TRIAL FULL TITLE	GLUTENSCREEN Early diagnosis of coeliac disease in the Preventive Youth Health Care Centres in the Netherlands
EUDRACT NUMBER	n.a.
SAP VERSION	1.0
ISRCTN NUMBER	n.a.
SAP VERSION DATE	11 May 2021
TRIAL STATISTICIAN	Dr. N. van Geloven
TRIAL CHIEF INVESTIGATOR	Prof. dr. M. L. Mearin
SAP AUTHOR	M. L. Mearin, C. R. Meijer, N. van Geloven

1. SAP Signatures

I give my approval for the attached SAP entitled "GLUTENSCREEN. Early diagnosis of coeliac disease in the Preventive Youth Health Care Centres in the Netherlands" dated April 22, 2021.

Chief Investigator

Name: Prof. dr. M.L. Mearin

Signature: _____

Date: _____

Statistician

Name: Dr. N. van Geloven

Signature:

Date: _____

2. Abbreviations and Definitions

CD	Celiac disease
DPSU	Dutch Paediatric Surveillance Unit
ESPGHAN	European Society of Paediatric Gastroenterology, Hepatology and Nutrition
GFD	Gluten free diet
HLA	Human Leukocyte Antigen
LUMC	Leiden University Medical Centre
GLUTENSCREEN	Early diagnosis of coeliac disease in the Preventive Youth Health Care Centres in the Netherlands; project funded by ZonMW Proposal/Contract no: 531002001
POC	Point of care test

SAP	Statistical analysis plan
tTGA	Anti-tissue transglutaminase antibodies
YHCCs	Youth Health Care Centres

3. Introduction

Coeliac disease (CD) is an immune-mediated systemic disorder elicited by the ingestion of gluten containing cereals (among others wheat, rye and barley) from the normal diet in genetically susceptible individuals (Husby 2012, Husby 2020). CD is treated with a gluten-free diet (GFD). CD is the most common immune-mediated food-related disease in Europe and beyond and affects as many as 1% of the general population (Lindfors 2019). This means that in The Netherlands there are approximately 150.000 CD cases, from whom 33.000 children, but most of them are undiagnosed and thus untreated (Mearin oratie). In the Netherlands for every child diagnosed with CD, there are seven who have unrecognized, and therefore, untreated disease (George 1995; Steens 2015; Jansen 2018; Meijer 2021. Submitted). This is partially due to the variable clinical presentation and symptoms, including asymptomatic CD cases. Untreated disease is associated with long-term complications, such as growth restriction, delayed puberty, neuropsychiatric disturbances, associated autoimmune disease, miscarriages, small-for-date-births, osteoporosis, and, rarely, malignancy (Jansen 2015, Kiefte-de Jonge 2013, Lindfors 2019). Untreated CD increases the overall mortality risk, reduces the quality of life and yields extensive negative economic consequences, thereby presenting a resource challenge for current and future health systems (Lindfors 2019). The increasing incidence of CD over the last half-century has resulted in rising interest in identifying preventive strategies (Meijer 2018). Recent prospective studies show that CD develops very early in life and that treatment of CD patients detected by early diagnosis results in health improvement (Vriezinga 2014; Meijer *In preparation*). The current standard health care is unable to solve the problem of underdiagnose of CD and secondary prevention by early diagnosis and treatment may only be achieved on a large scale by mass screening or by early and active case-finding (Meijer 2018).

The general objective of GLUTENSCREEN is to perform a novel case-finding project to detect CD in 12 months-4 years old symptomatic children who visit the Young Health Care Centres (YHCCs) in the well-described region of Kennemerland, The Netherlands, using a rapid point of care test (POC) for determination of CD specific antibodies against the enzyme tissue transglutaminase (TGA). The aim of GLUTENSCREEN is to show that case-finding of CD in young children attending the YHCCs is feasible, efficient, cost-effective and well accepted by the population. To achieve this GLUTENSCREEN will compare the results of the case-finding strategy to the outcome of current healthcare in the diagnosis of CD in children.

DESIGN. Prospective intervention cohort study.

Subjects: All children aged 12 months-4 years attending scheduled visits to the YHCCs in the Kennemerland region. At the YHCC a questionnaire on CD-related symptoms will be checked (annex 1). If one or more CD-associated symptoms from Annex 1 are present or if the YHCC signals growth restriction, the child is eligible for the POC test. At that moment the child is considered part of the case finding population, see block 4 in the Figure.

Control population

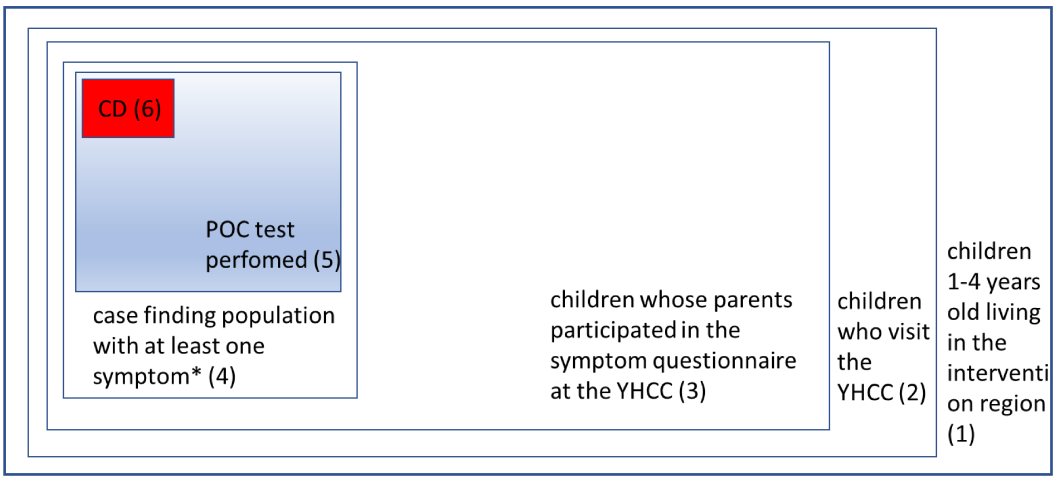
A national control group is based on the data reported by Dutch Paediatric Surveillance Unit (DPSU), with CD diagnosed according to the current standard of care. DPSU is a unique registry of the Dutch Society of Paediatrics, comprising of all Dutch paediatric practices. Under it, Dutch paediatricians are asked to report newly diagnosed cases of certain diseases (CD in our case) monthly during the time of this case-finding project and the year after this project. DPSU respondents have a 90% mean response rate. Cases reported from the study region of Kennemerland are not included in the control group. The cases of clinically diagnosed CD in the study region will be identified by the data of the YHCC, so not through the DPSU.

SAMPLE SIZE

Sample size calculation was based on the overall crude incidence rate of clinically diagnosed CD of 1.56/1000 live births and in an estimated incidence rate of 0.62/1000 child years in children aged 12 months-4 years (Csizmadia 1999; Steens 2005, Jansen 2018). Per 1000 children followed for an average of 2,5 years, we expect 1.56 cases of clinically diagnosed CD. We assume that in the Dutch population outside the case-finding project, the incidence of children with a diagnosis of CD equals 0.62/1000 children years and that in the case-finding population the incidence of children that will be detected (incidence rate) will be at least 8 times as high. With 2.5 years follow-up, 5434 subjects followed for case-finding are sufficient to detect an incidence rate ratio of 8, with 80% power, using an alpha level of 0.05. We expected 60% of the children to be symptomatic, and 60% participation of those symptomatic children in the case-finding population, so 15,100 children needed to be requested for participation, in order to obtain 5434 children available for case-finding using a rapid POC test. Since the population in the YHCCs in Kennemerland is approximately 12,000 children/year with additional 4,000 added per year and 2.5 years of follow-up will be sufficient to achieve the sample size. CD will be diagnosed after positive POC test following accepted guidelines.

Figure Design

**including growth restriction. In addition, subjects should meet in- and exclusion criteria.*



INTERVENTION:

After informed consent the validated POC test to determine CD specific antibodies against the enzyme tissue transglutaminase2 (TG2A, Celiac Quick Test; BioHit Oyj, Finland) which is also suitable for Immunoglobulin A (IgA)- deficient patients, will be performed. If the POC test is positive, the child will be referred to the paediatrician-gastroenterologist at the Leiden University Medical Centre (LUMC) for definitive diagnosis of CD according to the guideline of the European Society for Pediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN) (Husby 2012/2020).

OUTCOMES:

1. The incidence rate of new CD diagnoses in the study region Kennemerland (block 1 in the Figure) in comparison to the rest of the Netherlands

2. Number (percentage) of children with a positive POC test in the tested children (block 5 in the Figure).

3. Number (percentage) of children with a positive POC test in which the diagnosis of CD is confirmed.

4. Feasibility and cost-effectiveness of active case-finding for CD in the YHCCs. All costs of active case-finding, diagnostics and treatment of CD and the potential short and long term consequences of the disease will be calculated for the settings with and without active case-finding. Healthcare use will be valued according to the Dutch guideline for costing research using patient's questionnaires and data from the literature.

5. Ethical acceptability: by questionnaires on parental satisfaction and focus groups.

This statistical analysis plan (SAP) refers to OUTCOMES 1, 2 and 3:

4. Purpose of the analyses

The primary purpose of the analysis is to study the incidence rate of new CD diagnoses in the study region (block 1 in the Figure) and compare it with the incidence rate in children 1-4 years in the rest of the Netherlands, obtained from the DPSU in the as much as possible the same period (DPSU registration ended April 2021).

Secondary purpose is to detect variables that are related to the early development of CD in children visiting the YHCCs. These variables will be studied by comparing the children with a new CD diagnosis in both the study region and the rest of the Netherlands with the children who had a negative POC test.

5. Study Objectives and Endpoints

5.1 Study Objectives

The primary objectives of this study are to establish the feasibility, effectiveness and costs of early diagnosis of CD by active case-finding in children attending the YHCCs in a well described study region in the Netherlands, using a POC test for TG2A determination.

In this SAP we focus on the effectiveness aspect. In particular the objective is to estimate the incidence rate of new CD diagnoses in the study region and compare it with the incidence rate in rest of the Netherlands, obtained from DPSU in the same period.

Secondary objective is to detect variables that are associated with the early development of CD in children visiting the YHCCs with respect to the background information and symptoms as provided by the YHCCs, both for the detected CD cases as (codified/anonymised) for the children with a negative case-finding (negative POC test) as presented in annex 2. Taking into account the information from the literature (Lindfors 2019) and our previous results from the PreventCD cohort (Vriezinga 2014, Auricchio 2020, Meijer *In preparation*) the risk of developing CD is expected to differ according to:

a. Gender
b. Age
c. Positive family history of CD
d. Growth restriction based on standardized weight and height, defined as reduction of >0.25 SD per year in height and/or weight/age or a reduction >1 SD per year for weight/height
e. Symptoms as specified in annex 1

5.2 Endpoints

The primary aim of this analysis in GLUTENSCREEN is to establish the incidence rate of new CD diagnoses in the study region and to compare it with the incidence rate of CD diagnosed by the standard of care in the rest of the Netherlands, as obtained from DPSU in the same period.

The primary endpoint of this analysis is the incidence rate of new CD diagnoses by

- Case-finding
- Standard of care

CD diagnosis is established according to the criteria of ESPGHAN (Husby 2012/2020).

The age at diagnosis of CD (see below for precise definition) is defined
for the case-finding population as
 Date of first immunoglobulin A (IgA) TGA >7x normal value at the LUMC or date of diagnostic small bowel biopsies
For the standard of care population (DPSU) as
 Date of diagnosis of CD as reported by the paediatrician to the DPSU. As the exact date of diagnosis is not recorded in the DPSU, it will be estimated by averaging the date of first presentation at the clinic and the date of reporting of the CD case to the DPSU.

6. Study Methods

6.1 General Study Setup

Study population (block 1 in the Figure)
All children aged 12 months-4 years living in the study region Kennemerland during the timing of the study. These children are invited to regular visits to the YHCCs.

Case finding population (block 4 in the Figure)
The vast majority of children from the study population attend the visits at the YHCC. At the YHCCs a questionnaire on CD-related symptoms will be checked (annex 1). If one or more CD-associated symptoms are presented or when the YHCC registers growth restriction, the child is eligible for the POC test and thereby is included in the case finding population if they meet the following in- and exclusion criteria.

Inclusion-Exclusion Criteria
In order to be eligible for the case finding population, a subject must meet all of the following criteria: age 12 months to 4 years, not diagnosed with CD, not on a GFD, one or more CD-associated symptoms (annex 1) or growth restriction, parents have a sufficient knowledge of Dutch language, informed consent.

7 General Considerations

7.1 Timing of Analyses

Data for this analysis will be frozen after all the children have been tested in the YHCCs as by protocol, this is, during 2,5 years starting February 2019 (first child included), plus additional moths lost by the Corona crisis (March-August 2020) and 3 extra months for CD diagnosis at the LUMC and internal data control.

7.2 Analysis Populations

Since response rate may vary by province we will also separate describe the incidence rate of the control region with highest response and compare the incidence in the study region to that high incidence control region.

7.3 Subgroup analyses
The risk of developing CD is expected to differ between:

- Case-finding v. s. standard of care
- Gender
- Age
- Family history of CD
- Related disease history (e.g. diabetes type I, full list presented in annex 2)
- All symptoms in annex 1 (including growth restriction (in weight and/or height))

8 Statistical analyses

8.1 Primary endpoint: the incidence rate of new CD diagnoses in the study region Kennemerland (block 1 in the figure) in comparison to the rest of the Netherlands

The primary analysis of this study will be a comparison between the incidence rate of new CD diagnoses in the study region during the time of the study in comparison to the rest of the Netherlands.

The incidence rate in the study region will be calculated dividing the total number of definite diagnoses found by the total number of children years aged 1-4 living in the region according to Statistics Netherlands. Cases found outside of the case-finding will be included in calculating this rate.

The incidence rate in the rest of the Netherlands will be calculated using the number of definite diagnoses found in the DSUP registration from children not living in Kennemerland at the time of diagnosis. This number will be divided by the total number of child years aged 1-4 from children living in that region during the period of DSUP registration according to Statistics Netherlands.

The estimated incidence rate in the study region will be reported along with a 95% two sided confidence interval and a test will be performed comparing this rate with the rate in the rest of the Netherlands assuming the latter has no sampling variability (so using the incidence rate in the rest of the Netherlands as a fixed reference value).

<u>Supplementary analyses for the primary endpoint:</u>

- a similar comparison will be made with the region with highest reported incidence in DPSU

- we will split up the incidence found in the study region according to

- age of the child (1,2, 3, 4 years old) in comparison to incidence rate in same age groups in the rest of the Netherlands. For the denominator we will use (estimated) age specific number of child years in the study region and in the rest of the Netherlands.

- whether the case was found during a first time the child participated in the case finding and received the POC or during a second or later time. For the denominator we will use the total number of child years 1-3 years old in the study region.

- Calendar period. We will report separately the incidence found from start of study to corona stop (March 19th – Aug 15th 2020) and after that (Aug 15th 2020-end of study). The denominator in these separate periods will be calculated by scaling the number of child years proportionally to the calendar time. The primary analysis will include all periods together, including the corona stop period where no children were invited to participate. We chose this approach as also in the rest of the Netherlands due to the corona situation, fewer diagnosis than usual are expected during the total study period.

8.2 Number (percentage) of children with a positive POC test in the tested children (block 5 in the figure).

This will be calculated as the number of positive POC tests found within the study divided by the number of performed POC tests. This percentage will be reported along with a 2 sided 95% confidence interval.

8.3 Number (percentage) of children with a positive POC test in which the diagnosis of CD is confirmed.

This will be calculated as the number of children with a definite diagnosis out of the number of children with positive POC test. This percentage will be reported along with a 2 sided 95% confidence interval.

8.4 Subgroup analyses

To study whether the risk of developing CD is expected to differ between Gender, Age, Family history of CD, related diseases, growth restriction (in weight and/or height) and all symptoms in annex 1, we will use generalized (logistic) linear models applying generalized estimation equations to adjust inference for the correlation between observations from patients who participated multiple times. We will express the predictive value of the mentioned factors as odds ratios together with 95% confidence intervals, both from univariable analysis and from one multivariable analysis that includes all factors. For this multivariable model we will use derived variables (e.g. an indicator whether any symptom or any related disease was present) to limit the number of parameters that need to be estimated. In an exploratory analysis the correlation between the risk of developing CD and vaccination for Rotavirus will be analysed.

9 Changes during the study

9.1 Re-evaluation of sample size during preparation of SAP

During preparation of the SAP we made the comparison groups explicit (Figure). We then noted that the original sample size calculation (Section 3) was not targeted at the primary comparison groups (see Section 8.1). We therefore re-formulated the power calculation in light of the primary planned analysis. Note that we kept the sample size as originally planned.

New formulation of power calculation: We assume that in the Dutch population outside the case-finding project, the incidence of children 1-4 years old with a diagnosis of CD equals 0.62/1000 children years. With 2.5 years inclusion period, we expected 5434 children taking the POC test would give high power (about 95%) to detect an at least two times higher incidence rate in the study region

(alpha 5%) . We expected 60% of the children to be symptomatic, and 60% participation of those symptomatic children in the POC testing, so 15,100 children would need to be requested for participation, in order to obtain 5434 children available for case-finding using a rapid POC test. Since the population in the YHCCs in the Kennemerland region is approximately 12,000 children/year with additional 4,000 added per year and 2.5 years of study duration was considered sufficient to achieve sufficient sample size.

9.2 Re-evaluation of sample size due to COVID pandemic

After in March 2020 the study had to be interrupted for 5 months due to the COVID pandemic, the sample size calculation was re-evaluated based on the results up to that moment, including the number of CD cases found in the study region in the first year of the study. Based on this evaluation, it was decided that the original inclusion period of 2.5 years could be retained.

10 References

1. Meijer CR, van den Akker ME, van Bodegom L, Escher JC, van Overveld F, Rings EHHM, Smit L, de Vries M, Mearin ML, Early diagnosis of coeliac disease in the Preventive Youth Health Care Centres in the Netherlands: study protocol of a case-finding study (GLUTENSCREEN). In preparation.
2. Husby S et al. ESPGHAN guidelines for the diagnosis of coeliac disease. JPGN 2012:54:136-60.
3. Husby S, Koletzko S, Korponay-Szabó I et al. European Society Paediatric Gastroenterology, Hepatology and Nutrition Guidelines for Diagnosing Coeliac Disease 2020. J Pediatr Gastroenterol Nutr. 70(1):141-156.
4. Lindfors K, Ciacci C, Kurppa K, Lundin K, Makharia G, Mearin ML, Murray JA, Verdu EF, Kaukinen K. Coeliac disease, Nat Rev Dis Primers. 2019 Jan 10;5(1):3.
5. George EK, Mearin ML, van der Velde EA, Houwen RH, Bouquet J, Gijsbers CF, et al (1995) Low incidence of childhood celiac disease in The Netherlands. Pediatr Res 37(2):213-8
6. Steens RF, Csizmadia CG, George EK, Ninaber MK, Hira Sing RA, Mearin ML (2005) A national prospective study on childhood celiac disease in the Netherlands 1993-2000: an increasing recognition and a changing clinical picture. J Pediatr 147(2):239-43.
7. Jansen M, van Zelm M, Groeneweg M, Jaddoe V, Dik W, et al. The identification of celiac disease in asymptomatic children: the Generation R Study. J Gastroenterol. 2018: 53(3):377-386
8. Meijer CR. Et al, Efficient implementation of the 'non-biopsy approach' for the diagnosis of childhood celiac disease in the Netherlands. a national prospective evaluation 2010-2013. submitted 2021
9. Jansen MA et al. Growth Trajectories and Bone Mineral Density in Anti-Tissue Transglutaminase Antibody-positive Children: The Generation R Study. Clin Gastroenterol Hepatol 2015;13(5):913-20.
10. Kiefte-de Jong JC et al. Levels of antibodies against tissue transglutaminase during pregnancy are associated with reduced fetal weight and birth weight. Gastroenterology 2013;144(4):726-35.
11. Meijer CR et al. Celiac disease Prevention. Frontiers in pediatrics. 2018
12. Vriezinga SL et al. Randomized Feeding Intervention in Infants at High Risk for Celiac Disease. NEJM 2014;371:1304-15.
13. Csizmadia CGDS et al. An iceberg of childhood coeliac disease in the Netherlands. Lancet 1999;353:813-4.
14. Auricchio R, Stellato P, Bruzzese D, Cielo D, Chiurazzi A, Galatola M, Castilljeo G, Crespo Escobar P, Gyimesi J, Hartman C, Kolacek S, Koletzko S, Korponay-Szabo I, Mearin ML, MeijerC, Pieścik-

Lech M, Polanco I, Ribes-Koninckx C, Shamir R, Szajewska H, Troncone R, Greco L. Growth rate of coeliac children is compromised before the onset of the disease. Arch Dis Child 2020;105:964–968

15. Meijer C, Verburgt C, Auricchio R, Castillejo G, Crespo Escobar P, Gyimesi J, Hartman C, Kolacek S, Koletzko S, Korponay-Szabo I, Martinez Ojinaga Nodal E, Pieścik-Lech M, Polanco I, Ribes Koninckx C, Shamir R, Szajewska H, Szillat P, Troncone R, Werkstetter K, Putter H, Mearin ML. Children from coeliac families benefit from early diagnosis and treatment: an analysis of the PREVENTCD cohort. In preparation.

Annex 1

Questions Coeliac Disease (to fill in by parent(s))

1.	Does your child suffer from abdominal pain longer than 3 weeks (at least twice a week)?	Yes / No
2.	Does your child have abdominal bloating?	Yes / No
3.	Is your child regularly constipated and not responsive to laxatives?	Yes / No
4.	Does your child have diarrhoea longer than 2 weeks?	Yes / No
5.	Does your child suffer from vomiting longer than 3 weeks (at least twice a week)?	Yes / No
6.	Do you find your child easily tired in a way he/she is hindered in daily activities?	Yes / No
7.	Does your child regularly have aphthous stomatitis (mouth ulcers)?	Yes/ No
8.	Is your child regularly irritated (longer than 3 weeks, at least twice a week)?	Yes / No
9.	Does your child eat gluten?	Yes / No
10.	Has your child been diagnosed with Coeliac Disease?	Yes / No

Question (to fill in by the healthcare provider)

1.	Is growth (height and / or weight) restricted?	Yes/No