) structures initial coding, while inductive open coding identifies concepts beyond the a priori frame, which are then reconciled with the matrix. Because the matrix is unconstrained,

categories outside the initial framework can be created, preserving openness to emergent meaning; the interplay between frame and data is handled abductively.

Trustworthiness and researcher positionality

Rigor is established through trustworthiness criteria specific to qualitative content analysis (Elo et al., 2014), addressing credibility, dependability, confirmability, and transferability across the preparation, organization, and reporting phases: prolonged engagement over the 24-week follow-up, an audit trail through the sampling journal and complete MaxQDA documentation of coding decisions, independent review of a subset of coding by an external researcher, member checking of selected interpretations at week 24, and thick description of sample, context, and sampling decisions. The principal qualitative researcher (DS) is both investigator and a person with CeD; this insider position is treated as a productive asset for theoretical sensitivity and rapport, managed through systematic reflexivity and transparent reporting of positionality (Berger, 2013).

Integration

Integration is treated as a continuous design principle rather than a terminal step (Bazeley & Vogl, 2026). It operates at three levels: connecting at the sampling stage, merging at the level of interpretation through joint displays (Creswell & Plano Clark, 2017; Fetters et al., 2013), and meta-inference, at which conclusions drawn from the two strands are brought into a single interpretive account (Tashakkori & Teddlie, 2010).

Connecting: at sampling

From week 12, response trajectories from the quantitative strand (tTG-IgA, fecal GIP, symptom scores, VAS) together with the concepts emerging from ongoing content analysis jointly inform the selection of participants for in-depth interviews (see qualitative strand). This connection is directional: quantitative and qualitative data determine whom to interview, but qualitative findings never alter what or when quantitative outcomes are measured. The trial protocol therefore remains fixed, preserving the internal validity of the randomized strand, and the interview guide is held constant, so that connecting sampling does not compromise comparability.

Merging: at interpretation

Once each strand has been analyzed independently, quantitative trajectories and qualitative categories are brought together in joint displays at three predefined integration points (weeks 12, 18, and 24). Each joint display juxtaposes quantitative results (group-wise tTG-IgA and GIP trajectories, CSI and VAS trajectories) with qualitative findings (emergent categories) and is examined explicitly for convergence, where both strands confirm the same phenomenon from different angles; complementarity, where each illuminates a different facet; and divergence, where the strands appear to contradict one another. Divergence is treated not as error but as an analytically productive signal that typically points to contextual factors invisible to either strand alone (drawing on dialectical pragmatism; (Johnson, 2017)).

Inference: meta-inference

The integrated findings are synthesized into a meta-inference, a conclusion that answers the overarching mixed methods question, which neither strand could address alone: whether and how a rGFD can be safely and meaningfully implemented in the daily lives of adults with celiac disease? The trial strand establishes whether relaxation is clinically safe; the qualitative strand illuminates the practices, meanings, and material and social conditions that shape whether and how it can be lived. Member checking of selected interpretations at week 24 feeds this stage.

Discussion

This protocol describes, to our knowledge, among the first mixed methods designs embedded within an experimental framework with the qualitative and quantitative strands carrying equal methodological weight, in a dietary intervention in CeD. It addresses a gap left by an evidence base dominated by fixed-dose challenge studies, acute biomarker experiments, and cross-sectional contamination analyses, none of which corresponds to the condition that actually determines whether relaxation is safe and accepted: distributed, chronic, food-borne gluten exposure under a deliberately relaxed diet, evaluated prospectively against a histological endpoint.

Strengths

Several features of the design merit emphasis. First, the sequence of baseline endoscopy, histopathological confirmation of eligibility, and only then randomization, ensures that the primary endpoint at baseline is structurally independent of group assignment, since no allocation exists at the time of the initial biopsy, removing a source of baseline bias that is difficult to control where randomization precedes the first outcome measurement.

Second, the use of an active, genuinely beneficial comparator rather than a sham or usual-care control means that any between-group difference in the primary endpoint can be attributed to the specific content of the intervention rather than to differences in attention, educational exposure, or perceived intensity; both arms receive materials of equivalent format and dietitian contact, grounded in identical cross-contamination evidence.

Third, assessor blinding is achieved at several independent levels: the sponsor-investigator performing endoscopy, the histopathologist assessing the primary endpoint, the laboratory conducting blood and stool analyses, and the research assistant conducting study visits. No single unblinded actor can influence outcome ascertainment. This multi-level approach, in which study personnel are blinded to allocation, to outcomes, or to both according to their role, follows the recommendations for blinding in behavioral trials, where uniform masking of the kind achievable with placebo-controlled pharmacological designs is not possible (Friedberg et al., 2010).

Fourth, the mixed methods architecture reflects the complexity of dietary behavior as an intervention target and aligns with the Medical Research Council framework for complex interventions (Skivington et al., 2021), with a substantive grounding in social practice theory that is load-bearing rather than ornamental: it motivates the choice of inductive–deductive content analysis and positions the intervention as the de-implementation of an avoidance practice embedded in daily routines, material arrangements, and social relations. Finally, the systematic stakeholder engagement conducted as a separately registered sub-project, using the 7P framework (Concannon et al., 2019; Concannon et al., 2012), embeds patient and public involvement across the trial lifecycle, so that intervention design, materials, and outcome selection reflect the perspectives of those most directly affected.

Practical and operational considerations

Several operational aspects shaped the design and warrant transparent description. The rate-limiting constraint on the study is not candidate supply but endoscopic capacity: sixty procedures (baseline and week 24) must be accommodated within routine clinical service, which determines both the recruitment pace and the single-center setting. Recruitment is therefore governed by a defined window and an explicit stop criterion (30 participants or 31 October 2026, whichever occurs first), and the single site, while limiting immediate generalizability, secures procedural consistency in intervention delivery and outcome assessment.

Maintaining blinding in a dietary trial requires deliberate operational separation of roles. All allocation-revealing tasks, namely randomization and the distribution of the group-specific educational materials, are carried out exclusively by the unblinded investigator (DS). Study visits and interviews are conducted by both DS and the blinded research assistant (VD), using instruments and interview guides that are identical across arms; access rights in the electronic data capture system prevent VD from viewing allocation or safety-monitoring data. Non-disclosure of the comparator arm is set out in advance in the informed consent form and resolved by debriefing at study completion, under ethics-committee approval.

The individual safety switch is operationalized through scheduled tTG-IgA monitoring from week 6; a participant crossing the threshold is moved to the comparator diet, continues in follow-up until antibody normalization, and is analyzed within the originally allocated group, with more than two switches constituting a study-level stop criterion.

Intervention fidelity cannot be observed directly, since dietary behavior in everyday life is not amenable to continuous monitoring. This is addressed through a complementary, multi-level exposure strategy (the weekly food diary, fecal gluten immunogenic peptides, and fecal anti-htTG secretory IgA) together with verification that the educational video has been viewed. These sources capture different parts of the exposure spectrum rather than a single complete measure of adherence, and are interpreted accordingly. The qualitative strand adds its own operational demands: iterative, concept- and indicator-driven sampling from week 12, systematic documentation in a sampling journal, and the management of

researcher positionality. Data capture and monitoring are handled through REDCap under a risk-based monitoring plan overseen by the CTU Hirslanden.

Limitations

Several limitations follow directly from the pilot nature of the study. The sample of thirty participants was determined by feasibility and endoscopic capacity rather than by a power calculation; the trial is consequently not powered for confirmatory clinical inference, estimates of between-group differences will be imprecise, and the clinical outcomes are interpreted by estimation rather than hypothesis testing.

The study population is deliberately selected and may not represent newly diagnosed or non-adherent patients, and self-referral introduces a degree of volunteer bias. The 24-week duration may not capture long-term or cumulative effects; this is a particular consideration given that the long-term clinical significance of low-dose immune activation reported at doses far below labeling thresholds (Daveson et al., 2026) remains unknown, and it is precisely such cumulative effects that a definitive trial would need to address.

Complete masking of treatment assignment is not achievable in a dietary intervention, as participants necessarily know their own recommendations and those familiar with cross-contact research may infer their allocation; the active comparator, the equipoise framing, and multi-level assessor blinding mitigate but do not eliminate this. A proportion of interviews is conducted by the unblinded investigator, so that knowledge of allocation cannot be excluded as an influence on questioning or on the interactional dynamics of those interviews. This is mitigated by the group-blind construction of the interview guides, by the parallel conduct of interviews by a blinded research assistant, and by the reflexivity procedures described above, but it remains a limitation of the qualitative strand.

Combining the quantitative and qualitative strand has its own limitations: integration juxtaposes a whole-sample quantitative signal with a purposively selected qualitative subset, connected sampling means the strands are not fully independent at merging, and the pilot size renders any meta-inference provisional rather than confirmatory.

A further limitation follows from the intervention's design: it permits, but does not prescribe, higher gluten exposure. Should both arms converge on comparable exposure, the trial would characterize the acceptability of a relaxed decision rule without providing an informative test of its mucosal safety. Exposure data are therefore reported alongside the primary estimate rather than assumed, and the direction of any such convergence would itself inform whether a definitive trial requires a more directive intervention.

Finally, the fecal GIP assay, used deliberately as a threshold detector, does not quantify trace-level exposure, and neither tTG-IgA nor anti-htTG sIgA serves as a surrogate for mucosal immune activation.

Contribution and outlook

The trial is intended as the empirical and methodological foundation for a larger, confirmatory, multicenter study. It will generate the variance and effect estimates required to power such a trial, together with feasibility data on recruitment and safety-switch frequency, and it will refine the intervention, the biomarker strategy and the mixed methods integration framework. In this sense the study makes a dual contribution: it produces preliminary evidence on the safety and acceptance of dietary relaxation within a GFD in CeD, and it establishes a methodological template for the rigorous de-implementation and evaluation of complex dietary interventions. Should the findings point toward safety and acceptability, they would justify a definitive trial, one capable of confirming whether the burden of maximal vigilance can be calibrated to actual physiological risk, the proportionality question with which this work began.

Abbreviations

AIM: Acceptability of Intervention Measure

anti-htTG sIgA: anti-human tissue transglutaminase secretory immunoglobulin A

BASEC: Business Administration System for Ethics Committees

CD-FAB: Celiac Disease Food Attitudes and Behaviours scale

CeD: celiac disease

CONSORT: Consolidated Standards of Reporting Trials

CSI: Celiac Symptom Index

CTU: Clinical Trials Unit

DRKS: Deutsches Register Klinischer Studien (German Clinical Trials Register)

ELISA: enzyme-linked immunosorbent assay

GFD: gluten-free diet

GIP: gluten immunogenic peptides

GRAMMS: Good Reporting of A Mixed Methods Study

GRIPP2: Guidance for Reporting Involvement of Patients and the Public, version 2

ICH: International Council for Harmonisation

IEL: intraepithelial lymphocyte

IgA: immunoglobulin A

ppm: parts per million (mg/kg)

REDCap: Research Electronic Data Capture

rGFD: relaxed gluten-free diet

RR: relative risk

SPIRIT: Standard Protocol Items: Recommendations for Interventional Trials

tTG-IgA: tissue transglutaminase immunoglobulin A

VAS: visual analogue scale

Vh:Cd: villus height to crypt depth ratio

Declarations

Sponsor-Investigator

PD Dr. med. Jonas Zeitz, Swiss Celiac Center at GastroZentrum Hirslanden, Zurich, Switzerland and University of Zurich, Faculty of Medicine, Zurich, Switzerland

Ethics approval and consent to participate

This study was approved by the responsible cantonal ethics committee (Kantonale Ethikkommission Zürich; BASEC no. 2026-00721). Under the Swiss Clinical Trials Ordinance the trial is categorized as a category B clinical trial and is conducted as a low-risk study under a risk-based monitoring plan, in accordance with the Human Research Act and the principles of ICH Good Clinical Practice. Personal data are pseudonymized at entry and stored in access-controlled REDCap; data are handled in accordance with the Swiss Federal Act on Data Protection and retained as required by the Human Research Act. The accompanying stakeholder engagement process is registered as a separate sub-project (Req-2025-01479). Written informed consent is obtained by the sponsor-investigator or investigator from all participants at the beginning of the screening visit before any study-specific procedures. The single-blind design involves non-disclosure of the comparator arm; this incomplete disclosure, together with debriefing at study completion, was reviewed and approved by the ethics committee and is described to participants in the consent form. The trial is prospectively registered in the German Clinical Trials Register (DRKS00039577).

Substantial protocol amendments are submitted to the cantonal ethics committee via the BASEC portal and reflected in the trial registry before implementation; non-substantial amendments are documented

and notified in accordance with the Swiss Clinical Trials Ordinance.

Trial insurance

Trial-related harm is covered by Zurich Versicherung in accordance with the Human Research Act and the Clinical Trials Ordinance; after trial completion, participants continue their usual specialist care.

Consent for publication

Not applicable.

Availability of data and materials

Data sharing is not applicable to this article, as no datasets were generated or analyzed for the present protocol. Upon completion of the trial, de-identified data, statistical analysis plan and study materials are available from the corresponding author on reasonable request and will be deposited on the Open Science Framework (<https://osf.io/r7ka4/overview>), subject to the consent provided by participants and applicable data-protection requirements.

Competing interests

This study is funded in part by Dr. Falk Pharma GmbH; this relationship is disclosed as a potential competing interest. Immundiagnostik AG (Bensheim, Germany), the manufacturer of the two fecal assays used in this trial, provided the assay kits at a reduced price and supplied one assay free of charge. Analytica (Zurich, Switzerland) performs the blood and stool analyses at a reduced rate. Medica (Zurich, Switzerland) supports this study financially and performs the histopathological assessment free of charge; E. Marques Maggio is affiliated with Medica. Codegestalt provided free use of the Zettelbox software and technical support for this study; D. Studerus contributes to the development of this software in an advisory and developmental capacity and has no financial interest in Codegestalt. None of these parties had any role in the design of the study, in the selection of endpoints or assays, or in the analysis or interpretation of data, and none will have any role in the decision to publish.

D. Studerus has celiac disease; this personal experience is disclosed as a non-financial interest and is managed through the reflexivity procedures described in the Methods. The authors declare that they have no further competing interests.

Funding

This work is supported by a Candoc grant from the University of Zurich (decision number FK-26-038) awarded to D. Studerus and by contributions from the Hirslanden Forschungsfoundation, Medica Zürich, and Falk Pharma AG. In-kind support was provided by Codegestalt, which granted free use of the Zettelbox software together with technical support, and by Threema, which granted free use of the Threema Work application. Any remaining project costs are covered by the investigators themselves. None of the funders had any role in the design of the study or in the writing of this protocol, and none will

have any role in the collection, analysis, or interpretation of the data or in the decision to submit future reports for publication.

Authors' information (optional)

D. Studerus is an Advanced Practice Dietitian at the Swiss Celiac Center, GastroZentrum Hirslanden, and a PhD candidate in Care and Rehabilitation Science at the University of Zurich; her dual position as clinician-researcher and person with CeD informs the study's design and its reflexive approach to the qualitative strand.

References

1. Bazeley, P., & Vogl, S. (2026). *Integrating Analyses in Mixed Methods Research*. SAGE Publications, Incorporated. <https://books.google.ch/books?id=woCe0QEACAAJ>
2. Berger, R. (2013). Now I see it, now I don't: researcher's position and reflexivity in qualitative research. *Qualitative Research*, 15(2), 219–234. <https://doi.org/10.1177/1468794112468475>
3. Bowling, A. (2005). Just one question: If one question works, why ask several? *Journal of Epidemiology and Community Health*, 59(5), 342. <https://doi.org/10.1136/jech.2004.021204>
4. Cadenhead, J. W., Wolf, R. L., Lebwohl, B., Lee, A. R., Zybert, P., Reilly, N. R., Schebendach, J., Satherley, R., & Green, P. H. R. (2019). Diminished quality of life among adolescents with coeliac disease using maladaptive eating behaviours to manage a gluten-free diet: a cross-sectional, mixed-methods study. *J Hum Nutr Diet*, 32(3), 311–320. <https://doi.org/10.1111/jhn.12638>
5. Catassi, C., Fabiani, E., Iacono, G., D'Agate, C., Francavilla, R., Biagi, F., Volta, U., Accomando, S., Picarelli, A., De Vitis, I., Pianelli, G., Gesuita, R., Carle, F., Mandolesi, A., Bearzi, I., & Fasano, A. (2007). A prospective, double-blind, placebo-controlled trial to establish a safe gluten threshold for patients with celiac disease. *Am J Clin Nutr*, 85(1), 160–166. <https://doi.org/10.1093/ajcn/85.1.160>
6. Chan, A. W., Boutron, I., Hopewell, S., Moher, D., Schulz, K. F., Collins, G. S., Tunn, R., Aggarwal, R., Berkwits, M., Berlin, J. A., Bhandari, N., Butcher, N. J., Campbell, M. K., Chidebe, R. C. W., Elbourne, D. R., Farmer, A. J., Fergusson, D. A., Golub, R. M., Goodman, S. N.,... Hróbjartsson, A. (2025). SPIRIT 2025 statement: Updated guideline for protocols of randomised trials. *PLoS Med*, 22(4), e1004589. <https://doi.org/10.1371/journal.pmed.1004589>
7. Charmaz, K. (2014). *Constructing Grounded Theory*. SAGE Publications. https://books.google.ch/books?id=v_GGAwAAQBAJ
8. Codex Alimentarius, C. (2026). *Guidelines on the use of precautionary allergen labelling*. <https://www.fao.org/fao-who-codexalimentarius/codex-texts/list-standards/en/>
9. Concannon, T. W., Grant, S., Welch, V., Petkovic, J., Selby, J., Crowe, S., Synnot, A., Greer-Smith, R., Mayo-Wilson, E., Tambor, E., Tugwell, P., & Multi Stakeholder Engagement, C. (2019). Practical Guidance for Involving Stakeholders in Health Research. *J Gen Intern Med*, 34(3), 458–463. <https://doi.org/10.1007/s11606-018-4738-6>

10. Concannon, T. W., Meissner, P., Grunbaum, J. A., McElwee, N., Guise, J. M., Santa, J., Conway, P. H., Daudelin, D., Morrato, E. H., & Leslie, L. K. (2012). A new taxonomy for stakeholder engagement in patient-centered outcomes research. *J Gen Intern Med*, 27(8), 985–991. <https://doi.org/10.1007/s11606-012-2037-1>
11. Cook, G. A. (1993). *George Herbert Mead: The Making of a Social Pragmatist*. University of Illinois Press. <https://books.google.ch/books?id=Z1Ki8BMXJRIC>
12. Creswell, J. W., & Plano Clark, V. L. (2017). *Designing and Conducting Mixed Methods Research*. SAGE Publications.
13. Daveson, A. J. M., Craig, E., Vitak, A., Ware, R. S., Schafer, J., Sehgal, A., Bose, U., Colgrave, M. J., Hardy, M. Y., Tye-Din, J. A., & Anderson, R. P. (2026). A Randomized Double-Blind, Placebo-Controlled Dose-Response Study to Assess the Gluten Threshold Dose in Celiac Disease. *Gastroenterology*. <https://doi.org/10.1053/j.gastro.2026.03.011>
14. Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L., & Lancaster, G. A. (2016). CONSORT 2010 statement: extension to randomised pilot and feasibility trials. *BMJ*, 355, i5239. <https://doi.org/10.1136/bmj.i5239>
15. Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative Content Analysis: A Focus on Trustworthiness. *Sage Open*, 4(1), 2158244014522633. <https://doi.org/10.1177/2158244014522633>
16. Elo, S., & Kyngas, H. (2008). The qualitative content analysis process. *J Adv Nurs*, 62(1), 107–115. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>
17. Fetter, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: principles and practices. *Health Serv Res*, 48(6 Pt 2), 2134–2156. <https://doi.org/10.1111/1475-6773.12117>
18. Food & Agriculture Organization of the United Nations, & World Health Organization. (2025). *Summary report of the ad hoc Joint FAO/WHO Expert Consultation on Risk Assessment of Food Allergens: Reference dose(s) for cereals containing gluten or gluten, Rome, 3–7 November 2025*. <https://openknowledge.fao.org/items/0a455421-491b-4c44-a786-83c51452bdef>
19. Friedberg, J. P., Lipsitz, S. R., & Natarajan, S. (2010). Challenges and recommendations for blinding in behavioral interventions illustrated using a case study of a behavioral intervention to lower blood pressure. *Patient Educ Couns*, 78(1), 5–11. <https://doi.org/10.1016/j.pec.2009.04.009>
20. Gholmie, Y., Lee, A. R., Satherley, R. M., Schebendach, J., Zybert, P., Green, P. H. R., Lebwohl, B., & Wolf, R. (2023). Maladaptive Food Attitudes and Behaviors in Individuals with Celiac Disease and Their Association with Quality of Life [Article]. *Digestive Diseases and Sciences*, 68(7), 2899–2907. <https://doi.org/10.1007/s10620-023-07912-6>
21. Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, 24(2), 105–112. <https://doi.org/https://doi.org/10.1016/j.nedt.2003.10.001>

22. Johnson, R. B. (2009). Comments on Howe: Toward a More Inclusive “Scientific Research in Education”. *Educational Researcher*, 38(6), 449–457. <https://doi.org/10.3102/0013189x09344429>
23. Johnson, R. B. (2017). Dialectical Pluralism: A Metaparadigm Whose Time Has Come. *Journal of Mixed Methods Research*, 11(2), 156–173. <https://doi.org/10.1177/1558689815607692>
24. Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14–26. <https://doi.org/10.3102/0013189x033007014>
25. Julious, S. A. (2005). Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceutical Statistics*, 4(4), 287–291. <https://doi.org/10.1002/pst.185>
26. Kien, C., Griebler, U., Schultes, M.-T., Thaler, K. J., & Stamm, T. (2021). Psychometric Testing of the German Versions of Three Implementation Outcome Measures. *Global Implementation Research and Applications*, 1(3), 183–194. <https://doi.org/10.1007/s43477-021-00019-y>
27. Lancaster, G. A., & Thabane, L. (2019). Guidelines for reporting non-randomised pilot and feasibility studies. *Pilot and Feasibility Studies*, 5(1), 114. <https://doi.org/10.1186/s40814-019-0499-1>
28. Lebwohl, B., Ma, C., Lagana, S. M., Pai, R. K., Baker, K. A., Zayadi, A., Hogan, M., Bouma, G., Cellier, C., Goldsmith, J. D., Lundin, K. E. A., Pinto-Sanchez, M. I., Robert, M. E., Rubio-Tapia, A., Sanders, D. S., Schaeffer, D. F., Semrad, C. E., Silvester, J. A., Verdu, E. F.,... Murray, J. A. (2024). Standardizing Randomized Controlled Trials in Celiac Disease: An International Multidisciplinary Appropriateness Study. *Gastroenterology*, 166(1), 88–102. <https://doi.org/10.1053/j.gastro.2023.08.051>
29. Lebwohl, B., Sanders, D. S., & Green, P. H. R. (2018). Coeliac disease [Review]. *The Lancet*, 391(10115), 70–81. [https://doi.org/10.1016/S0140-6736\(17\)31796-8](https://doi.org/10.1016/S0140-6736(17)31796-8)
30. Lee, E. C., Whitehead, A. L., Jacques, R. M., & Julious, S. A. (2014). The statistical interpretation of pilot trials: should significance thresholds be reconsidered? *BMC Med Res Methodol*, 14, 41. <https://doi.org/10.1186/1471-2288-14-41>
31. Leffler, D. A., Dennis, M., Edwards George, J., Jamma, S., Cook, E. F., Schuppan, D., & Kelly, C. P. (2009). A validated disease-specific symptom index for adults with celiac disease. *Clin Gastroenterol Hepatol*, 7(12), 1328–1334, 1334.e1321-1323. <https://doi.org/10.1016/j.cgh.2009.07.031>
32. Ludvigsson, J. F., Bai, J. C., Biagi, F., Card, T. R., Ciacci, C., Ciclitira, P. J., Green, P. H., Hadjivassiliou, M., Holdaway, A., van Heel, D. A., Kaukinen, K., Leffler, D. A., Leonard, J. N., Lundin, K. E., McGough, N., Davidson, M., Murray, J. A., Swift, G. L., Walker, M. M.,... British Society of, G. (2014). Diagnosis and management of adult coeliac disease: guidelines from the British Society of Gastroenterology. *Gut*, 63(8), 1210–1228. <https://doi.org/10.1136/gutjnl-2013-306578>
33. Ludvigsson, J. F., Leffler, D. A., Bai, J. C., Biagi, F., Fasano, A., Green, P. H., Hadjivassiliou, M., Kaukinen, K., Kelly, C. P., Leonard, J. N., Lundin, K. E., Murray, J. A., Sanders, D. S., Walker, M. M., Zingone, F., & Ciacci, C. (2013). The Oslo definitions for coeliac disease and related terms. *Gut*, 62(1), 43–52. <https://doi.org/10.1136/gutjnl-2011-301346>
34. Mayes, C., Williams, J., & Sassano, A. (2026). Normal food practices negotiating the interaction among norms practices and food environments. *Health Sociology Review*.

<https://doi.org/10.1080/14461242.2026.2625708>

35. Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350, h1258. <https://doi.org/10.1136/bmj.h1258>
36. Morgan, D. L. (2007). Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research*, 1(1), 48–76. <https://doi.org/10.1177/2345678906292462>
37. O'Cathain, A., Murphy, E., & Nicholl, J. (2008). The quality of mixed methods studies in health services research. *J Health Serv Res Policy*, 13(2), 92–98. <https://doi.org/10.1258/jhsrp.2007.007074>
38. Parchman, M. L., Henrikson, N. B., Blasi, P. R., Buist, D. S., Penfold, R., Austin, B., & Ganos, E. H. (2017). Taking action on overuse: Creating the culture for change. *Healthc (Amst)*, 5(4), 199–203. <https://doi.org/10.1016/j.hjdsi.2016.10.005>
39. Proctor, E., Silmere, H., Raghavan, R., Hovmand, P., Aarons, G., Bunger, A., Griffey, R., & Hensley, M. (2011). Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Adm Policy Ment Health*, 38(2), 65–76. <https://doi.org/10.1007/s10488-010-0319-7>
40. Rasmussen, S. N., Vigre, H., & Madsen, C. B. (2022). Risk of consuming products with or without precautionary wheat or gluten labelling for persons with coeliac disease. *Food and chemical toxicology: an international journal published for the British Industrial Biological Research Association*, 166, 113231. <https://doi.org/https://dx.doi.org/10.1016/j.fct.2022.113231>
41. Rostami-Nejad, M., Asri, N., Olfatifar, M., Khorsand, B., Hour, H., & Rostami, K. (2023). Systematic Review and Dose-Response Meta-Analysis on the Relationship between Different Gluten Doses and Risk of Coeliac Disease Relapse. *Nutrients*, 15(6). <https://doi.org/10.3390/nu15061390>
42. Rubio-Tapia, A., Hill, I. D., Semrad, C., Kelly, C. P., & Lebwohl, B. (2023). American College of Gastroenterology Guidelines Update: Diagnosis and Management of Celiac Disease [Article]. *American Journal of Gastroenterology*, 118(1), 59–76. <https://doi.org/10.14309/ajg.0000000000002075>
43. Satherley, R. M., Higgs, S., & Howard, R. (2017). Disordered eating patterns in coeliac disease: a framework analysis. *J Hum Nutr Diet*, 30(6), 724–736. <https://doi.org/10.1111/jhn.12475>
44. Satherley, R. M., Lerigo, F., Higgs, S., & Howard, R. (2022). An interpretative phenomenological analysis of the development and maintenance of gluten-related distress and unhelpful eating and lifestyle patterns in coeliac disease. *Br J Health Psychol*, 27(3), 1026–1042. <https://doi.org/10.1111/bjhp.12588>
45. Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant*, 52(4), 1893–1907. <https://doi.org/10.1007/s11135-017-0574-8>

46. Schatzki, T. R. (2010). *Site of the Social: A Philosophical Account of the Constitution of Social Life and Change*. Pennsylvania State University Press.
47. Schatzki, T. R., Knorr-Cetina, K., & von Savigny, E. (2001). *The Practice Turn in Contemporary Theory*. Routledge.
48. Sekhon, M., Cartwright, M., & Francis, J. J. (2017). Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res*, 17(1), 88. <https://doi.org/10.1186/s12913-017-2031-8>
49. Simon, E., Molero-Luis, M., Fueyo-Diaz, R., Costas-Batlle, C., Crespo-Escobar, P., & Montoro-Huguet, M. A. (2023). The Gluten-Free Diet for Celiac Disease: Critical Insights to Better Understand Clinical Outcomes. *Nutrients*, 15(18). <https://doi.org/10.3390/nu15184013>
50. Singh, P., Arora, A., Strand, T. A., Leffler, D. A., Catassi, C., Green, P. H., Kelly, C. P., Ahuja, V., & Makharia, G. K. (2018). Global Prevalence of Celiac Disease: Systematic Review and Meta-analysis. *Clin Gastroenterol Hepatol*, 16(6), 823–836 e822. <https://doi.org/10.1016/j.cgh.2017.06.037>
51. Skivington, K., Matthews, L., Simpson, S. A., Craig, P., Baird, J., Blazeby, J. M., Boyd, K. A., Craig, N., French, D. P., McIntosh, E., Petticrew, M., Rycroft-Malone, J., White, M., & Moore, L. (2021). A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. *BMJ*, 374, n2061. <https://doi.org/10.1136/bmj.n2061>
52. Smith, J. D., Li, D. H., & Rafferty, M. R. (2020). The Implementation Research Logic Model: a method for planning, executing, reporting, and synthesizing implementation projects. *Implement Sci*, 15(1), 84. <https://doi.org/10.1186/s13012-020-01041-8>
53. Staniszewska, S., Brett, J., Simera, I., Seers, K., Mockford, C., Goodlad, S., Altman, D. G., Moher, D., Barber, R., Denegri, S., Entwistle, A., Littlejohns, P., Morris, C., Suleman, R., Thomas, V., & Tysall, C. (2017). GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *Res Involv Engagem*, 3, 13. <https://doi.org/10.1186/s40900-017-0062-2>
54. Streiner, D. L., Norman, G. R., & Cairney, J. (2014). *Health Measurement Scales: A practical guide to their development and use*. Oxford University Press. <https://doi.org/10.1093/med/9780199685219.001.0001>
55. Studerus, D., Lee, A. R., Hugo, T., Heim, P., Jossen, J., Scharl, M., & Zeitz, J. (2025). Understanding Cross-Contamination in a Gluten-Free Diet: A Scoping Review. *Clinical Nutrition ESPEN*. <https://doi.org/10.1016/j.clnesp.2025.102847>
56. Taavela, J., Koskinen, O., Huhtala, H., Lahdeaho, M. L., Popp, A., Laurila, K., Collin, P., Kaukinen, K., Kurppa, K., & Maki, M. (2013). Validation of morphometric analyses of small-intestinal biopsy readouts in celiac disease. *PLOS ONE*, 8(10), e76163. <https://doi.org/10.1371/journal.pone.0076163>
57. Tashakkori, A., & Teddlie, C. (2010). *SAGE Handbook of Mixed Methods in Social & Behavioral Research*. SAGE Publications. <https://books.google.ch/books?id=VU1uQ8FNKLYC>
58. Teare, M. D., Dimairo, M., Shephard, N., Hayman, A., Whitehead, A., & Walters, S. J. (2014). Sample size requirements to estimate key design parameters from external pilot randomised controlled trials: a simulation study. *Trials*, 15(1), 264. <https://doi.org/10.1186/1745-6215-15-264>

59. Timmermans, S., & Tavory, I. (2012). Theory Construction in Qualitative Research: From Grounded Theory to Abductive Analysis. *Sociological Theory*, 30(3), 167–186. <https://doi.org/10.1177/0735275112457914>
60. Vargas, F. M., Cardoso, L. T., Didone, A., Lima, J. P. M., Venzke, J. G., & de Oliveira, V. R. (2024). Celiac Disease: Risks of Cross-Contamination and Strategies for Gluten Removal in Food Environments. *Int J Environ Res Public Health*, 21(2). <https://doi.org/10.3390/ijerph21020124>
61. Warde, A. (2005). Consumption and Theories of Practice. *Journal of Consumer Culture*, 5(2), 131–153. <https://doi.org/10.1177/1469540505053090>
62. Warde, A. (2016). *The Practice of Eating*. Polity Press. <https://books.google.ch/books?id=PJJsCwAAQBAJ>
63. Weiner, B. J., Lewis, C. C., Stanick, C., Powell, B. J., Dorsey, C. N., Clary, A. S., Boynton, M. H., & Halko, H. (2017). Psychometric assessment of three newly developed implementation outcome measures. *Implement Sci*, 12(1), 108. <https://doi.org/10.1186/s13012-017-0635-3>
64. Wolf, R. L., Lebwohl, B., Lee, A. R., Zybert, P., Reilly, N. R., Cadenhead, J., Amengual, C., & Green, P. H. R. (2018). Hypervigilance to a Gluten-Free Diet and Decreased Quality of Life in Teenagers and Adults with Celiac Disease. *Dig Dis Sci*, 63(6), 1438–1448. <https://doi.org/10.1007/s10620-018-4936-4>

Table

Table 1 Feasibility outcomes, assessment timepoints, and a priori progression criteria

Domain	Assessment	Criterion / reported metric
Recruitment	Binary	≥ 24 of the target 30 participants enrolled by 31 October 2026, corresponding to a mean enrolment rate of ≥ 6 participants/month over the mid-July to end-October window.
Safety (individual safety switches)	Binary	>2 individual safety switches constitutes a study-level stop criterion ; ≤2 permits continuation (<i>defined in Randomization, blinding, and safety switch</i>)
Retention / paired-biopsy completion	Descriptive	Proportion completing the 24-week protocol; proportion with paired baseline and week-24 biopsies (the completion rate critical for a Vh:Cd-based definitive trial)
Adherence	Descriptive	Estimated exposure within the intended range: • for the intervention arm, estimated intake within the contamination range with fecal GIP below the limit of quantification; • for the comparator arm, GIP non-detectable and fecal anti-htTG sIgA <100 U/L
Acceptance	Descriptive	AIM scores (anticipated at baseline; experienced at weeks 6–24)
Data completeness	Descriptive	Proportion of expected outcome data captured (REDCap)

Progression criteria were defined a priori for each feasibility domain. In line with current guidance for pilot and feasibility trials, they are intended to inform the design of a definitive trial rather than to serve as automatic stop/go rules: where a criterion is not met, the specified adaptation is considered before concluding that a definitive trial is not warranted. The safety domain is the exception and additionally functions as a predefined ethical stopping rule. Assessment timepoints correspond to the study visits shown in Fig. 1.

Abbreviations: GIP gluten immunogenic peptides, htTG sIgA human tissue transglutaminase secretory immunoglobulin A, REDCap Research Electronic Data Capture, Vh:Cd villous height-to-crypt depth ratio

Figures

	TRIAL PERIOD					
	Enrollment		Post-randomization			Close-out
TIMEPOINT	-t to 0 Screening ^a	0	t ₁ Week 6	t ₂ Week 12	t ₃ Week 18	t ₄ Week 24
ENROLLMENT:						
Eligibility screen ^b	X					
Informed consent	X					
Serologic eligibility (tTG-IgA in reference range)	X					
Histologic eligibility (Vh:Cd ≥2.0)		X				
Randomization		X				
INTERVENTION						
rGFD arm ^c		X	→			X
Optimized GFD ^c		X	→			X
ASSESSMENTS:						
Upper endoscopy & duodenal biopsy — Vh:Cd		X				X
Marsh-Oberhuber grade & IEL count		X				X
Blood: tTG-IgA (safety)		X*	X	X	X	X
Stool: GIP & anti-htTG slgA		X	X	X	X	X
PROs ^d		X*	X	X	X	X
Acceptability (AIM) ^e		X	X	X	X	X
Estimated gluten intake ^f			X	X	X	X
Semi-structured interview ^g		X	X	X	X	X
Safety-switch evaluation (tTG-IgA >10 U/mL) ^{h,i}			X	X	X	X
Adverse-event assessment			X	X	X	X

Figure 1

Participant timeline: Schedule of enrollment, interventions, and assessments

* repeated at V0 only if >2 weeks have elapsed since the screening draw; otherwise the screening value is carried forward as baseline

- +** Individual switch to the control arm if tTG-IgA >10 U/mL from week 6; study-level stop criterion if >2 switches ($\approx$ >13% of the intervention group). At t4, tTG-IgA elevation triggers continued safety follow-up rather than a switch
- b** Eligibility screen = (CeD diagnosis, medical history, demographics), Symptom screen for exclusion (CSI >30), Physical examination, Pregnancy test (childbearing potential) – all performed after signing the informed consent as study specific procedures
- c** Educational material (group-specific video + handout)
- d** Celiac Symptom Index + VAS 0-100 mm for fr-QoL, gluten anxiety, risk willingness
- e** 0 captures anticipated acceptability (rated before intervention onset); T1–T4 capture experienced acceptability. Reported and interpreted as distinct constructs on a common scale (Sekhon et al. 2017; Weiner et al. 2017), not pooled as a single repeated measure
- f** focussed food diary (weekly, at home), estimation based on our previous work
- g** All 30 participants at V0 and T1; from T2 (week 12) a theoretically sampled subset (~10–16 per timepoint) for in-depth interviews. Interview guide held constant across timepoints for comparability.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Add1ImplementationResearchLogicModel.pdf](#)
- [Add2InclusionExclusionCriteria.docx](#)
- [Add3SPIRITChecklist.docx](#)
- [Add4GRAMMSChecklist.docx](#)