OPEN ACCESS

Original research

Endoscopic pyloromyotomy for the treatment of severe and refractory gastroparesis: a pilot, randomised, sham-controlled trial

Jan Martinek ,[1] Rastislav Hustak ,[2,3] Jan Mares ,[4] Zuzana Vackova ,[1] Julius Spicak,[1] Eva Kieslichova,[5] Marie Buncova,[6] Daniel Pohl ,[7] Sunil Amin ,[8] Jan Tack [9]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2022-326904).

For numbered affiliations see end of article.

Correspondence to
Professor Jan Martinek, Department of Hepatogastroenterology, Institute of Clinical and Experimental Medicine, Praha 140 21, Czech Republic; jan.martinek@volny.cz

Received 4 January 2022
Accepted 5 April 2022
Published Online First 25 April 2022

Linked

► http://dx.doi.org/10.1136/gutjnl-2022-327545

Check for updates

To cite: Martinek J, Hustak R, Mares J, *et al*. *Gut* 2022;**71**:2170–2178.

ABSTRACT

Objective Endoscopic pyloromyotomy (G-POEM) is a minimally invasive treatment option with promising uncontrolled outcome results in patients with gastroparesis.

Design In this prospective randomised trial, we compared G-POEM with a sham procedure in patients with severe gastroparesis. The primary outcome was the proportion of patients with treatment success (defined as a decrease in the Gastroparesis Cardinal Symptom Index (GCSI) by at least 50%) at 6 months. Patients randomised to the sham group with persistent symptoms were offered cross-over G-POEM.

Results The enrolment was stopped after the interim analysis by the Data and Safety Monitoring Board prior to reaching the planned sample of 86 patients. A total of 41 patients (17 diabetic, 13 postsurgical, 11 idiopathic; 46% male) were randomised (21 G-POEM, 20-sham). Treatment success rate was 71% (95% CI 50 to 86) after G-POEM versus 22% (8–47) after sham (p=0.005). Treatment success in patients with diabetic, postsurgical and idiopathic gastroparesis was 89% (95% CI 56 to 98), 50% (18–82) and 67% (30–90) after G-POEM; the corresponding rates in the sham group were 17% (3–57), 29% (7–67) and 20% (3–67).
Median gastric retention at 4 hours decreased from 22% (95% CI 17 to 31) to 12% (5–22) after G-POEM and did not change after sham: 26% (18–39) versus 24% (11–35). Twelve patients crossed over to G-POEM with 9 of them (75%) achieving treatment success.

Conclusion In severe gastroparesis, G-POEM is superior to a sham procedure for improving both symptoms and gastric emptying 6 months after the procedure. These results are not entirely conclusive in patients with idiopathic and postsurgical aetiologies.

Trial registration number NCT03356067; ClinicalTrials.gov.

Significance of this study

What is already known on this subject?

⇒ Gastroparesis (GP) is a gastric motility disorder with a complex pathophysiology, which is not completely understood.
⇒ Pylorospasm is believed to play a role in the pathophysiology of GP.
⇒ Endoscopic pyloromyotomy (G-POEM) is a new minimally invasive procedure with promising uncontrolled clinical results in patients with GP.

What are the new findings?

⇒ In this randomised and sham controlled trial that included 41 patients with severe GP, symptomatic improvement at 6 months was achieved in 71% of the patients after G-POEM compared with 22% after the sham procedure. Moreover, 75% of the patients achieved symptomatic improvement 6 months after cross-over G-POEM, which was offered to patients without treatment success after the sham procedure.
⇒ The trial was terminated early due to a significant result and given the risk of general anaesthesia in patients in the sham group.
⇒ Gastric emptying improved after G-POEM but did not change after the sham procedure.
⇒ The trial was not sufficiently powered to assess the effectiveness of G-POEM in the aetiology subgroups. Our results cannot be considered as fully conclusive in patients with idiopathic and postsurgical aetiologies.
⇒ The study design did not allow for the assessment of the relationship between symptom improvement at 6 months and changes in gastric emptying, as these two parameters were obtained at different time points after the procedure.

How might it impact on clinical practice in the foreseeable future?

⇒ The spectrum of treatment methods that can be offered to patients with severe and refractory GP can be extended by G-POEM.

INTRODUCTION

Gastroparesis (GP) is a gastric motility disorder defined by the presence of upper abdominal symptoms and delayed gastric emptying in the absence of organic obstruction.[1 2] Two important aetiologies are diabetes mellitus and GP following oesophageal

or gastric surgery (postsurgical GP). In up to half of the patients with GP, no underlying aetiology can be identified, and these patients are referred to as having idiopathic GP. The symptoms include nausea, vomiting, early satiety, postprandial fullness, bloating and abdominal pain. In severe cases, GP may lead to weight loss, poor nutritional status and increased mortality.[1–4]

The pathophysiology of GP is multifactorial and incompletely understood. Delayed gastric emptying is a defining feature, and gastric hypomotility due to several underlying mechanisms is believed to play a major role as well.[1] However, the symptomatic benefit of prokinetic agents is often disappointing.[5 6] An inappropriately spastic pyloric muscle has also been suggested as another important pathophysiological factor.[1 7]

Because of this complex pathophysiology, effective treatment for GP is a clinical challenge, especially in patients with severe and refractory disease.[1 2 5 6 8] Treatment options consist of dietary measures, administration of prokinetics and antiemetics, compensation of underlying disease, nutritional support and other methods such as gastric electrical stimulation, but none of these options is supported by strong scientific evidence.[1 2 8] Pylorus-directed therapies (botulinum toxin injection, balloon dilation, surgical pyloroplasty) constitute another approach.[8 9] Their common aim is to decrease pyloric tone, which is thought to be increased in patients with GP.[7] However, these therapies have not been recognised as a standard mainly due to a lack of scientific evidence for their clinical efficacy.[9] Endoscopic pyloromyotomy (G-POEM) is a new pylorus-directed minimally invasive therapy, consisting of purely endoscopic myotomy. A multitude of non-randomised and non-controlled studies has shown promising clinical efficacy and high safety of G-POEM.[10–14] We performed a randomised trial comparing the clinical efficacy of G-POEM versus a sham procedure in patients with severe and refractory GP.

METHODS

Trial design

We performed a randomised and prospective trial at two European centres (Prague, Czech Republic; Trnava, Slovakia) comparing G-POEM with a sham procedure. All patients signed informed consent prior to enrolment. Patients were followed up for 6 months when treatment allocation was revealed and patients in the control group were offered cross-over G-POEM if they did not achieve treatment success. These patients were followed-up for another 6 months. Study design is summarised in online supplemental tables S4a and S4b and figure S1 in the Supplementary Appendix.

The trial was investigator initiated, was approved by the ethics committee at both centres and was performed in accordance with the provisions of the Declaration of Helsinki. No industry support was received except for a supply of Endoflip balloons by Medtronic.

An independent Data and Safety Monitoring Board (DSMB) surveilled the trial in terms of ethical consideration, patient's safety and data management. On-site data monitoring to ensure the proper conduct of the trial was provided by Axon CRO (online supplemental table S12). All coauthors have reviewed and approved the final manuscript.

Patients

Eligible were patients older than 18 years who suffered from severe (Gastroparesis Cardinal Symptom Index (GCSI) >2.3) and refractory (symptom duration>6 months) GP, which had to be confirmed by a gastric emptying study (GES; scintigraphy, abnormal gastric retention at 2 hours and/or 4 hours on a standardised sulphur colloid solid-phase GES, for details, see online supplemental table S7). Abnormal GES was defined as gastric retention greater than 60% at 2 hours and/or 10% at 4 hours after meal ingestion.[15] Main exclusion criteria were absence of a previous therapy trial with at least one prokinetic drug, major oesophageal or gastric surgery and previous pyloromyotomy or pyloroplasty. A complete list of inclusion/exclusion criteria is displayed in online supplemental tables S6a and S6b.

Randomisation

Patients were randomly assigned in a 1:1 ratio and the randomisation was stratified according to the performing centre, sex and aetiology of GP using randomly permuted six-patient blocks. A dedicated nurse performed the randomisation and a treatment allocation was revealed before the procedure just after induction into general anaesthesia.

Interventions

All patients were admitted to the hospital 1 day prior to the intervention and an upper endoscopy was performed to check for and eventually to clean the stomach from food residues. The patients randomised to the G-POEM group underwent G-POEM (under general anaesthesia) comprising four principal steps: (1) submucosal injection followed by mucosal incision 4–5 cm proximal to the pyloric channel, (2) creation of a submucosal tunnel towards a pyloric ring, (3) a complete myotomy 2–3 cm long, (4) closure of the incision with endoscopic suturing system or endoscopic clips.

The patients randomised to the control group underwent upper GI endoscopy under general anaesthesia, lasting at least 40 min. All procedures were performed by one experienced endoscopist with sufficient experience in submucosal endoscopy. Further details of the G-POEM are provided in the protocol (online supplemental file 1).

Before the procedure, the patients received parenteral antibiotics (or placebo in the control group), and after the procedure, the patients were administered a proton pump inhibitor (or placebo) intravenously on postoperative days 0 and 1 and then all patients received a proton pump inhibitor orally for at least 4 weeks.

We measured pyloric distensibility using the principle of impedance planimetry (Endoflip,[16 17]) before a procedure (G-POEM or sham) and two times after G-POEM. The first two measurements were performed under general anaesthesia with administration of opioids, the third measurement at 3 months was performed under sedation with midazolame. As Endoflip technology was not available when the trial started, measurements are available from patient No. 17 onwards. For details on this measurement, see online supplemental table S8 and figures S2a, S2b.

Trial follow-up

Clinical data were collected at follow-up visits at 3 and 6 months after G-POEM/sham procedure and 3 and 6 months after cross-over G-POEM. Patient-reported outcomes were assessed by means of follow-up appointments by dedicated trial personnel who were not aware of the treatment-group assignments. Objective evaluation by means of endoscopy, GES and Endoflip measurement was performed at 3 months after G-POEM/sham procedure and at 3 months after cross-over G-POEM. Online supplemental tables S4a and S4b and figure S1 provide an overview of the plan of the study assessment. We are further following

the patients to assess both clinical and objective parameters at 12, 24 and 36 months.

No medication was forbidden during the trial and patients were allowed to take prokinetics, antiemetics, antidepressants or other treatments on as needed basis. However, the prokinetics (and anticholinergics) had to be withdrawn at least 3 days before the GES. Pylorus-directed interventions were not allowed during the follow-up.

Primary and secondary outcomes

The primary outcome was the proportion of patients with treatment success at 6 months after the procedure in the intention to treat (ITT) cohort. Treatment success was defined as a decrease of at least 50% in the total GCSI (online supplemental table S5 in the Supplementary Appendix).[18 19]

The primary (null) hypothesis was that G-POEM leads to treatment success in the same proportion of patients as the sham procedure.

Secondary clinical outcomes included proportion of patients with treatment success at 3 months after G-POEM/sham, at 3 and 6 months after cross-over G-POEM, change in GCSI and PAGI-SYM score (online supplemental table S9)[20] and Quality of Life evolution assessed by using the validated PAGI—QoL questionnaire (online supplemental table S10).[21 22]

Prespecified objective outcomes included the change in gastric emptying after G-POEM/sham procedure/cross-over G-POEM and changes in pyloric distensibility and cross-sectional area. Further secondary endpoints included analysis of adverse events and procedure details.

The statistical analysis plan was described in the protocol and specified that clinically relevant exploratory subgroup analyses would be performed. Exploratory subgroups were defined according to aetiology of GP (diabetic, postsurgical, idiopathic).

Statistical analysis

The sample size calculation was based on the conservative estimation that the expected treatment success of G-POEM would be 50% of treated patients compared with 20% in the sham group at a significance level of 0.05 and a study power of 0.8. We planned to randomise 86 patients accounting for a 15% drop off. After the first interim analysis (performed, in accordance with the protocol after 40% (n=34) patients completed 6 months follow-up), the DSMB recommended to stop further enrolment as the analysis showed a highly significant result (p=0.003) in favour of the active treatment arm. The Board considered it ethically controversial to complete the originally planned number of enrolled patients given the risks of general anaesthesia in patients in the control group.

Analyses of the treatment success (main outcome), GCSI, PAGI-SYM, PAGI-QoL scores and GES were performed on the ITT population with the values missing for some of the 41 patients imputed using multiple imputation.[23] At most three values (7%) were missing for any of the variables in the imputation model. The 6-month GCSI value defining the treatment success was imputed in one patient (sham group). A sensitivity analysis was performed for the main outcome using the per-protocol population (PP). Analyses of the GCSI subscores and Endoflip measurements were analysed on available data basis.

The difference in treatment success between the G-POEM and sham groups was assessed using a logistic regression model, which resulted in the only p value presented in this report evaluating the only confirmatory hypothesis. All the remaining results are presented as point estimates (medians, means, HRs) with 95% CIs. We adopted this approach to provide more information and to prevent inadequate interpretations of p values due to the multiple testing. The reader can still identify statistically significant results as those having CIs entirely below or above zero. The CIs are presented without a correction for multiple testing.

CIs for the proportions of treatment success were calculated using the Wilson method and combined with multiple imputation according to Lott and Reiter.[24] The CIs for continuous variables were constructed by smoothed bootstrapping. For a detailed description of statistical analysis, see online supplemental table S11 in the Supplementary Appendix.

The statistical analyses were performed using R V.4.1.2 (packages tidyr 1.1.4, mice 3.13.0, Hmisc 4.6.0, ggpubr 0.4.0, ggplot 2 3.3.5).

Patient and public involvement

Patients and/or public were not involved in the design, conduct or reporting of this trial.

RESULTS

Patients

Between November 2017 and February 2021, a total of 41 patients were randomised (table 1, figure 1) and these patients represent the ITT cohort. Forty patients underwent the assigned procedure (21 G-POEM and 19 sham) while 1 male patient in the control group withdrew consent. One G-POEM could not be completed due to severe submucosal fibrosis. The per-protocol population (PP) comprised 39 patients (20 G-POEM, 19 sham procedure). Fifteen patients, who were originally randomised to the control group, were offered cross-over G-POEM, and 12 of them agreed to undergo it. All these patients received the procedure and completed the 6-month follow-up. Online supplemental table S12 in the Supplementary Appendix shows the distribution of patients between the two centres. Demographic data, symptom scores and Gastrointestinal Quality of Live scores at baseline were similar in both treatment groups (table 1). Procedural data are provided in online supplemental table S18. Surgeries on patients with postsurgical GP included fundoplication or refundoplication (n=12) and laparoscopic Heller myotomy (n=1).

Treatment success

In the intention-to-treat population, 15 of 21 patients (71%, 95% CI 50% to 86%) in the active treatment group and 4 of 20 patients (22%, 95% CI 8% to 47%, one patient imputed) in the control group had treatment success at the 6-month follow-up (the primary endpoint, figure 2, online supplemental table S19). In the per-protocol population (sensitivity analysis), the treatment success was achieved in 14 of 20 patients (70%, 95% CI 48% to 85%) in the G-POEM group and 4 of 19 patients (21%, 95% CI 9% to 43%) in the control group (figures 2 and 3 and online supplemental table S19). Three months after the assigned intervention, treatment success was present in 57% (95% CI 36% to 76%) in the G-POEM group and 22% (95% CI 8% to 47%) in the control group (online supplemental table S19 and figures S4 and S12).

Nine out of 12 patients (75%, 95% CI 47% to 91%) achieved treatment success 6 months after cross-over G-POEM (figure 2, online supplemental table S19).

In an analysis of treatment success with a logistic regression model, the OR for success at 6 months in the G-POEM group, as

Table 1 Demographic and clinical characteristics of the patients at baseline

Characteristic	G-POEM arm	Control (sham) arm
Number of patients	21	20
Sex—number (%)		
Female	11 (52.4)	11 (55.0)
Male	10 (47.6)	9 (45.0)
Age—median (Q1–Q3) (years)	43 (30 – 51)	51 (45 – 56)
BMI – median (Q1–Q3)(kg/m^2)	22 (19 – 28)	26 (21 – 28)
Aetiology—number (%)		
Diabetic; (diabetes type I/diabetes type II, number)	9 (42.9); (8/1)	8 (40.0); (6/2)
Post-surgical	6 (28.6)	7 (35.0)
Idiopathic	6 (28.6)	5 (25.0)
Previous therapy—number (%)		
Metoclopramide	12 (57.1)	10 (50.0)
Itopride	11 (52.4)	10 (50.0)
Domperidone	9 (42.9)	7 (35.0)
Other prokinetics	3 (14.3)	2 (10.0)
Enteral feeding via nasojejunal/nasogastric tube	3 (14.3)	1 (5.0)
Recurrent hospitalisation for gastroparesis-related symptom	8 (38.1)	7 (35.0)
Baseline GCSI score—median (Q1–Q3)*	3.5 (3.2–3.7)	3.2 (2.6–3.4)
Baseline PAGI-QOL score—median (Q1–Q3)†	2.1 (1.7–2.7)	2.5 (1.4–2.8)
Baseline 4 hours GES retention—median (Q1–Q3)(%)‡	22 (17–32)	26 (16–42)
Pre-procedure DI 40 mL—median (Q1–Q3)(mm^2/mm Hg)§	5.8 (4.8–9.8)	5.6 (3.5–6.2)

Q1–Q3—the first and the third quartile (representing the middle half of the observed values), the difference between Q3 and Q1 is the inter-quartile range.
*GCSI is a validated score assessing symptoms severity in patients with gastroparesis, consisting of nine items (symptoms) and three subscales (nausea/vomiting subscale, postprandial fullness/early satiety subscale and the bloating subscale). Each item can be graded from 0 (no symptom) to 5 (maximally severe symptoms). The total GCSI is calculated as the average of all three subscale averages. GCSI value ranges from 0 (no symptoms) to 5 (maximally severe symptomatology). The index evaluates symptoms during the last 14 days. Only patients with GSCI > 2.3 (indicating severe disease) were eligible for enrolment.
†PAGI-QOL score—a validated QoL questionnaire measures quality of life outcomes in patients with upper gastrointestinal disorders. It contains 30 items with five subscales (daily activities, clothing, diet/food habits, relationship, psychological well-being and distress). A total score is calculated by averaging subscales scores, its value ranges from 0 (perfect QoL) to 5 (worse QoL).
‡GES—is a validated method to demonstrate delayed gastric emptying in patients with gastroparesis. In this trial, all GES were performed according to a standardised method for measuring gastric emptying by scintigraphy; a low-fat, egg-white meal with imaging at 0, 1, 2 and 4 hours after meal ingestion was used for each patient. Only patients with a retention of Tc-99m > 60% at 2 hours and/or≥10% at 4 hours on a standardised sulphur colloid solid phase were eligible for enrolment.
§DI—pyloric distensibility is one among several parameters obtained from measurement of pyloric distensibility by using impendance planimetry principle (Endoflip). Values below 10 mm^2/mm Hg are thought to demonstrate a pylorospasm. In this trial, not all patients underwent Endoflip measurement as the method was not available when the trial started.
DI, Distensibility Index; GCSI, Gastroparesis Cardinal Symptom Index; GES, gastric emptying study; G-POEM, endoscopic pyloromyotomy; PAGI-QOL, Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life.

compared with the control group, was 9.0 (95% CI 2.0 to 40.2, p=0.005) (table 2).

In patients with diabetic GP, the treatment success in the G-POEM group at 6 months was 89% (95% CI 56% to 98%, eight of nine patients), in postsurgical GP 50% (95% CI 18% to 82%, three of six patients) and in idiopathic GP 67% (95% CI 30% to 90%, four of six patients). The corresponding rates of treatment success in the sham group were 17% (95% CI 3% to 57%, one of seven plus one patient imputed), 29% (95% CI 7% to 67%, two of seven patients) and 20% (95% CI 3% to 67%, one of five patients) (figure 2, online supplemental table S19).

Exploratory analyses suggest that male gender, gastric retention at 4 hours below 20% and post G-POEM pyloric distensibility > 13 mm^2/mm Hg at 40 mL may predict a treatment success (table 2).

Secondary outcomes—symptoms and QoL

The median GCSI decreased in the G-POEM group from a baseline value of 3.5 (95% CI 3.2 to 3.7) to 1.4 (95% CI 0.9 to 1.9) at 3 months and 1.1 (95% CI 0.5 to 1.5) at 6 months postintervention, while in the sham group, it decreased from 3.2 (95% CI 2.8 to 3.4) to 2.5 (95% CI 1.9 to 3.1) at 3 months and 2.5 (95% CI 1.9 to 3.2) at 6 months (figure 3). The median reduction from baseline to 6 months was 2.4 (95% CI 2.0 to 2.8) in the active arm and 0.7 (95% CI 0.0 to 1.2) in the sham group (online supplemental table S20). Evolution of GCSI subscores is displayed in online supplemental table S21 and figures S5 and S6.

After cross-over G-POEM, GCSI significantly decreased from 2.8 (95% CI 2.5 to 3.7) to 1.0 (95% CI 0.6 to 1.7), the reduction from baseline was 2.1 (95% CI 1.3 to 2.6) (figure 3, online supplemental table S20, figures S5 and S6). Gastrointestinal Quality of Life Index score decreased from 2.1 (95% CI 1.7 to 2.5) to 0.8 (95% CI 0.6 to 1.5) showing a significant median reduction by 1.1 (95% CI 0.1 to 1.6) in the G-POEM group. In the sham group, the score decreased from 2.5 (95% CI 1.5 to 2.9) to 1.7 (95% CI 1.2 to 2.4) with a median reduction by 0.4 (95% CI −0.1 to 0.8). After cross-over G-POEM, the score decreased from 2.2 (95% CI 1.3 to 3.3) to 1.6 (95% CI 0.7 to 2.3) with the median reduction by 0.3 (95% CI 0.1 to 1.6) (online supplemental table S20, figure S8).

Secondary outcomes—objective parameters

Gastric retention at 4 hours decreased significantly after G-POEM but did not change after the sham procedure. Furthermore, gastric retention significantly decreased after the cross-over procedure. There was no correlation between GCSI and gastric retention at 3 months (r=0.15 (95% CI −0.18 to 0.42)). Detailed results on gastric emptying are shown at figure 4, online supplemental figures S9, S10, S12 and table S20.

Distensibility index at 40 mL (mm^2/mm Hg) increased from a baseline value of 7.6 (95% CI 6.0 to 9.3) to 12.7 (95% CI 11.4 to 14.3) after the procedure and to 13.1 (95% CI 11.3 to 15.7) at 3 months. The corresponding values for cross-sectional area (CSA, mm^2, 40 mL) were 144 (95% CI 125 to 165), 199 (95% CI 177 to 219) and 206 (95% CI 185 to 234). Detailed analysis of Endoflip measurements is provided in online supplemental table S22 and figure S11.

Safety

Ten serious adverse events (SAEs) occurred, 7 after G-POEM (five in the G-POEM group, two after cross-over G-POEM) and 3 in the sham group. All three SAEs in the sham group were not related to the procedure but rather to GP itself or to a newly diagnosed achalasia (online supplemental tables S15, S16).

Three SAEs were related to the G-POEM procedure (9% of all G-POEMs performed). One patient developed abdominal pain 1 day after G-POEM and was diagnosed with a gastric ulcer

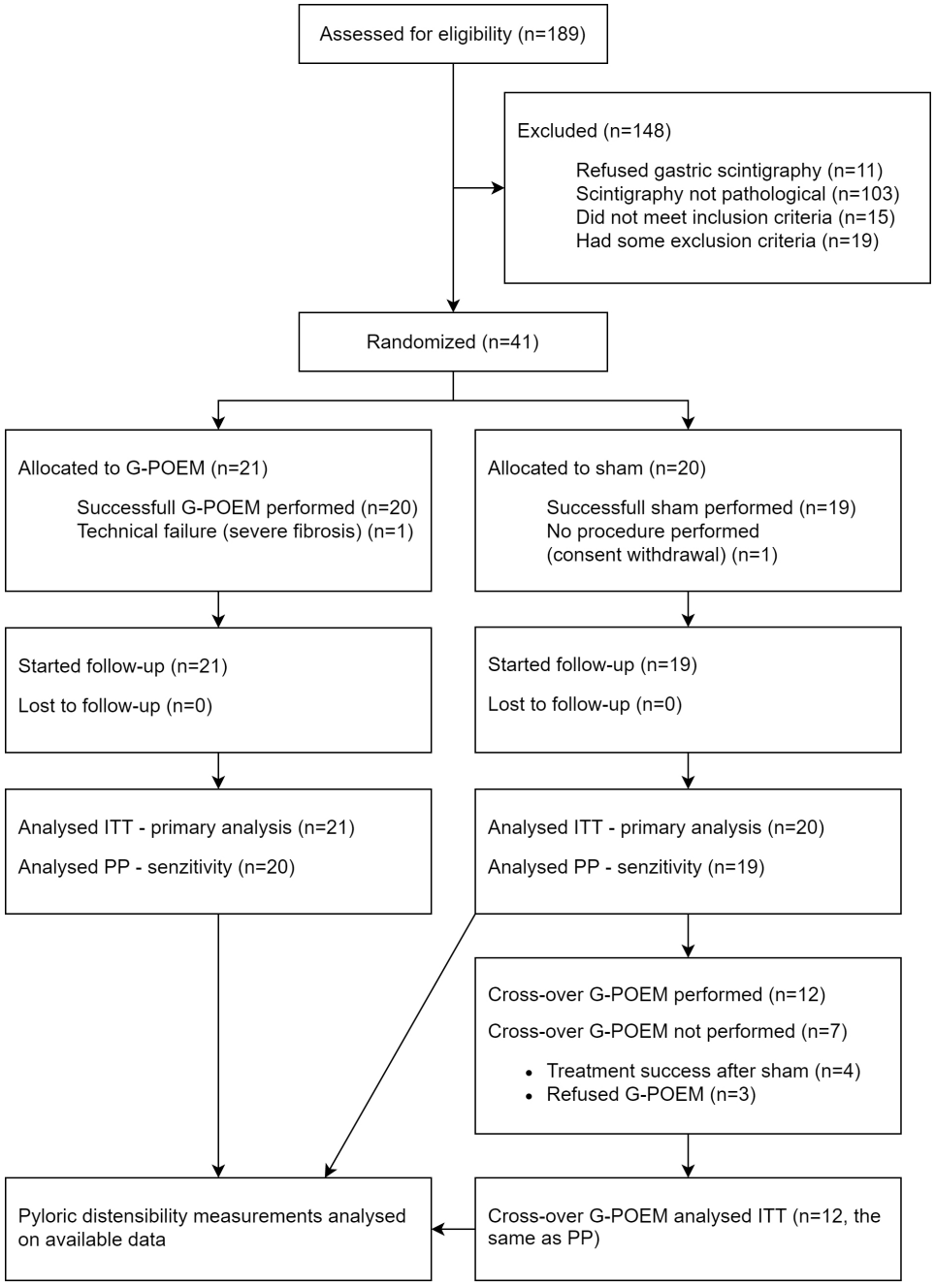


Figure 1 Flowchart demonstrating screening, enrolment, randomisation, follow-up and cross-over procedure with the subsequent follow-up. Eligible patients from the two centres were randomly assigned to either G-POEM or the sham procedure consisting of endoscopic examination under general anaesthesia. The length of the follow-up was 6 months when the treatment allocation was revealed. A total of 12 patients, who did not have treatment success after the sham procedure and agreed with a cross-over endoscopic pyloromyotomy (G-POEM), underwent the procedure and were followed up for another 6 months. The intention to treat (ITT) analysis comprises 41 patients, the per protocol (PP) analysis 39 patients.

near the pylorus. Conservative management was successful. Another patient had a mucosal injury during the G-POEM at the level of myotomy and was kept longer in the hospital as a precautionary measure. The third patient developed moderate dumping syndrome 3 months after the cross-over G-POEM with a need for hospitalisation, resulting in a complete resolution (online supplemental tables S15 and S16).

DISCUSSION

To our knowledge, this is the first prospective trial to compare the clinical effectiveness of G-POEM with a sham procedure in patients suffering from severe and refractory GP. Six months after the procedure, a significant treatment effect was achieved in 71% of patients in the active arm compared with 22% in the control group. Furthermore, treatment success was achieved in 75% of patients after cross-over G-POEM. G-POEM was associated with both improved gastric emptying and increased pyloric distensibility.

The two main mechanisms responsible for GP are believed to be postprandial gastric hypomotility and an abnormal control of pyloric muscle contractility resulting in pylorospasm.[1–3 7] Current treatment of GP is comprised of symptomatic measures

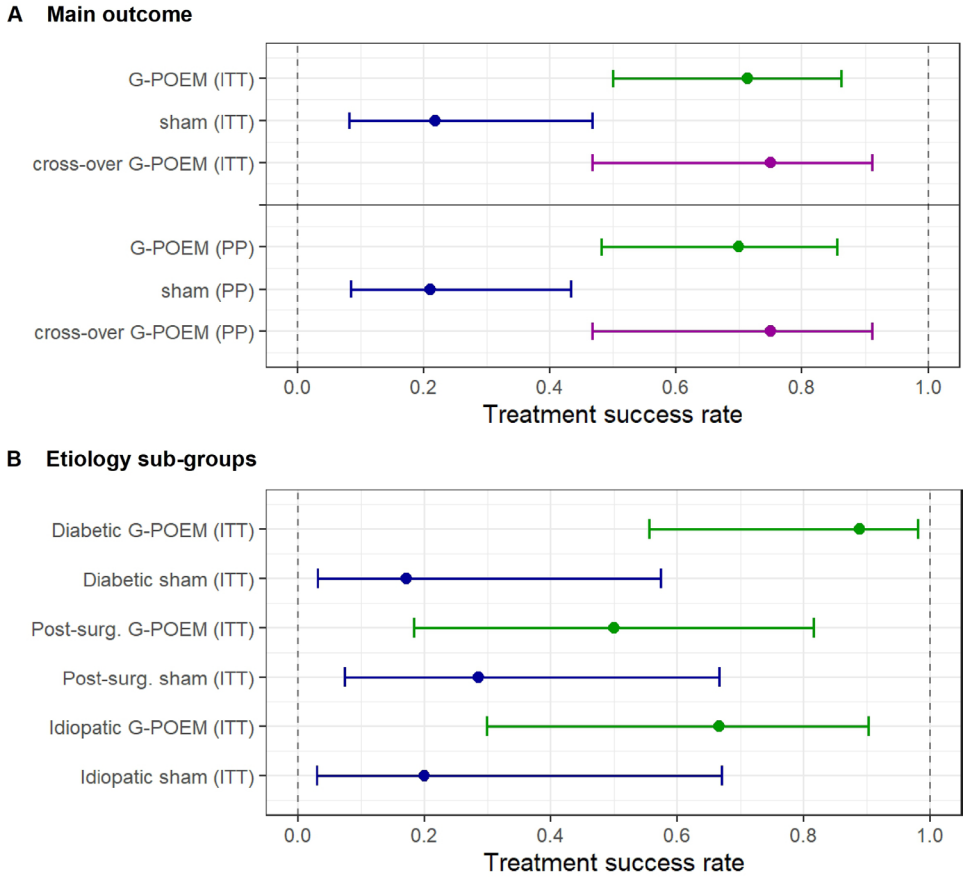


Figure 2 Treatment success at 6 months after the assigned procedure (main outcome), after the crossover G-POEM (A) and treatment success in sub-groups by aetiology of gastroparesis (B). The plot shows rates of treatment success with 95% CIs, where the clinical success is defined as reduction of the total Gastroparesis Cardinal Symptom Index (GCSI) score by at least 50% from the baseline. For the cross-over endoscopic pyloromyotomy (G-POEM), GCSI at 6 months after the sham procedure was considered as baseline. The results analysed on the intention to treat (ITT) population (N=41, N-Di-G-POEM=9, N-Di-Sham=8, N-PS-G-POEM=6, N-PS-Sham=7, N-Id-G-POEM=6, N-Id-Sham=5, 1 GCSI value (2 %) imputed in diabetic GP patient in the sham group) are supplemented by the main outcome analysis on the per protocol (PP) population (N=39).

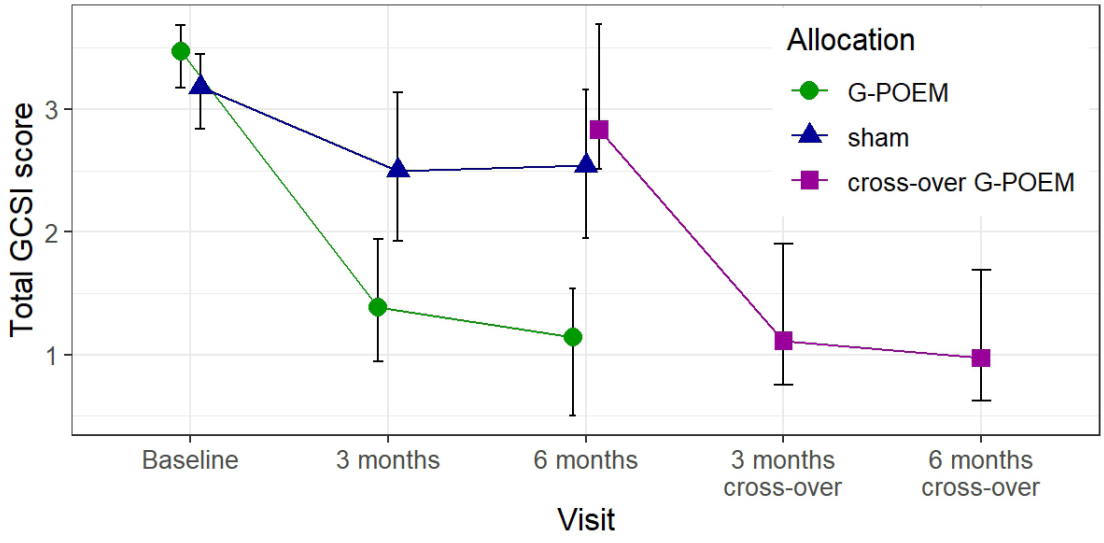


Figure 3 Evolution of the Gastroparesis Cardinal Symptom Index (GCSI) total score. Point estimates of medians with 95% CIs calculated on the intention to treat (ITT) population are shown for patients after the endoscopic pyloromyotomy (G-POEM) procedure (green circles, N=21), sham procedure (blue triangles, N=20, imputed 1 value (5 %) for 3 months and 1 value (5 %) for 6 months), and cross-over G-POEM procedure (purple squares, N=12). For the cross-over G-POEM group, the value at 6 months reflects only the data for the patients in this group (who subsequently underwent the cross-over G-POEM procedure). The GCSI score may range from 0 (no symptoms) to 5 (maximally severe symptoms).

Table 2 Primary treatment success comparison G-POEM versus sham at 6 months and predictors of treatment success at 6 months

Variable	OR*	95% CI for OR*	P value
Allocation G-POEM	9.0	2.0 to 40.2	0.005
Gender male	4.0	1.0 to 15.8	
Age >47 years	0.69	0.19 to 2.52	
Baseline GCSI >2.6	2.6	0.4 to 16.4	
Baseline GES 4 hours >20 %	0.24	0.06 to 0.93	
Baseline distensibility (DI) >8 mm^2/mm Hg†	3.6	0.5 to 33.6	
Post G-POEM distensibility (DI) >13 mm^2/mm Hg‡	6.0	0.66 to 136.8	

Each variable was tested as a predictor of treatment success in a separate logistic regression model. Only one p value for the main outcome is presented. The analyses of distensibility were performed on available data with N=19 for baseline distensibility and N=16 for post G-POEM distensibility. The remaining analyses used the ITT population with N=41, one treatment success value was imputed and there were no missing data in the predictor variables.
*Single parameter statistical significance can be judged by the CI for OR lying entirely below (reduced chance of treatment success) or above (increased chance for treatment success) the value of 1.
†Includes sham patients who did not undergo cross-over G-POEM, data for 40 mL filling.
‡Primary G-POEM and cross-over G-POEM combined, data for 40 mL filling.
DI, Distensibility Index (Endoflip measurement); GCSI, Gastroparesis Cardinal Symptom Index; G-POEM, endoscopic pyloromyotomy; ITT, intention to treat.

(dietary adjustments, antiemetics, nutritional support) as well as causal treatments targeting the proven or assumed pathophysiological mechanisms (prokinetic drugs, gastric electrical stimulation, pylorus-directed therapies).[1–3 5 6 8 9] Despite the existence of several options, treatment of GP is often partially effective or ineffective.[4–6 8 9 25]

Pylorospasm is believed to be an important pathophysiological factor, which was first demonstrated in 1986 by finding an increase of baseline pyloric tone in 60% of symptomatic diabetic patients.[7] Further evidence came from experimental studies, showing, for example, an insulin-sensitive reversible loss of neuronal nitric oxide synthase responsible for relaxation of the pylorus in diabetic mice.[26] To date, the main evidence has been brought by several studies showing some clinical effectiveness of pylorus-directed therapies, including G-POEM.[10–14 27–29] G-POEM showed short and mid-term clinical efficacy in 56%–81% of patients and improved gastric emptying in several uncontrolled and non-randomised studies.[10–14 30]

Not all data corroborate the hypothesis that pylorospasm plays the dominant pathophysiologic role in patients with GP. For example, a substantial number of patients do not respond to G-POEM, and partial efficacy has been reported for several treatments that do not influence pyloric tone.[1–3 8 31] In addition, two placebo-controlled trials did not show a benefit of intrapyloric injection of botulinum toxin injection.[32 33]

G-POEM should be indicated in patients with proven pylorospasm. The key question, however, of how to select these candidates remains. Unfortunately, no specific GP symptom pattern or aspect is associated with pylorospasm. Furthermore, it is not known whether different aetiologies of GP are associated with a differential response to pylorus-directed treatment. Based on previous[28] as well as the current study, measurement of pyloric distensibility by impedance planimetry may be a promising tool for patient selection. However, to date, normal values have not yet been defined and the protocol of measurement is not well standardised.

Our trial demonstrated a favourable effect of G-POEM in unselected patients with GP. This was most clearly the case in patients with diabetic GP while in the smaller subgroups of patients with postsurgical and idiopathic GP, the differences between active and sham treatment were numerically lower. Of note, one female patient in the postsurgical group without treatment success after initial G-POEM underwent redo-G-POEM with an excellent effect. It may signify that either the first G-POEM was not done well, or, that a double myotomy may be required in some patients as suggested by one retrospective study.[34] It is noteworthy that had this postsurgical patient had treatment success with the initial G-POEM, the rate of treatment success in the postsurgical group would be 67% (95% CI 30 to 90).

Another prospective trial reported rather modest (56%) clinical effectiveness of G-POEM 12 months after the procedure.[12] There may be several explanations for this difference.

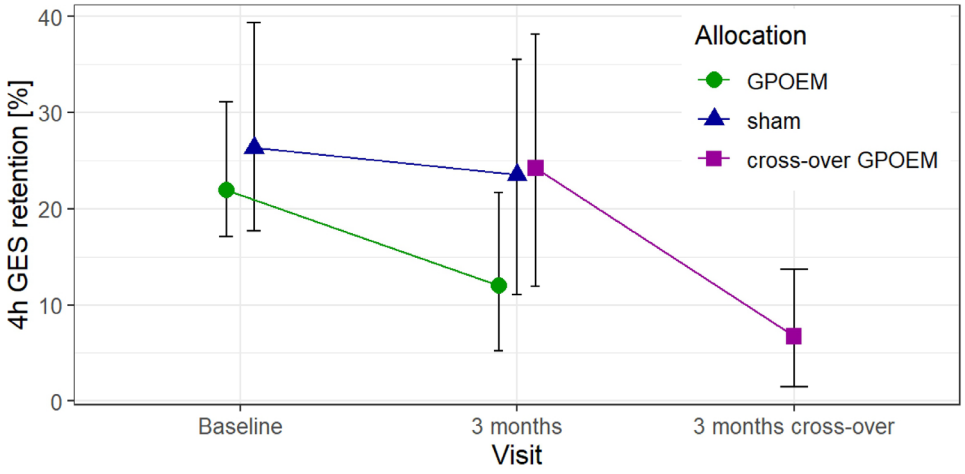


Figure 4 Evolution of gastric retention at 4 hours after meal ingestion on a standardised sulphur colloid solid-phase gastric emptying study (scintigraphy). Point estimates of medians with 95% CIs are shown for patients after the G-POEM procedure (green circles, N=21, imputed 2 values (10 %) for 3 months), sham procedure (blue triangles, N=20, imputed 1 value (5 %) for 3 months), and cross-over GPOEM procedure (purple squares, N=12). For the cross-over G-POEM group, the value at 3 months reflects only the data for the patients in this group (who subsequently underwent the cross-over G-POEM procedure). GES, gastric emptying study; G-POEM, endoscopic pyloromyotomy.

For example, in our trial, the largest subgroup of patients had diabetic GP (with a predominance of type I diabetes) and this aetiology responded best to G-POEM. In the study by Vosoughi *et al*, diabetic GP accounted for the smallest subgroup of patients (with a predominance of type II). Or, we enrolled patients with severe symptoms, the presence of which may predict good clinical effect of G-POEM.

Our primary endpoint was treatment success defined as a 50% reduction from the baseline symptom index. By contrast, in some previous studies, treatment success was defined as a decrease of the symptom score of at least one point.[10 12 14 30] However, this is a relatively low threshold, which may well be susceptible to spontaneous improvement and placebo effects. The choice of a 50% reduction sets a higher threshold, which is less susceptible to these confounders[35] and is clinically meaningful. Nevertheless, if we had defined our treatment success similarly to other studies, the difference between active treatment and control groups would not have changed (online supplemental table S19 and figure S3). We believe that there is a need to adopt a standard definition of treatment success, so that the studies are better comparable.

The treatment success was corroborated by two objective measurements. Patients after pyloromyotomy, in contrast to patients after the sham procedure, showed significantly improved gastric emptying, even if our results are in line with a lack of consistent reproducible relationships between global GP symptoms and gastric emptying delay.[36]

Pyloromyotomy also increased both pyloric distensibility and cross-sectional area. Unfortunately, Endoflip measurement was started midway through the trial as it was not available when the trial started. Furthermore, two out of three measurements were performed under general anaesthesia with opioids, which could have influenced measured values. Therefore, we cannot draw any firm conclusion with this respect. However, similarly to other two studies, we showed a trend that post-G-POEM distensibility > 13 mm Hg/mm^2 might predict treatment success.[17 37]

We experienced 10 SAEs, but only 3 were related to pyloromyotomy. Even if our results are in line with other studies reporting the occurrence of severe adverse events after G-POEM up to 6%,[38] one case of moderate dumping syndrome in our study and one case report of a severe refeeding syndrome in the literature[39] should be considered when performing pyloromyotomy as this procedure is not free of SAEs. Based on postprocedural symptoms evolution, our patients did not experience new onset or worsening of duodeno-gastric reflux, which is theoretically one of the SAEs with possible long-term sequalae.

Our study has several limitations. First, our follow-up is only 6 months after the procedure, and clinical recurrences may still occur after this time.[14] As such, we continue to follow-up our patients in the absence of further blinding. Longer blinded follow-up was not an option given the severity of the GP symptoms in our patients. Second, with the premature termination of our trial due to the significant results, we did not achieve the planned number of randomised patients. We followed the recommendation of DSMB given the risk of general anaesthesia in patients undergoing sham procedure. The lower number of enrolled patients did not influence the evaluation of the main endpoint but hampers the interpretation of results for the individual types of GP because of a lower number of subjects in postsurgical and idiopathic groups. Third, as we measured gastric emptying at a different time than primary endpoint, we could not accurately assess the relationship between the change in gastric emptying and symptomatic improvement. Future studies should reflect the need to determine the relationship between symptoms and gastric emptying. Fourth, we did not investigate relevant pathophysiological parameters (antroduodenal and small intestinal dysmotility, vagal function), all of which could play a role in development of symptoms or post-G-POEM adverse events. They might also identify a subgroup of patients less likely to respond to the G-POEM procedure. Fifth, all G-POEMs were performed by a single endoscopist, thus, limiting the generalisability of our results.

In conclusion, our results demonstrated that G-POEM is beneficial in a substantial proportion of patients with severe and refractory GP. These results may help expand the range of available treatment options for patients suffering from this debilitating disease. However, our results need to be confirmed, in particular, among patients with idiopathic and postsurgical aetiologies as the results in these two subgroups are not entirely conclusive. Finally, correct patient selection with an emphasis on long-term results should be the focus of future research.

Author affiliations
[1]Department of Hepatogastroenterology, Institute of Clinical and Experimental Medicine, Praha, Czech Republic
[2]Department of Internal Medicine, University Hospital Trnava, Trnava, Slovakia
[3]Institute of Physiology, Charles University in Prague, Prague, Czech Republic
[4]Department of IT and Biostatistics, Institute of Clinical and Experimental Medicine, Praha, Czech Republic
[5]Department of Anesthesiology and Intensive Care, Institute of Clinical and Experimental Medicine, Praha, Czech Republic
[6]Department of Nuclear Medicine, Institute of Clinical and Experimental Medicine, Praha, Czech Republic
[7]Division of Gastroenterology & Hepatology, University Hospital Zürich, Zürich, Switzerland
[8]Division of Gastroenterology, University of Miami Miller School of Medicine, Miami, Florida, USA
[9]Clinical and Experimental Medicine, University of Leuven, Leuven, Belgium

Twitter Sunil Amin @SunilAminMD

Contributors JMartinek: study concept and study design, patient's recruitment, performing all G-POEM procedures, follow-up endoscopies, drafting the manuscript; the guarantor. RH: study design, data collection, assisting the procedures, follow-up endoscopies, drafting the manuscript. JMares: statistical analysis, study design, drafting the manuscript. ZV: patient's recruitment and selection, measurement and analysis of pyloric distensibility measurement, follow-up endoscopies, critical review of the manuscript. JS: follow-up endoscopies, critical review of the manuscript. EK: anaesthesia during procedures, analysis of AEs, critical review of the manuscript. MB: gastric emptying studies—analysis and critical review of the manuscript. DP: Endoflip measurement and analysis, study design. SA: critical review of the manuscript, advisor. JT: study concept, study design, data analysis and interpretation, critical review of the manuscript.

Funding The work was supported by a Grant 17-28797A from the Czech Ministry of Health.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by: 1. Ethical Committee of the Institute for Clinical and Experimental Medicine and of the Thomayer's hospital in Prague (ID = title of the study). 2. Institutional Review Board of the University Hospital in Trnava, Slovakia (ID = title of the study). Approved in 2017. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Jan Martinek http://orcid.org/0000-0002-1415-4719
Rastislav Hustak http://orcid.org/0000-0001-8669-9024
Jan Mares http://orcid.org/0000-0002-4727-3685
Zuzana Vackova http://orcid.org/0000-0002-0443-5752
Daniel Pohl http://orcid.org/0000-0002-0855-1152
Sunil Amin http://orcid.org/0000-0002-3067-4730
Jan Tack http://orcid.org/0000-0002-3206-6704

REFERENCES

1 Camilleri M, Sanders KM. Gastroparesis. *Gastroenterology* 2022;162:68–87.
2 Schol J, Wauters L, Dickman R, *et al*. United European gastroenterology (UEG) and European Society for neurogastroenterology and motility (ESNM) consensus on gastroparesis. *United European Gastroenterol J* 2021;9:287–306.
3 Jung H-K, Choung RS, Locke GR, *et al*. The incidence, prevalence, and outcomes of patients with gastroparesis in Olmsted County, Minnesota, from 1996 to 2006. *Gastroenterology* 2009;136:1225–33.
4 Gourcerol G, Melchior C, Wuestenberghs F, *et al*. Delayed gastric emptying as an independent predictor of mortality in gastroparesis. *Aliment Pharmacol Ther* 2022;55:867–75.
5 Acosta A, Camilleri M. Prokinetics in gastroparesis. *Gastroenterol Clin North Am* 2015;44:97–111.
6 Vijayvargiya P, Camilleri M, Chedid V, *et al*. Effects of Promotility agents on gastric emptying and symptoms: a systematic review and meta-analysis. *Gastroenterology* 2019;156:1650–60.
7 Mearin F, Camilleri M, Malagelada JR. Pyloric dysfunction in diabetics with recurrent nausea and vomiting. *Gastroenterology* 1986;90:1919–25.
8 Lacy BE, Tack J, Gyawali CP. AGA clinical practice update on management of medically refractory gastroparesis: expert review. *Clin Gastroenterol Hepatol* 2022;20:491–500.
9 Weusten BLAM, Barret M, Bredenoord AJ, *et al*. Endoscopic management of gastrointestinal motility disorders - part 1: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 2020;52:498–515.
10 Kahaleh M, Gonzalez J-M, Xu M-M, *et al*. Gastric peroral endoscopic myotomy for the treatment of refractory gastroparesis: a multicenter international experience. *Endoscopy* 2018;50:1053–8.
11 Gonzalez J-M, Lestelle V, Benezech A, *et al*. Gastric per-oral endoscopic myotomy with antropyloromyotomy in the treatment of refractory gastroparesis: clinical experience with follow-up and scintigraphic evaluation (with video). *Gastrointest Endosc* 2017;85:132–9.
12 Vosoughi K, Ichkhanian Y, Benias P, *et al*. Gastric per-oral endoscopic myotomy (G-POEM) for refractory gastroparesis: results from an international prospective trial. *Gut* 2022;71:25–33.
13 Li P, Ma B, Gong S, *et al*. Gastric per-oral endoscopic myotomy for refractory gastroparesis: a meta-analysis. *J Gastrointest Surg* 2021;25:1108–16.
14 Abdelfatah MM, Noll A, Kapil N, *et al*. Long-Term outcome of gastric per-oral endoscopic pyloromyotomy in treatment of gastroparesis. *Clin Gastroenterol Hepatol* 2021;19:816–24.
15 Abell TL, Camilleri M, Donohoe K, *et al*. Consensus recommendations for gastric emptying scintigraphy: a joint report of the American neurogastroenterology and motility Society and the Society of nuclear medicine. *J Nucl Med Technol* 2008;36:44–54.
16 Desprez C, Roman S, Leroi AM, *et al*. The use of impedance planimetry (Endoscopic Functional Lumen Imaging Probe, EndoFLIP®) in the gastrointestinal tract: A systematic review. *Neurogastroenterol Motil* 2020;32:e13980.
17 Vosoughi K, Ichkhanian Y, Jacques J, *et al*. Role of endoscopic functional luminal imaging probe in predicting the outcome of gastric peroral endoscopic pyloromyotomy (with video). *Gastrointest Endosc* 2020;91:1289–99.
18 Revicki DA, Rentz AM, Dubois D, *et al*. Development and validation of a patient-assessed gastroparesis symptom severity measure: the gastroparesis cardinal symptom index. *Aliment Pharmacol Ther* 2003;18:141–50.
19 Revicki DA, Rentz AM, Dubois D, *et al*. Gastroparesis cardinal symptom index (GCSI): development and validation of a patient reported assessment of severity of gastroparesis symptoms. *Qual Life Res* 2004;13:833–44.
20 Rentz AM, Kahrilas P, Stanghellini V, *et al*. Development and psychometric evaluation of the patient assessment of upper gastrointestinal symptom severity index (PAGI-SYM) in patients with upper gastrointestinal disorders. *Qual Life Res* 2004;13:1737–49.
21 De La Loge C, Trudeau E, Marquis P, *et al*. Responsiveness and interpretation of a quality of life questionnaire specific to upper gastrointestinal disorders. *Clin Gastroenterol Hepatol* 2004;2:778–86.
22 Revicki DA, Rentz AM, Tack J, *et al*. Responsiveness and interpretation of a symptom severity index specific to upper gastrointestinal disorders. *Clin Gastroenterol Hepatol* 2004;2:769–77.
23 White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377–99.
24 Lott A, Reiter JP. Wilson confidence intervals for binomial proportions with multiple imputation for missing data. *Am Stat* 2020;74:109–15.
25 Pasricha PJ, Grover M, Yates KP, *et al*. National Institute of diabetes and digestive and kidney Diseases/National Institutes of health gastroparesis clinical research Consortium. functional dyspepsia and gastroparesis in tertiary care are interchangeable syndromes with common clinical and pathologic features. *Gastroenterology* 2021;160:2006–17.
26 Watkins CC, Sawa A, Jaffrey S, *et al*. Insulin restores neuronal nitric oxide synthase expression and function that is lost in diabetic gastropathy. *J Clin Invest* 2000;106:373–84.
27 Clarke JO, Snape WJ. Pyloric sphincter therapy: botulinum toxin, stents, and pyloromyotomy. *Gastroenterol Clin North Am* 2015;44:127–36.
28 Desprez C, Melchior C, Wuestenberghs F, *et al*. Pyloric distensibility measurement predicts symptomatic response to intrapyloric botulinum toxin injection. *Gastrointest Endosc* 2019;90:754–60.
29 Katzka DA, Camilleri M. Treating the pylorus in gastroparesis: the new riddle wrapped in the ultimate enigma? *Gastrointest Endosc* 2020;91:1300–2.
30 Gregor L, Wo J, DeWitt J, *et al*. Gastric peroral endoscopic myotomy for the treatment of refractory gastroparesis: a prospective single-center experience with mid-term follow-up (with video). *Gastrointest Endosc* 2021;94:35–44.
31 Ducrotte P, Coffin B, Bonaz B, *et al*. Gastric electrical stimulation reduces refractory vomiting in a randomized crossover trial. *Gastroenterology* 2020;158:506–14.
32 Arts J, Holvoet L, Caenepeel P, *et al*. Clinical trial: a randomized-controlled crossover study of intrapyloric injection of botulinum toxin in gastroparesis. *Aliment Pharmacol Ther* 2007;26:1251–8.
33 Friedenberg FK, Palit A, Parkman HP, *et al*. Botulinum toxin A for the treatment of delayed gastric emptying. *Am J Gastroenterol* 2008;103:416–23.
34 Abdelfatah MM, Li B, Kapil N, *et al*. Short-term outcomes of double versus single pyloromyotomy at peroral endoscopic pyloromyotomy in the treatment of gastroparesis (with video). *Gastrointest Endosc* 2020;92:603–9.
35 Pasricha PJ, Yates KP, Nguyen L, *et al*. Outcomes and factors associated with reduced symptoms in patients with gastroparesis. *Gastroenterology* 2015;149:1762–74.
36 Janssen P, Harris MS, Jones M, *et al*. The relation between symptom improvement and gastric emptying in the treatment of diabetic and idiopathic gastroparesis. *Am J Gastroenterol* 2013;108:1382–91.
37 Jacques J, Pagnon L, Hure F, *et al*. Peroral endoscopic pyloromyotomy is efficacious and safe for refractory gastroparesis: prospective trial with assessment of pyloric function. *Endoscopy* 2019;51:40–9.
38 Ichkhanian Y, Vosoughi K, Aghaie Meybodi M, *et al*. Comprehensive analysis of adverse events associated with gastric peroral endoscopic myotomy: an international multicenter study. *Surg Endosc* 2021;35:1755–64.
39 Abdelfatah MM. Refeeding syndrome: a fatal adverse event of gastric peroral endoscopic pyloromyotomy. *Gastrointest Endosc* 2021;94:425–6.

ENDOSCOPIC PYLOROMYOTOMY FOR THE TREATMENT OF SEVERE AND REFRACTORY GASTROPARESIS: A PILOT, RANDOMIZED, SHAM-CONTROLLED TRIAL

Martinek Jan, Hustak Rastislav, Mares Jan, Vackova Zuzana, Spicak Julius, Kieslichova Eva, Buncova Marie, Pohl Daniel, Amin Sunil, Tack Jan

Supplementary Appendix

Martinek J, *et al. Gut* 2022;0:1–9. doi: 10.1136/gutjnl-2022-326904

Suppl Table S1. Table of contents

SUPPLEMENTARY TABLES

Suppl Table S2. List of Contributors

Name	Affiliation	City, Country
Jan Martinek	Department of Hepatogastroenterology, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Rastislav Hustak	Department of Internal Medicine, University Hospital Trnava	Trnava, Slovak Republic
Jan Mares	Department of IT and biostatistic, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Zuzana Vackova	Department of Hepatogastroenterology, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Julius Spicak	Department of Hepatogastroenterology, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Eva Kieslichova	Department of Anesthesiology and Intensive Care, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Marie Buncova	Department of Nuclear Medicine, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Daniel Pohl	Division of Gastroenterology and Hepatology, University Hospital Zurich	Zurich, Switzerland
Amin Sunil	Division of Digestive Health and Liver Diseases, University of Miami Miller School of Medicine	Miami, Florida, USA
Jan Tack	Department of Gastroenterology and Hepatology, University Hospitals Leuven	Leuven, Belgium
Jan Usak	Department of Internal Medicine, University Hospital Trnava	Trnava, Slovak Republic
Martin Janicko	First Department of Internal Medicine, PJ Safarik University	Kosice, Slovak Republic
Tereza Malkova	Department of Hepatogastroenterology, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Monika Horackova	Department of Hepatogastroenterology, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Helena Pitelkova	Department of Hepatogastroenterology, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Adela Kreckova	Department of Hepatogastroenterology, Institute for Clinical and Experimental Medicine (IKEM)	Prague, Czech Republic
Gabriela Petranova	Axon CRO Ltd.	Prague, Czech Republic
Emilia Oleksakova	Department of Internal Medicine, University Hospital Trnava	Trnava, Slovak Republic
Eva Kolarovicova	Department of Internal Medicine, University Hospital Trnava	Trnava, Slovak Republic

Suppl Table S3. Authors contribution	
Jan Martinek	study concept and study design, patient´s recruitment, performing all G-POEM procedures, follow-up endoscopies, drafting the manuscript
Rastislav Hustak	study design, data collection, assisting the procedures, follow-up endoscopies, drafting the manuscript
Jan Mares	statistical analysis, study design, drafting the manuscript
Zuzana Vackova	patient´s recruitment and selection, measurement and analysis of pyloric distensibility measurement, follow-up endoscopies, critical review of the manuscript
Julius Spicak	follow-up endoscopies, critical review of the manuscript
Eva Kieslichova	anesthesia during procedures, analysis of AEs, critical review of the manuscript
Marie Buncova	gastric emptying studies - analysis and critical review of the manuscript
Daniel Pohl	Endoflip measurement and analysis, study design
Sunil Amin	critical review of the manuscript, language corrections
Jan Tack	study concept, study design, data analysis and interpretation, critical review of the manuscript
Jan Usak	patient´s recruitment and selection
Martin Janicko	statistical advisor, study design
Tereza Malkova	endoscopic nurse, assisting the procedures
Monika Horackova	head nurse, post-operative care, follow-up visits
Helena Pitelkova	head nurse, post-operative care, follow-up visits
Adela Kreckova	Study administration, CRFs - data collection
Gabriela Petranova	study monitoring (Axon CRO Ltd.)
Emilia Oleksakova	endoscopic nurse, assisting the procedures
Eva Kolarovicova	study nurse, postoperative care, follow-up visits

Suppl Table S4a. Study design for patients in the active (G-POEM) group

	Baseline	*POD 0 – day of G-POEM*	*POD 1*	*3M*	*6M*	*12M*	*24M*	*36M*
Scintigraphy	✓			✓	-	✓		✓ (optional)
Endoscopy	✓	✓	✓ (optional)	✓	-	✓ (optional)		
GCSI + PAGI-SYM + PAGI-QoL	✓			✓	✓	✓	✓	✓
Blood tests	✓		✓					
Endoflip		✓ *Before and after G-POEM*		✓				

GCSI = Gastroparesis Cardinal Symptom Index (see also suppl table S5)
PAGI-SYM = Patient Assessment of Upper Gastrointestinal Symptom Severity Index (see also suppl table S9)
PAGI QoL = Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life (see also suppl table S10)
POD = postoperative day
3M, 6M, 12M, 24M, 36M = 3 months, 6 months, 12 months, 24 months, and 36 months visit

Suppl Table S4b. Study design for patients randomized in the sham group*

	Baseline	*POD 0*	*POD 1*	*3M*	***6M****	*POD 0/* ***G-POEM***	***POD1***	*3M*	*6M*	*12M*	*24M*	*36M*
Scintigraphy	✓			✓	-			✓				✓ (optional)
Endoscopy	✓	✓	✓ (optional)	✓	-		✓ (optional)	✓		✓ (optional)		
GCSI + PAGI-SYM, PAGI-QoL	✓			✓	✓			✓	✓	✓	✓	✓
Blood tests	✓		✓				✓					
Endoflip		✓ *Before***		✓♦		✓ *After G-POEM*		✓				

***** *At 6 months, the patients will be offered cross-over G-POEM (if symptoms persist).*

****** *In patients having undergone Endoflip during the sham procedure, no Endoflip measurement will be repeated prior to G-POEM.*

♦ *sham measurement*

GCSI = Gastroparesis Cardinal Symptom Index (see also suppl table S5)
PAGI-SYM = Patient Assessment of Upper Gastrointestinal Symptom Severity Index (see also suppl table S9)
PAGI QoL = Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life (see also suppl table S10)
POD = postoperative day
3M, 6M, 12M, 24M, 36M = 3 months, 6 months, 12 months, 24 months, and 36 months visit

Suppl Table S5. Gastroparesis Cardinal Symptom Index (GCSI)

	Symptoms	Score					
		0	1	2	3	4	5
1	Nausea	None	Very mild	Mild	Moderate	Severe	Very severe
2	Retching	None	Very mild	Mild	Moderate	Severe	Very severe
3	Vomiting	None	Very mild	Mild	Moderate	Severe	Very severe
4	Stomach fullness	None	Very mild	Mild	Moderate	Severe	Very severe
5	Not able to finish a normal sized meal	None	Very mild	Mild	Moderate	Severe	Very severe
6	Feeling extensively full after meals	None	Very mild	Mild	Moderate	Severe	Very severe
7	Loss of appetite	None	Very mild	Mild	Moderate	Severe	Very severe
8	Bloating	None	Very mild	Mild	Moderate	Severe	Very severe
9	Stomach or belly visibly larger	None	Very mild	Mild	Moderate	Severe	Very severe

The GCSI consists of nine items and three subscales to measure symptoms related to gastroparesis. The nausea/vomiting subscale consists of the following three items: nausea, retching, and vomiting. The postprandial fullness/early satiety subscale consists of the following four items: stomach fullness, inability to finish a normal-sized meal, feeling excessively full after meals, and loss of appetite. The bloating subscale consists of the following items: bloating and stomach or belly visibly larger. The GCSI total score is constructed as the average of the three symptom subscales. Its value ranges from zero meaning no symptoms to five indicating maximally severe symptomatology (see ref. No. 18-19 in the main article).

Calculation:
Total GSCI score = arithmetic mean of the three symptom subscales. Subscores = arithmetic means of (1-3), (4-7) and (8-9)

Suppl Table S6a. Inclusion criteria	
1	**Refractory** (> 6 months) and severe (based on a validated total GCSI = Gastroparesis Cardinal Symptom Index) gastroparesis, with confirmed gastric emptying based on a gastric emptying study: standardized protocol of **scintigraphy in all patients (**performed less than 6 months prior to enrolment (see ref. No. 13 in the main article), or confirmed by a **validated gastric emptying breath test** The total GCSI score must be >2.3. • Abnormal gastric emptying is defined as retention of Tc-99 m >60% at 2 h and/or ≥10% of residual activity at 4 h on a standardized sulphur colloid solid-phase gastric emptying study. • Radiolabelled liquids emptying study will be reserved as alternative technique for patients with poor tolerance of solids during scintigraphy. Abnormal gastric emptying will represent >50% retention of radiolabelled content (e.g. In-111) at 1 hour. • Abnormal gastric empyting breath test based on a solid or malrange determination for the test used (e.g. T1/2 > 109 min).
2	Severe refractory disease is defined as GCSI >2.3 and failure or recurrence in patients who received available optimal pharmacological therapies.
3	Persons 18 years or older at the time of signing the informed consent
4	Signed informed consent

Suppl Table S6b. Exclusion criteria	
1	No previous attempt with at least one prokinetic drug
2	No previous attempt to withdraw anticholinergic agents and glucagon like peptide - 1 (GLP-1) and amylin analogues* in patients treated with these substances (see ref. 1-2)
3	Active treatment with opioids or a history of treatment with opioids within 12 months before enrolment
4	Previous gastric surgery BI or II, esophagectomy, gastric pull-through
5	Previous pyloromyotomy or pyloroplasty
6	Known eosinophilic gastroenteritis
7	Organic pyloric (or intestinal) obstruction (fibrotic stricture, etc.)
8	Sever coagulopathy
9	Esophageal or gastric varices and /or portal gastropathy
10	Advanced liver cirrhosis (Child B or Child C)
11	Active peptic ulcer disease
12	Pregnancy or puerperium
13	Malignant or pre-malignant gastric diseases (dysplasia, gastric cancer, GIST): patients with a history of such disease after its cure are eligible for enrolment
14	Any other condition, which in the opinion of the investigator would interfere with study requirements
15	Uncontrolled diabetes mellitus
16	Diagnosis of rumination syndrome or “eating” disorder (mental anorexia, bulimia nervosa) **
17	Severe constipation without using laxatives
18	Inability to obtain informed consent

* Attempts to normalize glycaemic control using amylin analogues (e.g., pramlintide) or GLP-1 analogues (e.g., exenatide) may result in delayed gastric emptying.

** The presence of a rumination syndrome or eating disorders (anorexia nervosa, bulimia) is an exclusion criterion. In case of doubts, a psychiatric examination should be performed

GIST = Gastrointestinal Stromal Tumor

Suppl Table S7. Gastric emptying study protocol (GES)

Scintigraphy protocol in all patients (see ref. No. 15 in the main article, protocol endorsed by both American Neurogastroenterology and Motility Society and American Nuclear Medicine Society; 2008); less than 6 months prior to randomization. Test begun with patients under fasting conditions for a minimum of 6 hours. A radiolabelled meal was prepared by adding 0.75 mCi ^{99m}Tc-sulfur colloid into 2 the liquid egg whites. Eggs were cooked in a microwave or on a hot non-stick skillet, the eggs were stirred once or twice during cooking until firm – to the consistency of an omelette. Then, the bread was toasted and jelly spread on the toasted bread.

- Gamma camera images was obtained immediately after meal ingestion and then at 1, 2, 3 and 4 hours. The geometric mean of delay-corrected counts was used to estimate the proportion of ^{99m}Tc emptied at each time point. Diagnostic criterion for gastroparesis is defined as the percentage of gastric retention >60% at 2 h and equal to or greater than 10% at 4 h or both. Half-time (T1/2) emptying time was also be calculated. In case of poor tolerance of solids during gastric scintigraphy, radiolabelled liquids were used (see inclusion criteria, suppl table S6a). At least 72 hours before gastric emptying test, narcotics and other medications that can delay gastric emptying should be discontinued. Other alternative meals were used for patients with egg allergies or egg´s intolerance, patients with gluten-sensitive enteropathy, according to the local principles.
- Items needed for Egg Beaters Gastric Emptying Scintigraphy: 118 mL of liquid egg whites (Egg Beaters; egg substitute): 99% real eggs, cholesterol free, fat free, low calorie (120 g Egg Beater, 60 kcal, approx. two large eggs), 2 slices of wheat bread (120 kcal), Strawberry jam (30 g, 74 kcal), Water (120 mL), Technetium-99m 0.75 mCi. The subject completed the sandwich meal quickly, within max. 10 minutes. Generally, the fasting glucose in diabetic patients should be between 75 and 275 mg/dL (4.2 to 15.3 mmol/l). Diabetic patients administered their insulin with meal ingestion, generally ½ what they took normally. The nutritional composition of the meal was 69-72% carbohydrate, 22-24% protein, 2% fat and 2% fibre.

Suppl Table S8. Pyloric distensibility (Endoflip®) measurement protocol
See also **Figures S2a** and **S2b** (see ref. No. 16, 17)

The pyloric distensibility measurement was performed using the Endoflip™ 1.0 Impedance Planimetry System (Medtronic, Minneapolis, MN, USA). The Endoflip system consists of a 24 cm long 3mm outer diameter catheter with highly compliant balloon attached to its tip surrounding 16 paired impedance sensors mounted on the catheter and a solid-state pressure transducer on the distal end of the catheter within the balloon.

A single-use catheter EF-325N with 8 cm long balloon was used for all measurements. The catheter was attached to both the monitor and a syringe automatically filling the balloon with conductive fluid. Based on the principle of impedance planimetry, excitation electrodes at either end of the balloon emit a continuous low electric current, the voltage is measured across the paired impedance planimetry electrodes by leveraging Ohm's law to provide measurement of cross-sectional area at intervals based on electrode spacing. Cross-sectional area together with the pressure data from the intra-balloon pressure transducer allow to calculate resistance to distention, i.e. distensibility.

The catheter was introduced into the pylorus under direct endoscopic control, a snare or forceps were used to navigate the catheter through the pylorus if necessary. Once adequate position was achieved, with the balloon straddling the pylorus (an hourglass shape image on the Endoflip monitor), the balloon was automatically (but under direct visual supervision of the performing physician) filled with fluid from an 80mL syringe to three balloon filling volumes 30 mL, 40 mL, 50 mL. At each of these volumes the following parameters were recorded: distensibility index (mm^2/mmHg), cross-sectional area (mm^2), balloon diameter (mm), and intra-balloon pressure (mmHg). The measurements were performed in between the peristaltic waves driven by the motor migrating complex. The additional time to the procedure was approximately 10 minutes.

Suppl Table S9. PAGI-SYM score (Patient Assessment of Gastrointestinal Disorders Symptom Severity Index)

Symptoms		Score					
		0	1	2	3	4	5
		None	Very mild	Mild	Modarate	Severe	Very severe
1	Heartburn (burning pain rising in your chest or throat) during the day						
2	Regurgitation or reflux (fluid or liquid from your stomach coming up into your throat) during the day						
3	Heartburn (burning pain rising in your chest or throat) when lying down						
4	Regurgitation or reflux (fluid or liquid from your stomach coming up into your throat) when lying down						
5	Feeling of discomfort inside your chest during the day						
6	Bitter, acid or sour taste in your mouth						
7	Feeling of discomfort inside yourchest at night (during sleep time)						
8	Vomiting						
9	Nausea (feeling sick to your stomach as if you were going to vomit or throw up)						
10	Retching (heaving as if to vomit, but nothing comes up)						
11	Stomach fullness						
12	Not able to finish a normal-sized meal						
13	Feeling excessively full after meals						
14	Loss of appetite						
15	Bloating (feeling like you need to loosen your clothes)						
16	Stomach or belly visibly larger						
17	Upper abdominal (above the navel) discomfort						
18	Upper abdominal (above the navel) pain						
19	Lower abdominal (below the navel) pain						
20	Lower abdominal (below navel) discomfort						

(see ref. No. 20 in the main article)

Questionnaire was developed to measure specific symptoms of patients with upper gastrointestinal disorders. It records 20 symptoms (6 subscales) and assesses their severity within the 2 weeks prior to the test. Subscale scores are calculated by averaging across items comprising the subscale; scores vary from 0 (none or absent) to 5 (very severe). The PAGI-SYM subscale scores have good internal consistency and test-retest reliability (18).

1 - 7 = heartburn/regurgitation
8 - 10 = nausea/vomiting
11 - 14 = post-prandial fullness/early satiety
15 - 16 = bloating
17 - 18 = upper abdominal pain
19 - 20 = lower abdominal pain

Calculation:

Total PAGY-SIM score = arithmetic mean of the six symptom subscales. Subscores = arithmetic means of (1-7), (8-10), (11-14), (15-16), (17-18) and (19-20)

Suppl Table S10. PAGI – QoL questionnaire (Patient Assessment of Upper Gastrointestinal Disorders – Quality of Life)

Questions mostly related to previous 2 weeks.
Most desirable option: 5 points / Less desirable option: 0 points

Symptoms	Score					
	0	1	2	3	4	5
	None of the time	A little of the time	Some of the time	A good bit of the time	Most of the time	All of the time
During the past 2 weeks, because of your Gastrointestinal problems, how often …						
1	have you had to depend on others to do your daily activities?					
2	have you avoided performing your daily activities?					
3	have you had difficulty concentrating?					
4	has it taken you longer than usual to perform your daily activities?					
5	have you felt tired?					
6	have you lost the desire to participate in social activities such as visiting friends or relatives?					
7	have you been worried about having stomach symptoms in public?					
8	have you avoided performing physical activities or sports?					
9	have you avoided traveling?					
10	have you felt frustrated about not being able to do what you wanted to do?					
11	have you felt constricted in the clothes you wear?					
12	have you felt frustrated about not being able to dress as you wanted to?					
13	have you felt concerned about what you can and cannot eat?					
14	have you avoided certain types of foods?					
15	have you restricted eating at restaurant or at someone's home?					
16	have you felt less enjoyment in food than usual?					
17	have you felt concerned that a change in your food habits could trigger your symptoms?					
18	have you felt frustrated about not beingable to choose the food you wanted to?					
19	have you left frustrated about not being able to choose the type of beverage you wanted to?					
20	has your relationship with yours pouseor partner been disrupted?					
21	has your relationship with your children or relatives been disrupted?					
22	has your relationship with your friends been disrupted?					
23	have you been in a bad mood?					
24	have you felt depressed?					
25	have you felt anxious?					
26	have you felt angry?					
27	have you felt irritable?					
28	have you felt discouraged?					
29	have you been stressed?					
30	have you felt helpless?					

(see ref. No. 21 in the main article)

The PAGI-QoL contains 30 items with five subscales:

(1) daily activities (1 – 10)
(2) clothing (11 – 12)
(3) diet/food habits (13 – 19)
(4) relationship (20 – 22)
(5) psychological well-being and distress (23 – 30)

The PAGI-QoL questionare contains of 30 items with five subscales: (1) daily activities; (2) clothing; (3) diet/food habits; (4) relationship; and (5) psychological well-being and distress. Each items are scored on a 6-point Likert scale, with response options ranging from 0 (none) to 5 (severe problem all of the time). Subscale scores are calculated by averaging the item responses.

Calculation:

Total PAGI-QoL score = arithmetic mean of the five subscales. Subscores = arithmetic means of (1-10), (11-12), (13-19), (20-22) and (23-30)

Suppl Table S11. Detailed description of the statistical analysis

The intention to treat population and early study termination

All the main analyses were performed on the Intention To Treat (ITT) population as specified in the protocol. The ITT population includes all randomized patients and evaluates them as members of the groups to which they were originally allocated regardless of the actual treatment received or any other protocol deviations. Since some values were missing (including a complete follow-up of one patient in the sham group who withdrew consent before receiving any procedure), these values had to be imputed to recover the ITT population. The sample size for all the analyses on ITT is 41 patients as the trial was terminated early for efficacy of G-POEM in a planned interim analysis. As stated in the report of the Data and Safety Monitoring Board, the decision to stop the enrolment was adopted based on a combination of two factors: 1) The interim result was truly highly significant in favor of G-POEM with p=0.003 (the final p-value for the main outcome presented in the results is different since more follow-up data accumulated after the enrollment was stopped at the interim analysis. Unfortunately, no exact strategy for early termination was indicated in the study protocol. Therefore, the conservative Haybittle-Peto boundary was considered indicating to stop the trial at p=0.001 for any number of interim analyses. This boundary was almost reached. 2) The second factor was the risk of general anesthesia for patients undergoing the sham procedure.

Imputation of missing data and confidence intervals

The imputation of missing values was performed by the multiple imputation procedure with chained equations. We only imputed some missing values for the 41 patients enrolled in the study. We did not impute values for the remaining 45 potential patients, who were not enrolled in the study due to the early study termination. Although the amount of missing data was rather low – at most 3 values missing in any of the variables evaluated on the ITT basis – we decided on imputation to adhere to the protocol. The assumption of data missing at random (MAR) was considered plausible and given the low proportion of missing data even its partial violation would not pose a significant threat of biased results. To further prevent any suspicion that our result could be heavily influenced by the imputation, we also provide analysis of treatment success (primary outcome) on the per protocol population (1 patient with technical failure and 1 with missing GCSI excluded) and also the worst case imputation (1 technical failure in the G-POEM group as failure and 1 missing GCSI follow u in sham as success).

Multiple imputation in simple terms: The chained equations approach allows imputation of missing values using the information from the observed values. The estimates of missing values are updated iteratively which allows one to deal with missing values in all included variables. The process of imputation is random to some extent. This is further used in the multiple imputation approach. Here, multiple (e.g. 100) different random versions of imputed datasets are created. The desired analysis is performed on each realization of the dataset. Finally, the estimates of desired statistics (e.g. the median) are combined from all the imputations and their confidence intervals are constructed while reflecting variability originating both from the observed data itself and from the uncertainty of the imputation process. The resulting values are an aggregate of all the different realizations of the imputation. As a result, for example the treatment success in the sham group is 22% in 20 patients, so the value does not correspond to any of 4/20 or 5/20. This reflects the fact that the patient with missing follow-up GCSI values was assigned a treatment success in some imputations and treatment failure in the others.

Our imputation model included the following variables: age; gender; etiology of gastroparesis; baseline, 3 months, and 6 months values of total GCSI score, total PAGI-SYM score, and total QOL score; and baseline and 3 months values of 4h GES retention and GES retention halftime. We imputed the GCSI scores. The treatment success of the imputed patient was evaluated afterwards. The allocation of the patient was not used for imputation. Otherwise, the model would be strongly forced to impute high GCSI values for a patient just based on allocation into the sham group. We imputed data for the main part of the analysis separate from the cross-over data.

For the estimation of treatment success, we made 100 imputed datasets and on each used the Wilson method for construction of confidence intervals for proportions. Compared to the normal approximation approach, this method can result in non-symmetrical confidence intervals, which is very relevant for example for the GES halftime with a clearly skewed distribution. To combine the Wilson confidence intervals from all imputation datasets we used the method by Lott and Reiter (see ref. 24), which is particularly designed for this purpose.

As the primary statistics for the continuous variables we used the median since normality of the data was rejected by the Shapiro-Wilk test for at least one dataset among the compared groups and time points for each investigated variable. The confidence intervals were obtained as 2.5 % and 97.5 % quantiles from 20000 bootstrapping iterations. The same approach was used for the correlation coefficient between GCSI and gastric emptying.

Bootstrapping in simple terms: Bootstrapping is based on the idea that the distribution of observed values is the best available estimate of the actual distribution for the investigated population. Therefore, we resample the data many times (in our case 20000 times) to estimate confidence intervals for our statistic (e.g. the median). When we have N values in our sample, resampling means randomly choosing N of these values with replacement. We can imagine this as writing each of the values on a paper ticket and putting them into a hat. We randomly draw a ticket N times, but each chosen ticket is returned into the hat before another draw. As a result, the resampled dataset contains certain values multiple times and some other are not present at all. On this dataset we calculate our statistic (the median). We then take the dataset of statistics (medians) from all iterations of the resampling and estimate the confidence interval limits by discarding 2.5 % of the lowest and 2.5 % of the highest values (medians).

In our case, the process of bootstrapping had two additional steps:

1. As the number of observations in our dataset is relatively small, the median can be highly influenced by the middle values since the extreme values have no effect on the median. This can in some cases lead to underestimation of the width of the confidence intervals. We face this issue by smoothing with a Gaussian noise with sigma given by the inter quartile range of the sample divided by the square root of N.
2. We sampled the original dataset including missing values and after resampling we imputed the missing values. With this approach, both variability from the data and from the imputation are reflected in the final confidence interval.

P-values and multiple testing

As the protocol indicated a regression-based approach for the evaluation of the main outcome, we used logistic regression to calculate the only p-value presented in the manuscript for the only primary outcome (as previously specified in the protocol). All the remaining results are presented as point estimates (medians and hazard ratios) with 95 % confidence intervals with accordance to the CONSORT statement. We hope that this will prevent inadequate interpretations of the results

in terms of the multiple testing problem, which we consider likely to happen if we presented uncorrected p-values for all the other outcomes.

Technically, methods of multiple testing correction could be applied. Nevertheless, there are many strategies with different results. Primarily, the decision of which variables should be included into the analysis (defining the family of tests over which the false positive rate is to be controlled) is of major importance. The multiple testing corrections are suited for situations, where many tests are performed without a pre-defined primary hypothesis or for situations where multiple primary hypotheses are aimed to be tested simultaneously in a single trial.

We are convinced that presenting uncorrected confidence intervals for the secondary outcomes is the best option as they both show the uncertainty of the actual presented value and allow the reader to judge the single-test statistical significance. Whenever a confidence interval for a difference lies entirely below or entirely above zero, this corresponds a to statistically significant decrease or increase. As no correction for multiple testing is applied (as it is a common standard), there is 95% confidence for each individual interval to contain the true value of the population statistic (e.g. the median), but not 95% confidence that all the intervals contain their respective true values. This is presumably understandable to the reader. In contrast, by presenting all the p-values a less statistically experienced reader could be tempted to just interpret any p-value below 5 % as a clear indication of a proven effect.

Suppl Table S12. Patients treated by a respective Trial Center (in and out of the trial) and Number of Monitoring Visits

Centre No.	Randomized patients	Patients underwent G-POEM	Patients underwent sham	Patients underwent cross-over G-POEM	Patients treated outside the trial during trial period (G-POEM)	Number of monitoring visits
IKEM	33	17	15	9	7	18
Trnava	8	3	4	3	3	3
Total	41	20	19	12	10	21

G-POEM = Gastric Per Oral Endoscopic Pyloromyotomy

Suppl Table S13. Screened and excluded (not enrolled) patients

Centre No.	Screened patients	Patients uderwent GES	Patients with positive GES	Patients did not fulfill inclusion criteria	Patients fulfilled at least one exclusion criterium
IKEM	147	136	57	8	15
Trnava	42	42	18	7	4
Total	189	178	75	15	19

GES = Gastric Emptying Study (scintigraphy)

Suppl Table S14. Definition of Adverse event (AE) / Serious Adverse Event (SAE)

An adverse event (AE) is any undesirable, unintentional or unanticipated event that occurs during use of the investigational device, whether or not considered related to the therapy. A serious adverse event (SAE) is an event that is: fatal, life-threatening, results in persistent or significant disability/incapacity, requires or prolongs inpatient hospitalization, requires an intervention (endoscopy, radiology, surgery, etc.) postoperatively. Abdominal pain requiring analgetics without a need for prolongation of hospitalization was not considered as adverse event. SAE had to be reported within 24 hours to the Prague study center and the Ethics Committees / IRB if applicable. AE/AES were documented on designated CRF forms.

Report of a Adverse Event Form

Hospital visits due to follow up visits are not considered to be SAE.

□ Initial report

□ Consecutive report

Date AE start: _____ / _____ / _____ (DDMMYY)

□ Expected event □ Unexpected event

Event related to G-POEM / SHAM procedure

□ No □ Possibly □ Yes

Complication: □ Perforation □ Bleeding □ Infection □ Other

Please describe complication:

……

……

Intervention required: □ No □ Yes

Please describe intervention

……

……

Medication required: □ No □ Yes

Medication(s):

……

……

Report of a Serious Adverse Event

Hospitalization or prolongation of hospital stay required (SAE):

□ Yes □ No

If yes, please report within 24 hours to the Prague study center and Ethics Committee/IRB if applicable!

Date of hospitalisation/ - prolongation _____ (DDMMYY)

Date hospital discharge _____ (DDMMYY)

……

……

□ Event resolved □ Event ongoing

□ Long term sequela □ Death □ Unknown

Description / comment: ……

Date AE stop: _____ / _____ / _____ (DDMMYY)

Suppl Table S15. Overall incidence of adverse events

Patient	Serious / non-serious	G-POEM / Sham / cross over G-POEM	Time of AEs occurrence after the allocated procedure	Adverse Event / Serious Adverse Event	Related to procedure
01-04	Serious	G-POEM	1 month	Hospitalisation due to vomiting (not related to gastroparesis), probably food toxin	no
01-08	Non-serious	G-POEM	4 months	Mild abdominal pain without need for analgetics	no
01-10	Serious	cross over G-POEM	POD 1	Sever abdominal pain, deep ulcer of the pylorus, prolonged hospitalisation for 6 days	yes
01-11	Non-serious	G-POEM	POD 0	small periprocedural perforation of duodenal mucosae without need for intervention, no need for prolonged hospitalisation	yes
01-26	Non-serious	cross over G-POEM	POD 0	Hyperglycemia (24 mmol/L) with mild metabolic acidosis	no
01-26	Non-serious	cross over G-POEM	POD 0	Small gastric serosal perforation during G-POEM without need for intervention without sequelae	yes
01-26	Non-serious	cross over GPOEM	6 months	Non-complicated Hp- positive gastric ulcer of stomach, eradication of Hp	no
01-28	Serious	Sham	3 months	Need for hospitalisation due to severe mycotic esophagitis not allowing to eat and newly diagnosed achalasia, pneumatic dilation of achalasia, NG tube placement for feeding, prolonged hospitalisation for 23 days	no
01-28	Non-serious	cross over G-POEM	3 months	Decompensation of achalasia, mycotic esophagitis, prolonged hospitalisation for 20 days	no
01-30	Serious	G-POEM	1 month	Vomiting, need for 3 days hospitalisation, temporary nasojejunal tube placement, mycotic esophagitis	no
01-30	Non-serious	G-POEM	4 months	Feeding intolerance, hyponutrition	no
01-31	Non-serious	G-POEM	3 months	Hypoglycemia, no dumping syndrome	no
01-32	Non-serious	Sham	2 months	Recurent abdominal pain, need for opioids	no

01-32	Non serious	Sham	5 months	Vomiting, abdominal pains, administration of prokinetics and opioids	no
01-32	Serious	Sham	5 months	Hospitalisation due to vomiting for 2 days, feeding intolerance, need for nasogastric tube placement and enteral nutrition	no
01-32	Non-serious	Sham	5 months	Nausea, diarrhea, feeding intolerance, need for painkillers (opioids)	no
01-32	Non-serious	cross-over G-POEM	1 month	Severe nausea, feeding intolerance, need for administration of parenteral prokinetics	no
01-32	Serious	cross-over G-POEM	4 months	Abdominal pains, weightloss, feeding intolerance, nasogastric tube placement, hospitalisation for 6 days, acute urinary retention, pains of ears.	no
01-35	Serious	G-POEM	POD 0	During G-POEM mucosal injury, prolonged hospitalisation for precautionary reasons for 7 days, no need for intervention or specific treatment	yes
01-38	Serious	G-POEM	3 months	Hospitalisation for 30 days due to hypocalcemia, examination before transplantation, diarrhea, hypoglycemia - confirmed dumping syndrome	yes
01-40	Serious	Sham	3 months	Hospitalisation for 1 day, abdominal pains, nausea	no
01-41	Serious	G-POEM	4 months	Hospitalisation for 6 days, due to intestinal infection – gastroenteritis	no

POD = postoperative day
Hp = Helicobacter pylori

Suppl Table S16. Summary of adverse events (AE)			
	G-POEM	**Sham**	**Cross-over G-POEM**
Serious AE – n			
Hospitalisation (required or prolonged) related to procedure	2	0	1
Need for additional endoscopic, radiological or surgical intervention	0	0	0
Hospitalisation (required or prolonged) not related to procedure	3	3	1
Live-threatening events	0	0	0
Death	0	0	0
Overall	**5**	**3**	**2**
Overall SAEs related to procedure	**2**	**0**	**1**
Overall SAEs not related to procedure	**3**	**3**	**1**
Non-serious AE – n			
Abdominal pain (not related to procedure)	1	1	0
Periprocedural serosal perforation	1	0	1
Nausea or vomiting, feeding intolerance (not related to procedure)	1	2	1
Decompansation of achalasia with mycotic esophagitis	0	0	1
Hypoglycemia/hyperglycemia	1	0	1
Gastric ulcer	0	0	1
Overall	**4**	**3**	**5**
Overall AEs related to procedure	**1**	**0**	**1**
Overall AEs not related to procedure	**3**	**3**	**4**

Suppl Table S17. Need for analgesics administration after G-POEM, sham procedure or cross-over G-POEM

	G-POEM	Sham	Cross-over
Number n (%)	10 (41%)	2 (10%)	4(33%)
Total number of procedures	21	20	12

Postprocedural pain necessitating administration of analgesics on postoperative days 0 or 1 was not considered as adverse event but rather a standard part of the postoperative course like with other similar procedures.

Suppl Table S18. Procedure details. The analysis was performed on the available data, one procedure length was missing. There was one technical failure of G-POEM procedure, which is included into the analysis but no closure was used in this case.

Procedure length	**G-POEM** n=21	**Cross-over G-POEM** n=12	**Sham** n=19
Mean	76 min	58 min	55 min
Standard deviation	41 min	17 min	9 min
Median	61 min	56 min	55 min
Minimal	35 min	40 min	40 min
Maximal	185 min	91 min	76 min
Length of myotomy	**G-POEM**	**Cross-over G-POEM**	**Sham**
Mean	27 mm	27 mm	-
Standard deviation	7 mm	4 mm	-
Median	30 mm	30 mm	-
Minimal	25 mm	20 mm	-
Maximal	30 mm	30 mm	-
Hospitalization after procedure	**G-POEM**	**Cross-over G-POEM**	**Sham**
Mean	1.9 days	2.4 days	1 day
Standard deviation	1.4 days	1.3 days	0 days
Median	1.5 days	2 days	1 day
Minimal	1 day	1 day	1 day
Maximal	7 days	6 days	1 day
Technical success	95% (20/21)	100% (12/12)	NA
Closure with endoclips (n)	9	8	NA
Closure with endoscopic suturing system, (n)	11	4	NA
Need for capnoperitoneum puncture	No	No	No
Other gas related adverse events	No	No	No
Anesthesia related adverse events	No	No	No

NA = not applicable

Suppl. Table S19. Treatment success for the primary outcome, sensitivity analysis and etiology subgroups. In this trial, treatment success was defined as a reduction by 50% from baseline GCSI for the primary G-POEM and sham procedure and as a reduction by 50% from the 6 months visit (after the sham procedure) for the cross-over G-POEM. In addition, table shows treatment success rates if the treatment success had been defined as a decrease of GCSI by 1 point (a common definition of treatment success). Subgroup analysis in different etiologies of gastroparesis after cross-over G-POEM was not performed because of small numbers of patients. For the ITT population, one of the 41 values (2 %) was multiply imputed (in the sham group). For the worst case scenario, we assumed treatment failure for the G-POEM patient with technical failure and treatment success in the sham patient with missing GSCI data.

Treatment success rate [%] (95% CI) at 6 months	**G-POEM**	N	**Sham**	N	**Cross-over G-POEM**	N
ITT population, GCSI reduction by 50 %	71 (50 – 86)	21	22 (8 – 47)	20	75 (47 – 91)	12
PP population, GCSI reduction by 50 %	70 (48 – 85)	20	21 (9 – 43)	19	75 (47 – 91)	12
Worst case scenario, GCSI reduction by 50 %	67 (45 – 83)	21	25 (11 – 47)	20	75 (47 – 91)	12
Diabetic etiology (ITT, reduction by 50 %)	89 (56 – 98)	9	17 (3 – 57)	8	Not applicable	
Post-surgical etiology (ITT, reduction by 50 %)	50 (18 – 82)	6	29 (7 – 67)	7	Not applicable	
Idiopathic etiology (ITT, reduction by 50 %)	67 (30 – 90)	6	20 (3 – 67)	5	Not applicable	
ITT population, GCSI reduction by 1 point	95 (76 – 99)	21	37 (19 – 60)	20	75 (47 – 91)	12
Treatment success rate [%] (95% CI) at 3 months	**G-POEM**	N	**Sham**	N	**Cross-over G-POEM**	N
ITT population, GCSI reduction by 50 %	57 (36 – 76)	21	22 (8 – 47)	20	58 (32 – 81)	12
PP population, GCSI reduction by 50 %	55 (34 – 74)	20	21 (9 – 43)	19	58 (32 – 81)	12
Worst case scenario, GCSI reduction by 50 %	52 (32 – 72)	21	25 (11 – 47)	20	58 (32 – 81)	12
Diabetic etiology (ITT, reduction by 50 %)	67 (35 – 88)	9	17 (3 – 57)	8	Not applicable	
Post-surgical etiology (ITT, reduction by 50 %)	33 (9 – 72)	6	43 (15 – 76)	7	Not applicable	
Idiopathic etiology (ITT, reduction by 50 %)	67 (30 – 90)	6	0 (0 – 43)	5	Not applicable	
ITT population, GCSI reduction by 1 point	76 (55 – 89)	21	42 (23 – 64)	20	67 (39 – 86)	12

ITT – intention to treat population (all patients evaluated according to their allocation, missing data multiply imputed)

PP – per protocol population (only patients following the study protocol)

GCSI – gastroparesis cardinal symptom index

N – number of patients in a given group

Suppl. Table S20. Evolution of variables in time. The table presents estimates of medians of various quantities at different time points in the study and differences between time points. The differences are calculated on a single patient level. The confidence intervals (CI) are not corrected for multiple testing. The analysis was performed on the ITT population with 21, 20, and 12 patients in the G-POEM, sham, and cross-over G-POEM groups, respectively. In total, 2 GCSI values (1 %), 3 PAGI-SYM values (2 %), 7 PAGI-QOL values (5 %) and 10 GES values (5 %) were imputed across all groups and time points.

	Values at visits			**Decrease from baseline***	
Variable – median (95% CI)	**Baseline**	**3 months**	**6 months**	**to 3 months**	**to 6 months**
GCSI – G-POEM	3.5 (3.2 – 3.7)	1.4 (0.9 – 1.9)	1.1 (0.5 – 1.5)	2.3 (1.3 – 2.6)	2.4 (2.0 – 2.8)
GCSI – sham	3.2 (2.8 – 3.4)	2.5 (1.9 – 3.1)	2.5 (1.9 – 3.2)	0.8 (0.1 – 1.2)	0.7 (0.0 – 1.2)
PAGI-SYM – G-POEM	2.7 (2.0 – 3.0)	0.9 (0.7 – 1.4)	0.7 (0.5 – 1.2)	1.5 (1.0 – 1.9)	1.5 (1.2 – 2.0)
PAGI-SYM – sham	2.8 (2.5 – 3.0)	2.0 (1.5 – 2.8)	2.0 (1.5 – 2.6)	0.7 (0.1 – 1.1)	0.5 (0.1 – 1.1)
PAGI-QoL – G-POEM	2.1 (1.7 – 2.5)	1.6 (0.9 – 2.5)	0.8 (0.6 – 1.5)	0.3 (-0.5 – 0.9)	1.1 (0.1 – 1.6)
PAGI-QoL – sham	2.5 (1.5 – 2.9)	1.9 (1.2 – 2.7)	1.7 (1.2 – 2.4)	0.4 (-0.2 – 0.7)	0.4 (-0.1 – 0.8)
BMI [kg/m2] – G-POEM	22 (19 – 26)	22 (20 – 25)	22 (21 – 26)	-0.4 (-1.2 – 0.5)	-0.7 (-1.8 – 0.2)
BMI [kg/m2] – sham	26 (21 – 28)	24 (21 – 27)	24 (21 – 28)	-0.4 (-1.0 – 0.4)	-0.7 (-1.2 – 0.4)
GES 4h retention [%] – G-POEM	22 (17 – 31)	12 (5 – 22)		12 (3 – 19)	
GES 4h retention [%] – sham	26 (18 – 39)	24 (11 – 35)		6 (-7 – 19)	
GES ret. halftime [min] – G-POEM	152 (127 – 185)	95 (77 – 118)		53 (5 – 94)	
GES ret. halftime [min] – sham	157 (128 – 263)	110 (82 – 158)		49 (-3 – 144)	

	Values at visits			**Decrease from baseline ***	
Variable – median (95% CI)	**Baseline (= 6 months visit after sham)‡**	**3 months after cross-over**	**6 months after cross-over**	**to 3 months after cross-over**	**to 6 months after cross-over**
GCSI – cross-over G-POEM	2.8 (2.5 – 3.7)	1.1 (0.7 – 1.9)	1.0 (0.6 – 1.7)	1.9 (1.1 – 2.4)	2.1 (1.3 – 2.6)
PAGI-SYM – cross-over G-POEM	2.2 (1.9 – 3.0)	0.8 (0.6 – 1.6)	0.5 (0.4 – 1.8)	1.3 (1.0 – 2.0)	1.6 (0.8 – 2.2)
PAGI-QoL – cross-over G-POEM	2.2 (1.3 – 3.3)	1.8 (0.9 – 2.6)	1.6 (0.7 – 2.3)	0.5 (-0.1 – 1.5)	0.3 (-0.1 – 1.6)
BMI [kg/m2] – cross-over G-POEM	22 (19 – 26)	22 (19 – 27)	22 (20 – 27)	0.0 (-1.0 – 0.9)	-0.2 (-1.1 – 0.5)
	Baseline (= 3 months visit after sham)‡	**3 months after cross-over**		**to 3 months after cross-over**	
GES 4h ret. [%] – cross-over G-POEM	24 (11 – 38)	7 (1 – 14)		13 (5 – 23)	
GES ret. halftime [min] – cross-over G-POEM	138 (83 – 178)	66 (32 – 154)		80 (29 – 179)	

* The table presents a decrease, so positive values indicate reduction of the score/measurement.
‡ For the cross-over procedure, values obtained at 6 months visit (at 3 months in case of gastric emptying study) after the sham procedure are considered as baseline value
GCSI = Gastroparesis Cardinal Symptom Index (see also suppl table S5)
PAGI-SYM = Patient Assessment of Upper Gastrointestinal Symptom Severity Index (see also suppl table S9)
PAGI QoL = Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life (see also suppl table S10)
GES = Gastric Emptying Study
ITT = Intention To Treat

Suppl. Table S21. Evolution of GCSI sub-scores in time. Means of GCSI subscales are presented at different time points in the study and differences between time points. The differences are calculated on a single patient level. The Nausea / vomiting subscale comprises of the questions 1 to 3, Fullness of questions 4 to 7 and Bloating of questions 8 and 9 (see Table S5). The confidence intervals (CI) are not corrected for multiple testing. The analysis was performed on the available data basis with N=21 for G-POEM, N=19 for sham, and N=12 for cross-over G-POEM.

	Values at visits			**Decrease from baseline***	
Variable – mean (95% CI)	**Baseline**	**3 months**	**6 months**	**to 3 months**	**to 6 months**
G-POEM					
Nausea / vomiting	3.3 (2.8 – 3.7)	1.3 (0.7 – 1.9)	0.8 (0.4 – 1.2)	2.1 (1.5 – 2.6)	2.5 (2.1 – 3.0)
Fullness	3.6 (3.4 – 3.9)	1.7 (1.3 – 2.0)	1.2 (0.8 – 1.7)	2.0 (1.6 – 2.4)	2.4 (2.0 – 2.9)
Bloating	3.5 (3.0 – 4.0)	1.5 (0.9 – 2.0)	1.4 (0.9 – 1.9)	2.0 (1.5 – 2.6)	2.1 (1.6 – 2.6)
Sham					
Nausea / vomiting	3.0 (2.5 – 3.4)	2.0 (1.4 – 2.4)	1.8 (1.4 – 2.4)	1.1 (0.7 – 1.7)	1.2 (0.6 – 1.8)
Fullness	3.4 (3.1 – 3.7)	3.0 (2.6 – 3.4)	2.9 (2.5 – 3.4)	0.4 (-0.2 – 0.9)	0.5 (0.1 – 1.0)
Bloating	3.3 (2.6 – 3.8)	2.6 (1.9 – 3.3)	3.0 (2.3 – 3.7)	0.7 (0.4 – 1.1)	0.3 (-0.3 – 1.0)
	Values at visits			**Decrease from 6 months after sham***	
Variable – mean (95% CI)	**Baseline (= 6 months visit after sham)‡**	**3 months after cross-over**	**6 months after cross-over**	**to 3 months after cross-over**	**to 6 months after cross-over**
Cross-over G-POEM					
Nausea / vomiting	2.2 (1.6 – 2.8)	0.7 (0.3 – 1.3)	0.6 (0.2 – 1.1)	1.5 (1.0 – 2.1)	1.6 (1.0 – 2.3)
Fullness	3.5 (3.1 – 3.9)	1.4 (0.8 – 2.0)	1.3 (0.9 – 1.6)	2.1 (1.4 – 2.8)	2.3 (1.8 – 2.7)
Bloating	3.5 (2.7 – 4.2)	1.9 (1.4 – 2.4)	1.5 (0.9 – 2.2)	1.7 (1.2 – 2.0)	2.0 (1.4 – 2.7)

* The table presents decrease, so positive values indicate reduction of the score/measurement.

‡ For the cross-over procedure, values obtained at 6 months visit after the sham procedure were considered as baseline values.

GCSI – gastroparesis cardinal symptom index

Suppl. Table S22. Endoflip® measurements – primary G-POEM and cross-over G-POEM combined. Means of pyloric distensibility are presented at different time points in the study and differences between time points. The differences are calculated on a single patient level. The confidence intervals (CI) are not corrected for multiple testing. The table presents only available data; the imputation model was not used as over 50% of data is missing because the measurement of pyloric distensibility was added after beginning of the trial.
Note that at pre-procedure, post-procedure, and follow-up time points there were 16, 17, and 15 (14 for filling volume 50 mL) values available. There were 14 values for the pre vs. post treatment difference and 12 values for the pre vs. follow-up difference.

	Values at visits			**Increase from pre-procedure** ‡	
Variable – mean (95% CI)	**Pre-G-POEM**	**Post-G-POEM**	**3 months follow-up**	**to post-G-POEM**	**to 3 months follow-up**
DI [mm2/mmHg] 30 mL filling	6.8 (5.2 – 8.4)	12.6 (10.3 – 14.9)	10.2 (8.6 – 11.8)	7.4 (6.0 – 9.0)	5.2 (3.4 – 7.5)
DI [mm2/mmHg] 40 mL filling	7.6 (6.0 – 9.3)	12.7 (11.4 – 14.3)	13.1 (11.3 – 15.7)	5.4 (3.7 – 7.0)	8.0 (5.5 – 10.2)
DI [mm2/mmHg] 50 mL filling	9.1 (6.5 – 12.4)	11.6 (9.5 – 14.1)	10.3 (8.2 – 12.4)	2.6 (0.7 – 4.4)	3.6 (1.2 – 6.0)
CSA [mm2] 30 mL filling	91 (75 – 107)	128 (114 – 142)	142 (111 – 176)	50 (35 – 64)	35 (10 – 58)
CSA [mm2] 40 mL filling	144 (125 – 165)	199 (177 – 219)	206 (185 – 234)	66 (36 – 99)	64 (37 – 83)
CSA [mm2] 50 mL filling	216 (180 – 247)	291 (267 – 319)	279 (246 – 306)	92 (63 – 120)	66 (18 – 110)

‡ The estimates of change are based only on cases where both relevant values were available. Therefore, the expected median difference does not have to correspond to the difference in medians for the corresponding visits. Please note that increase of both DI and CSA at all three filling volumes are significant.

DI – distensibility index
CSA – cross-sectional area

Supp. Table S23. List of pre-specified primary and secondary outcomes and post-hoc outcomes and other analyses with references	
Primary outcome	
Treatment success at 6 months	Table S19, Figures 2 (main document), S3
Secondary outcomes	
Treatment success at 3 months	Table S19, Figure S4
Treatment success in per-protocol population	Table S19, Figures 2 (main document), S3, S4
Treatment success in etiology sub-groups	Table S19, Figures 2 (main document), S3, S4
Treatment success predictors	Table S2 (main document)
GCSI score	Tables 1 (main document), S20, S21, Figures 3 (main document), S5, S6
PAGI-SYM score	Table S20, Figure S7
PAGI-QoL score	Tables 1 (main document), S20, Figure S8
GES 4h retention	Tables 1 (main document), S20, Figures 4 (main document), S9
GES retention halftime	Table S20, Figure S10
BMI	Tables 1 (main document), S20
Endoflip® DI and CSA (pyloric distensibility)	Table S22, Figure S11
Adverse events	Tables S14, S15, S16
Need for analgetics (pain analysis)	Table S17
Post hoc analyses	
GCSI by sub-scores	Table S21, Figure S6
GCSI and GES correlation at 3 months	Figure S12
Other analyses	
Baseline Demographic and Clinical characteristics	Table 1 (main document)
Procedure details	Table S18
Patients treated in centres in and out of the Trial	Tables S12
Screened and enrolled patients	Tables S13

GCSI = Gastroparesis Cardinal Symptom Index

PAGI-SYM = Patient Assessment of Gastrointestinal Disorders Symptom Severity Index

PAGI-QoL = Quality of Life Questionnaire

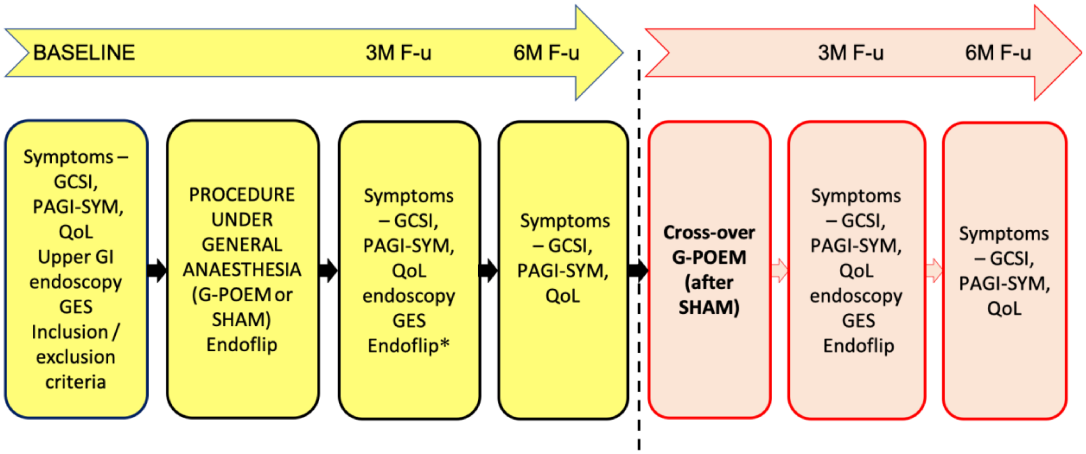
GES = Gastric Emptying Study

BMI = Body Mass Index

DI = distensibility

CSA= Cross-Sectional Area

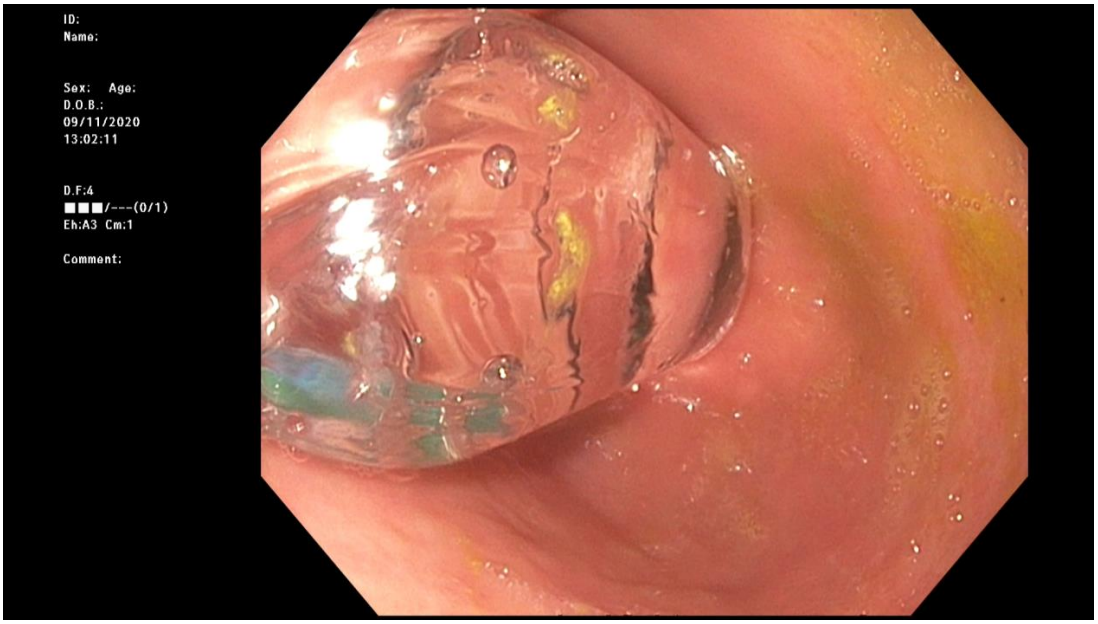
SUPPLEMENTARY FIGURES



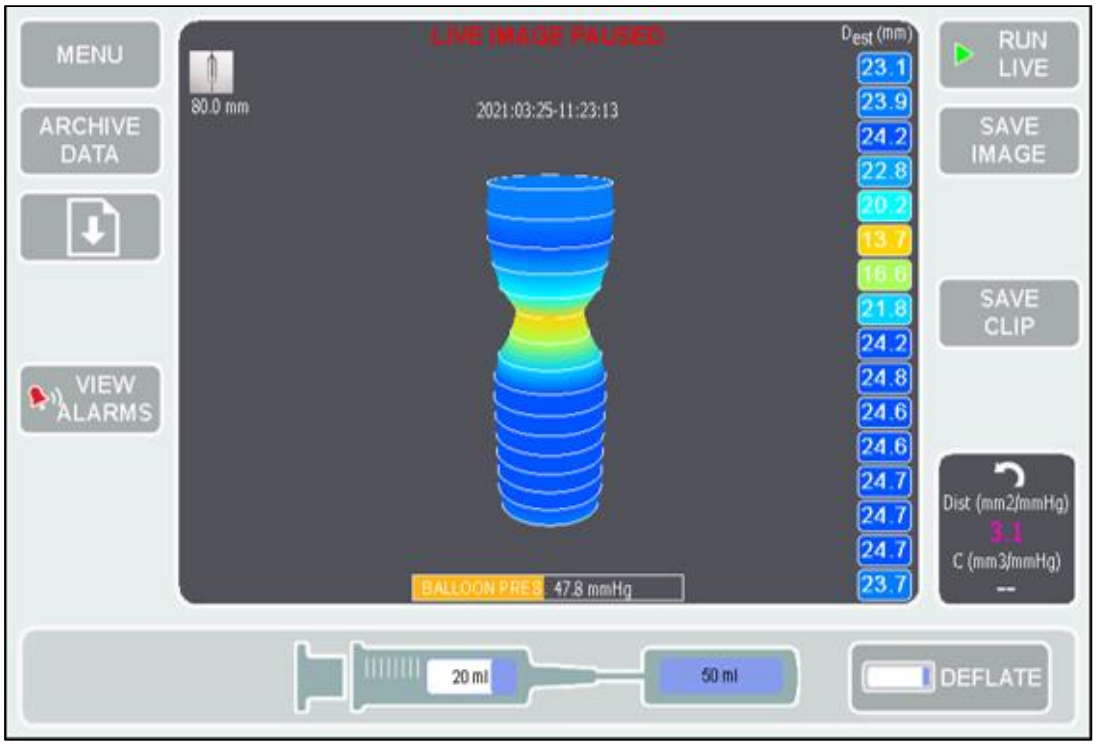
Suppl Figure S1. Study design showing the study course after randomisation (yellow boxes) and for patients who underwent cross-over G-POEM

* *In patients having undergone Endoflip during the sham procedure, no Endoflip measurement was repeated prior to G-POEM.*

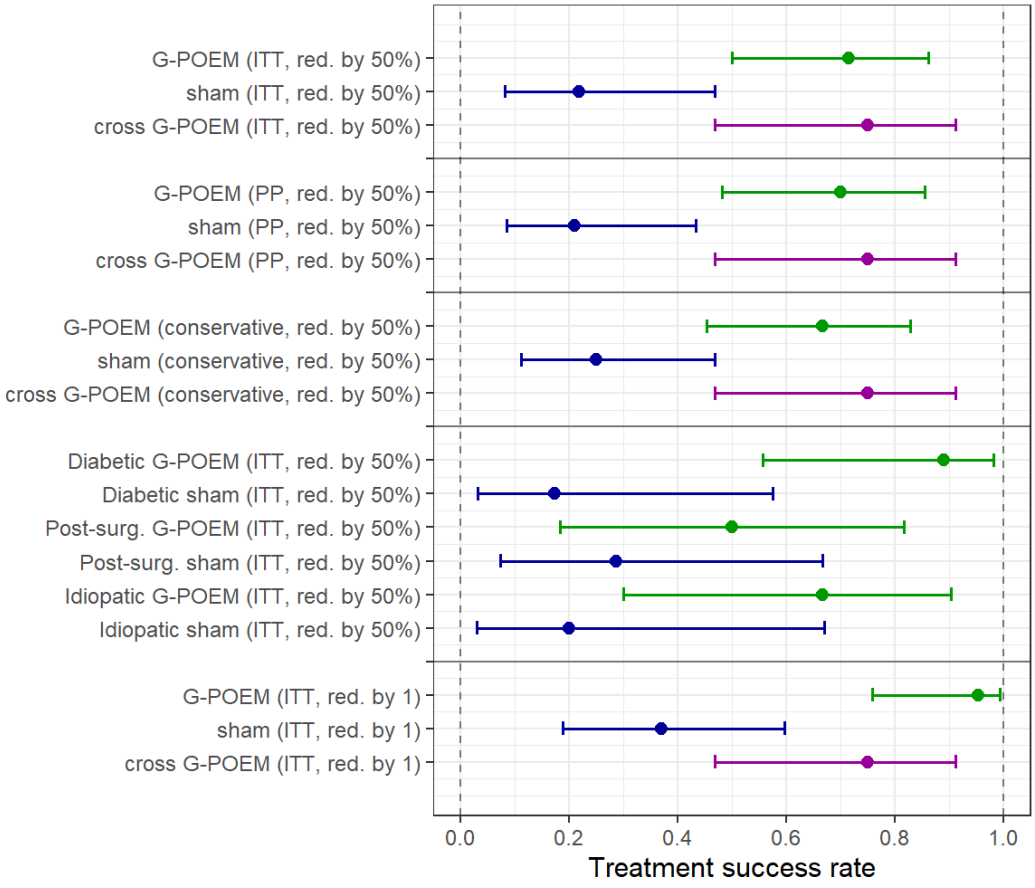
GCSI = Gastroparesis Cardinal Symptom Index (see also suppl table S5)
PAGI-SYM = Patient Assessment of Upper Gastrointestinal Symptom Severity Index (see also suppl table S9)
PAGI QoL = Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life (see also suppl table S10)
GES = Gastric Emptying Study
3M, 6M = 3 months, 6 months visit



Suppl Figure S2a. Measurement of pyloric distensibility. A balloon is introduced through the pylorus under endoscopic control and inflated automatically. Figure shows endoscopic image during measurement



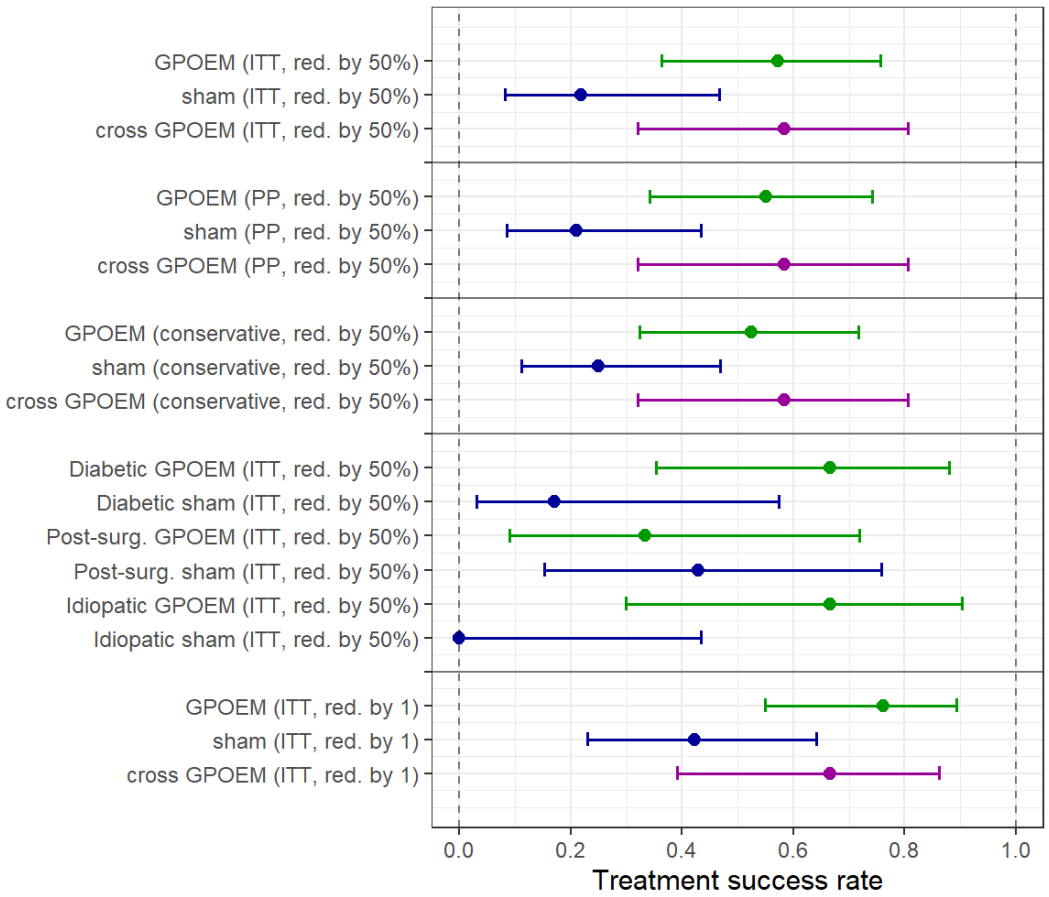
Suppl Figure S2b. Measurement of pyloric distensibility. An hourglass shape image on the Endoflip monitor during measurement. The narrowed place in the picture points to a pylorus. In the right down corner a value shows pyloric distensibility (3.1mm^2/mmHg)



Suppl Figure S3. Treatment success at 6 months after procedure, from top to bottom:

- the main outcome on the intention to treat (ITT) population with treatment success defined as reduction of the total GCSI score by 50% form baseline,
- treatment success evaluated on the per-protocol (PP) population (for cross-over the ITT population and PP population are the same),
- treatment success evaluated with the most conservative approach (worst case scenario), where the patient with technical failure of G-POEM is assigned failure and the sham patient who withdraw consent is assigned success (note, that overlap of confidence intervals does not exclude significant difference, which is 42% with 95% CI: 9% to 74% not containing zero),
- treatment success in sub-groups defined by etiology of gastroparesis,
- treatment success on the ITT population defined as reduction of the total GCSI score by 1 point from baseline.

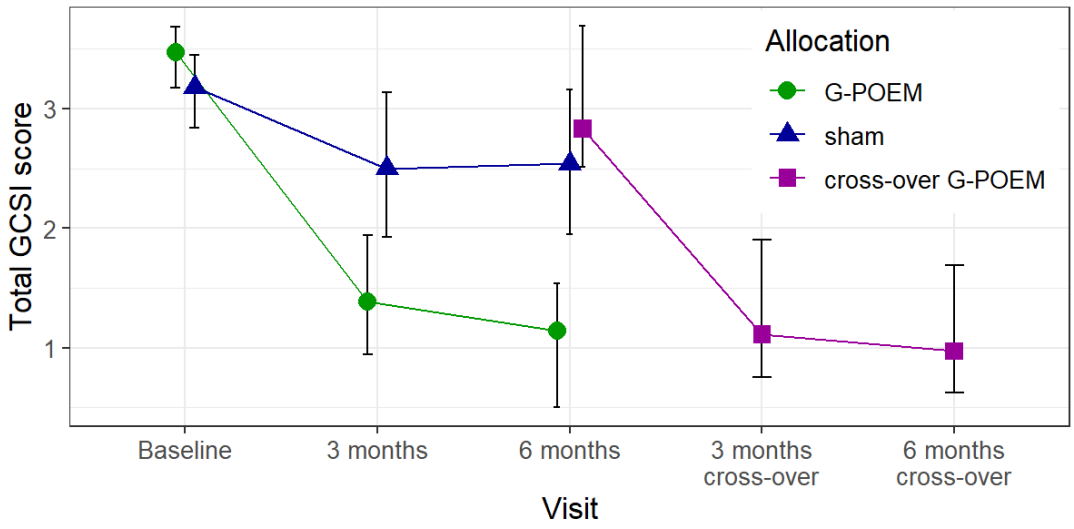
The results analyzed on the intention to treat (ITT) population (N=41, $N_{Di\text{-}G\text{-}POEM}$=9, $N_{Di\text{-}Sham}$=8, $N_{PS\text{-}G\text{-}POEM}$=6, $N_{PS\text{-}Sham}$=7, $N_{Id\text{-}G\text{-}POEM}$=6, $N_{Id\text{-}Sham}$=5, 1 value (2 %) imputed in diabetic GP patient in the sham group) are supplemented by the main outcome analysis on the per protocol (PP) population (N=39).



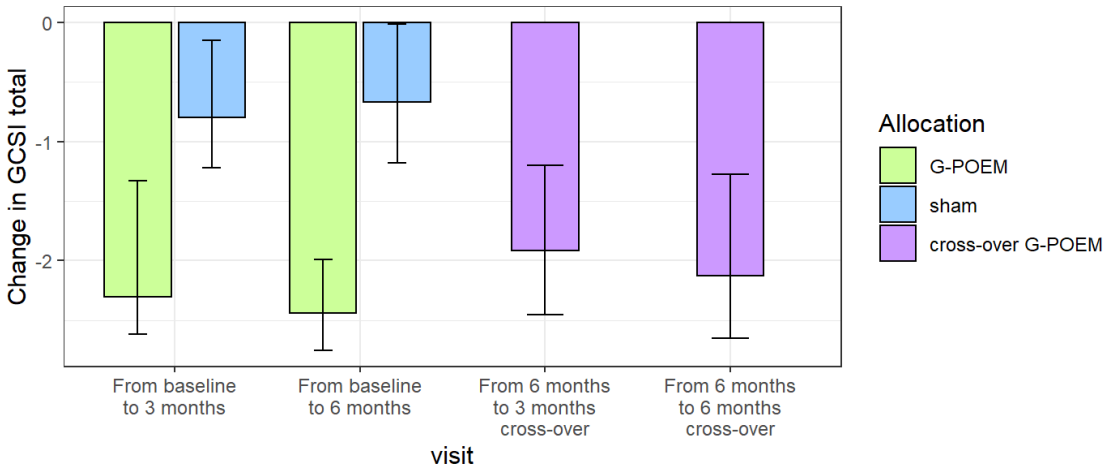
Suppl Figure S4. Treatment success 3 months after procedure, from top to bottom:

- treatment success in the G-POEM, sham and cross-over arms on the intention to treat (ITT) population with treatment success defined as reduction of the total GCSI score by 50% form baseline,
- treatment success evaluated on the per-protocol (PP) population,
- treatment success evaluated with the most conservative approach (worst case scenario), where the patient with technical failure of G-POEM is assigned failure (despite having success) and the sham patient who withdraw consent is assigned success,
- treatment success in sub-groups defined by etiology of gastroparesis (not evaluated for cross-over G-POEM due to low number of patients in groups),
- treatment success on the ITT population defined as reduction of the total GCSI score by 1 point form baseline.

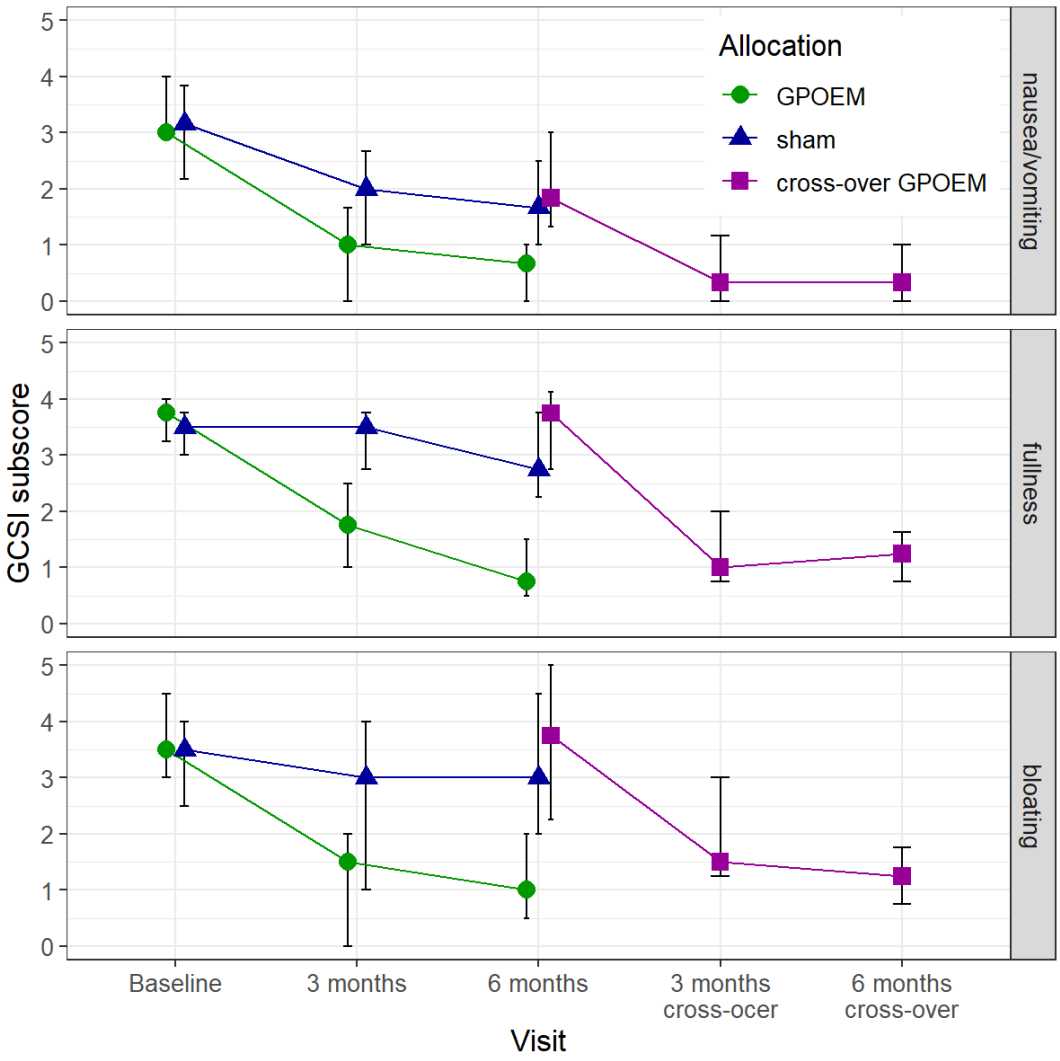
The results analyzed on the intention to treat (ITT) population (N=41, $N_{Di\text{-}G\text{-}POEM}$=9, $N_{Di\text{-}Sham}$=8, $N_{PS\text{-}G\text{-}POEM}$=6, $N_{PS\text{-}Sham}$=7, $N_{Id\text{-}G\text{-}POEM}$=6, $N_{Id\text{-}Sham}$=5, 1 value (2 %) imputed in diabetic GP patient in the sham group) are supplemented by the analysis on the per protocol (PP) population (N=39).



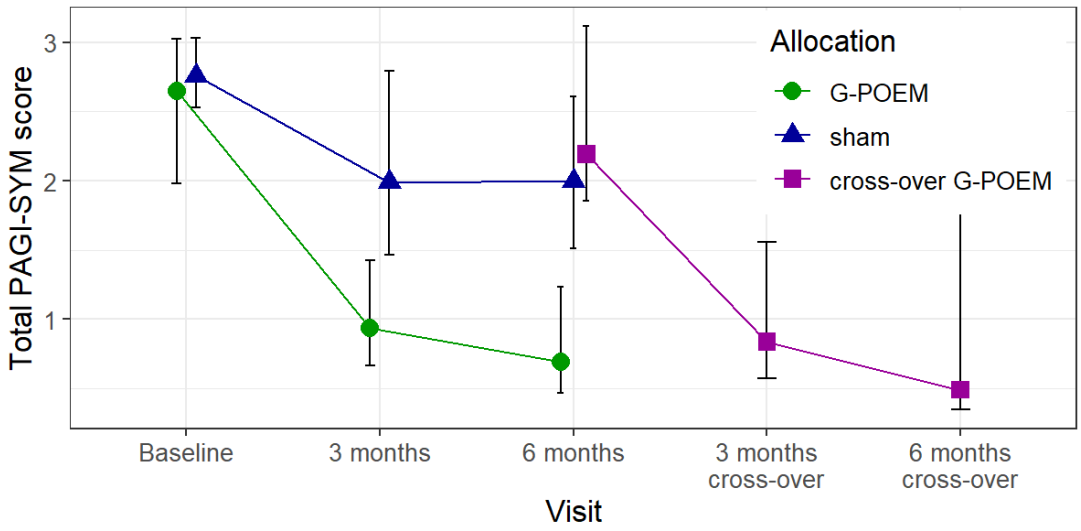
Suppl Figure S5a. Evolution of the GCSI total score. Point estimates of medians with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green circles, N=21), sham procedure (blue triangles, N=20, imputed 1 value (5 %) for 3 months and 1 value (5 %) for 6 moths), and cross-over G-POEM procedure (purple squares, N=12). For the cross-over G-POEM group the value at 6 months reflects only the data for the patients in this group (who subsequently underwent the cross-over G-POEM procedure). Points are connected for visual aid.



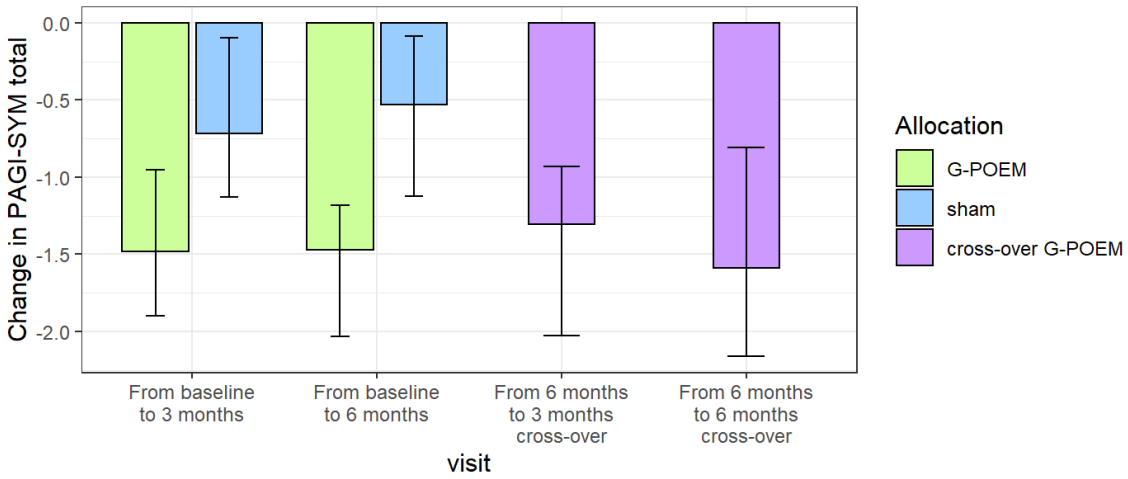
Suppl Figure S5b. Changes of the GCSI total score between visits. Point estimates of medians of differences between the specified visits with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green, N=21), sham procedure (blue, N=20, imputed 1 value (5 %) for 3 months and 1 value (5 %) for 6 months), and cross-over G-POEM procedure (purple, N=12) (patients who subsequently underwent the cross-over G-POEM procedure).



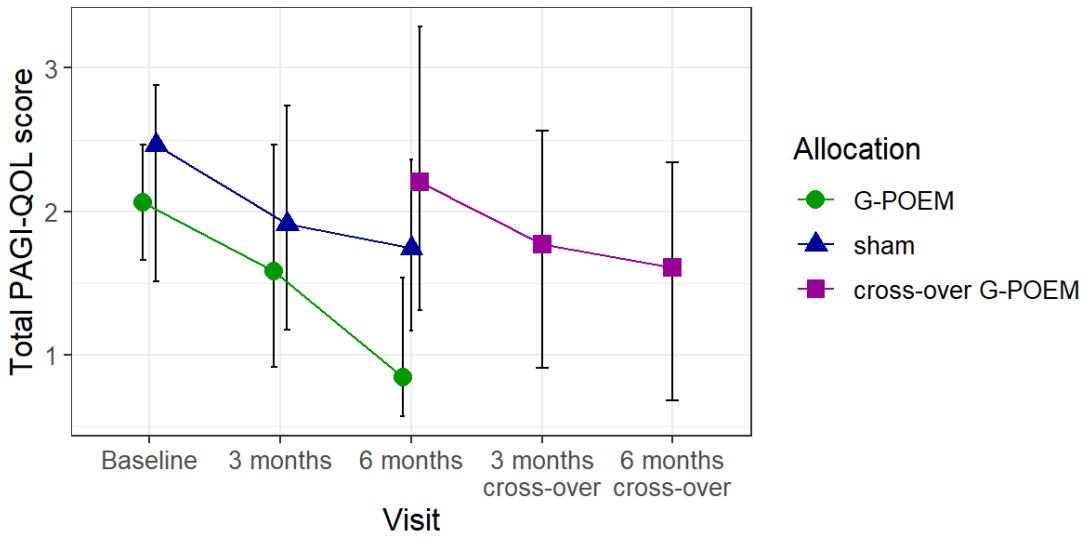
Suppl Figure S6. Evolution of the GCSI sub-scores. Point estimates of medians with 95% confidence intervals calculated on the available data are shown for patients after the G-POEM procedure (green circles, N=21), sham procedure (blue triangles, N=19), and cross-over G-POEM procedure (purple squares, N=12). For the cross-over G-POEM group the value at 6 months reflects only the data for the patients in this group (patients who subsequently underwent the cross-over G-POEM procedure). Points are connected for visual aid. The nausea / vomiting subscale comprises of the questions 1 to 3, Fullness of questions 4 to 7 and Bloating of questions 8 and 9, see Table S5.



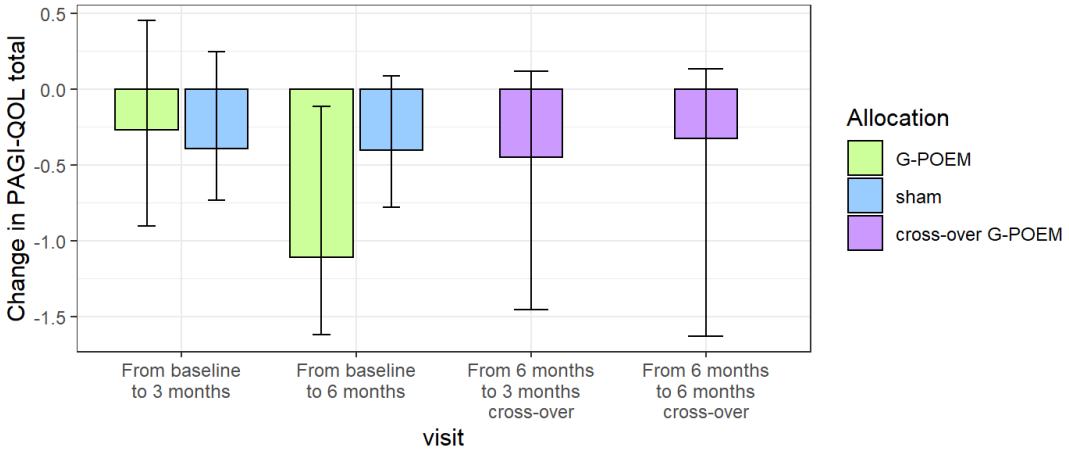
Suppl Figure S7a. Evolution of the PAGI-SYM total score. Point estimates of medians with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green circles, N=21), sham procedure (blue triangles, N=20, imputed 1 value (5 %) for 3 months and 1 value (5 %) for 6 months), and cross-over G-POEM procedure (purple squares, N=12). For the cross-over G-POEM group the value at 6 months reflects only the data for the patients in this group (patients who subsequently underwent the cross-over G-POEM procedure). Points are connected for visual aid.



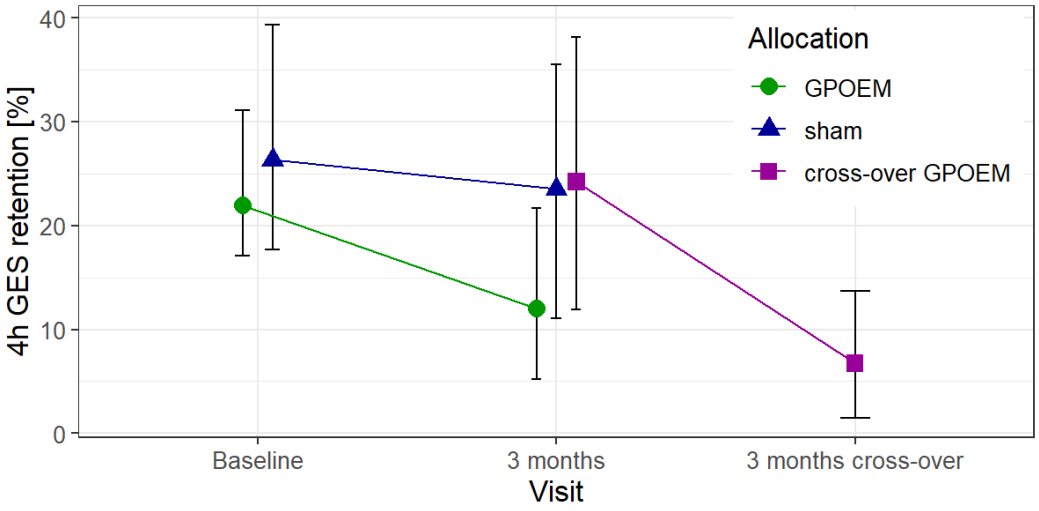
Suppl Figure S7b. Changes of the PAGI-SYM total score between visits. Point estimates of medians of differences between the specified visits with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green, N=21), sham procedure (blue, N=20, imputed 1 value (5 %) for 3 months and 1 value (5 %) for 6 months), and cross-over G-POEM procedure (purple, N=12) (patients who subsequently underwent the cross-over G-POEM procedure).



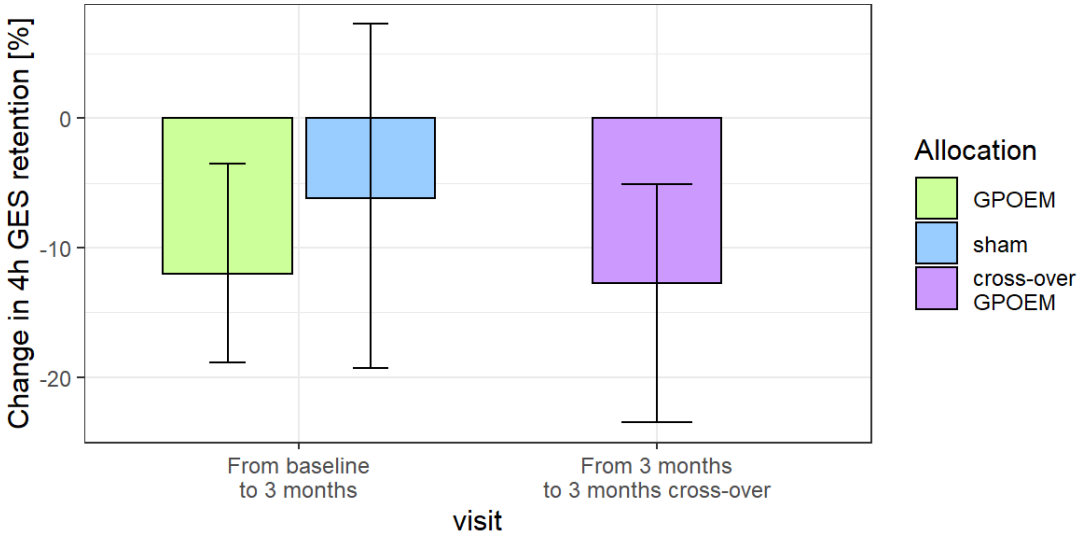
Suppl Figure S8a. Evolution of the PAGI-QoL total score. Point estimates of medians with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green circles, N=21, imputed 1 value (5 %) for 3 months and 1 value (5 %) for 6 months), sham procedure (blue triangles, N=20, imputed 1 value (5 %) for baseline, 1 value (5 %) for 3 months , and 1 value (5 %) for 6 months), and cross-over G-POEM procedure (purple squares, N=12). For the cross-over G-POEM group the value at 6 months reflects only the data for the patients in this group (patients who subsequently underwent the cross-over G-POEM procedure). Points are connected for visual aid.



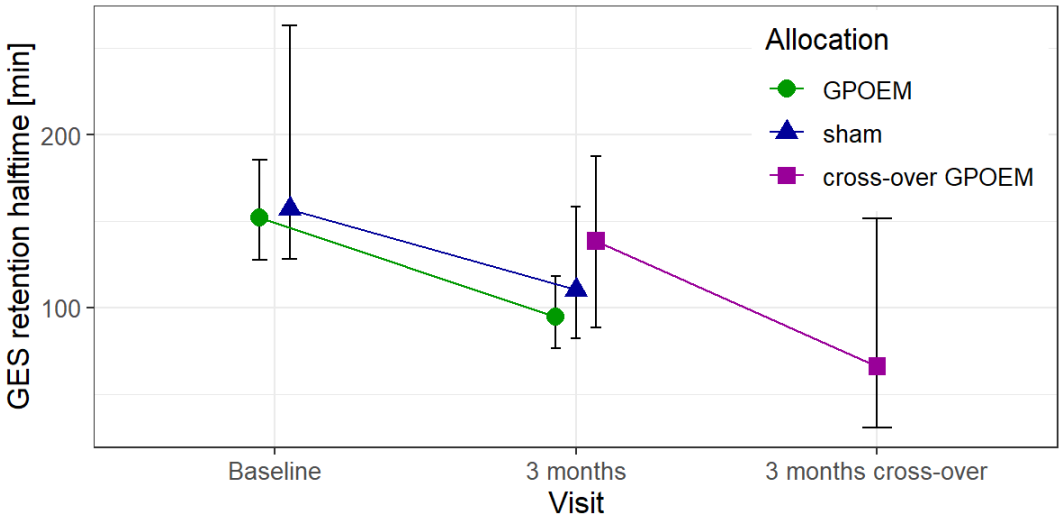
Suppl Figure S8b. Changes of the PAGI-QoL total score between visits. Point estimates of medians of differences between the specified visits with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green, N=21, imputed 1 value (5 %) for 3 months and 1 value (5 %) for 6 months), sham procedure (blue, N=20, imputed 1 baseline value (5 %), 1 value (5 %) for 3 months, and 1 value (5 %) for 6 months), and cross-over G-POEM procedure (purple, N=12) (patients who subsequently underwent the cross-over G-POEM procedure).



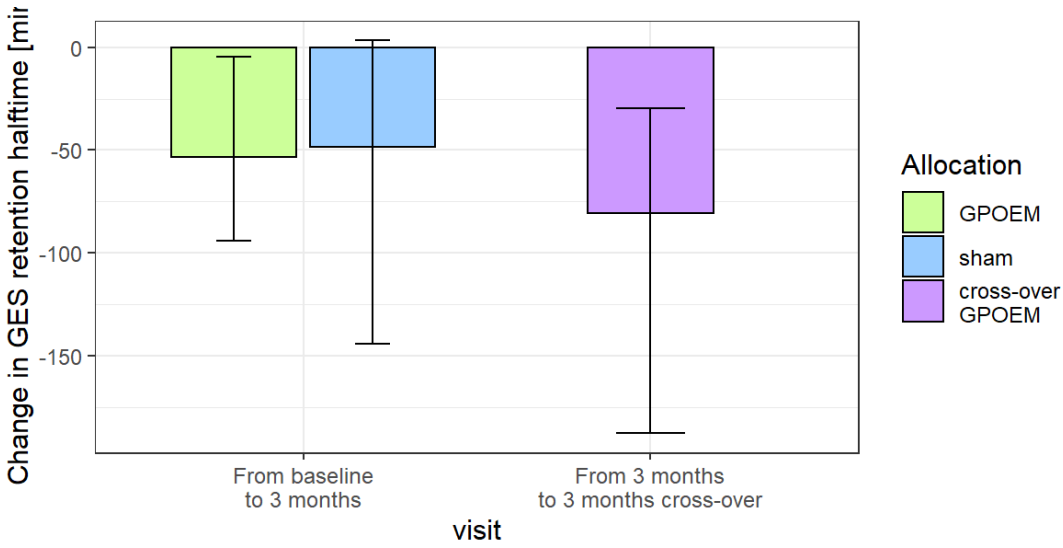
Suppl Figure S9a. Evolution of the GES 4h retention. Point estimates of medians with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green circles, N=21, imputed 2 values (10 %) for 3 months), sham procedure (blue triangles, N=20, imputed 1 value (5 %) for 3 months), and cross-over G-POEM procedure (purple squares, N=12). For the cross-over G-POEM group the value at 3 months reflects only the data for the patients in this group (patients who subsequently underwent the cross-over G-POEM procedure). Points are connected for visual aid.



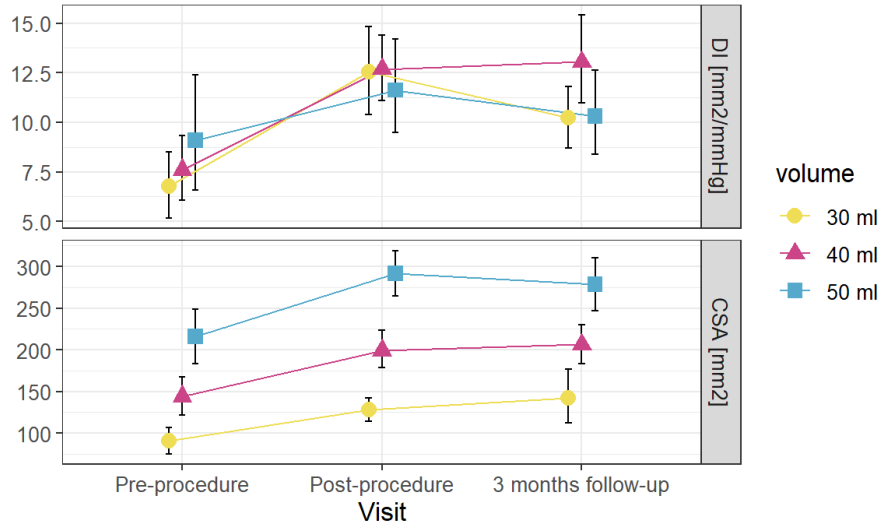
Suppl Figure S9b. Changes of GES 4h retention between visits. Point estimates of medians of differences between the specified visits with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green, N=21, imputed 2 values (10 %) for 3 months), sham procedure (blue, N=20, imputed 1 value (5 %) for 3 months), and cross-over G-POEM procedure (purple, N=12) (patients who subsequently underwent the cross-over G-POEM procedure).



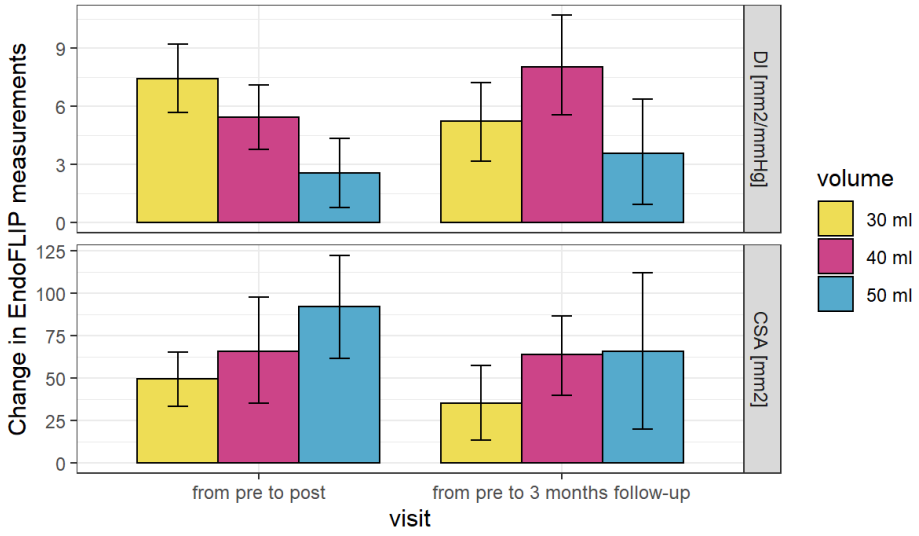
Suppl Figure S10a. Evolution of GES retention halftime. Point estimates of medians with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green circles, N=21, imputed 1 value (5 %) for 3 months), sham procedure (blue triangles, N=20, imputed 1 baseline value (5 %) and 1 value (5 %) for 3 months), and cross-over G-POEM procedure (purple squares, N=12). For the cross-over G-POEM group the value at 3 months reflects only the data for the patients in this group (patients who subsequently underwent the cross-over G-POEM procedure). Points are connected for visual aid.



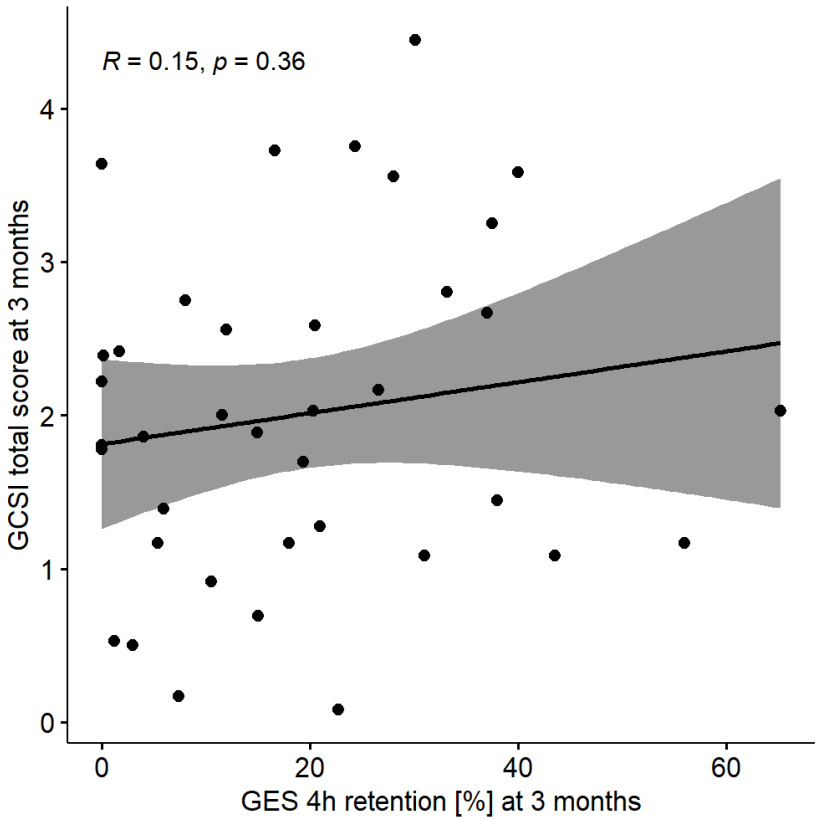
Suppl Figure S10b. Changes of GES retention halftime between visits. Point estimates of medians of differences between the specified visits with 95% confidence intervals calculated on the ITT population are shown for patients after the G-POEM procedure (green, N=21, imputed 1 value (5 %) for 3 months), sham procedure (blue, N=20, imputed 1 baseline value (5 %) and 1 value (5 %) for 3 months), and cross-over G-POEM procedure (purple, N=12) (patients who subsequently underwent the cross-over G-POEM procedure).



Suppl Figure S11a. Evolution of pyloric distensibility measurements (Endoflip) for different filling volumes – primary G-POEM and cross-over G-POEM combined**.** Point estimates of means for distensibility index (DI, top panel) and cross-sectional area (CSA, bottom panel) with 95% confidence intervals are shown for 30 mL (yellow circles), 40 mL (magenta triangles), and 50 mL (cyan squares) balloon fillings. The figure presents only available data; the imputation model was not used as for pre-procedure, post-procedure, and follow-up time points a total of 16, 17, and 15 (14 for 50 mL) values were available - this measurement was added after beginning of the trial. Points are connected for visual aid.



Suppl Figure S11b. Changes of measurements of pyloric distensibility by Endoflip for different filling volumes between visits – primary G-POEM and cross-over G-POEM combined**.** Point estimates of means of differences between the specified visits for distensibility index (DI, top panel) and cross-sectional area (CSA, bottom panel) with 95% confidence intervals are shown for 30 mL (yellow circles), 40 mL (magenta triangles), and 50 mL (cyan squares) balloon fillings. The figure presents only available data; the imputation model was not used as only 14 values were available for pre vs. post treatment difference and 12 for the pre vs. follow-up difference. The measurement of pyloric distensibility was added after beginning of the trial)



Suppl Figure S12. Correlation between GCSI total score and GES 4h retention at 3 months. All points for available data are plotted (no imputation performed) along with the linear regression line (black) and the corresponding confidence interval area (gray). The fact that also a decreasing line can be placed into the gray area indicates that there is no significant correlation. Correlation at 6 months can not be shown as GES was not measured at 6 months.

SUPPLEMENT

Finantial Support

a) The trial was finantially supported by a Grant 17-28797A from the Czech Ministry of Health
b) Medtronic provided for free Endoflip balloons for measurement of pyloric distensibility

This suplement contains the following items:

Martinek J, *et al. Gut* 2022;0:1–9. doi: 10.1136/gutjnl-2022-326904

A Randomized, Sham and Cross-Over-Controlled Trial of per-oral endoscopic pyloromyotomy (G-POEM) in patients with refractory gastroparesis

1. Administrative information:

Principal Investigators: Assoc. Prof. Jan Martínek & Prof. Thomas Rösch
Study nurse & Data management: Gabriela Petráňová
Phone: +420 26136 2617, email: gapc@ikem.cz, copy: jan.martinek@volny.cz

Contact for scientific queries
Email: jan.martinek@volny.cz
Email: t.roesch@uke.de

Contact for public queries: …….
Countries of recruitment: Czech Republic, Germany, Belgium, Nederland, Great Britain, Sweden, Denmark, Slovak Republic, USA, Romania.
While the study is ongoing, further centres might be included.
Commetties: (appendix 1)

ClinicalTrials.gov Identifier: NCT03356067
Date of registration:29.11.2017

Study type: Interventional
Allocation: randomized
Intervention model: parallel block assignment
Masking: sham-controlled

Date of first enrolment: December 2017
Target sample size: 86 patients

Recruitment status: Recruiting

Primary outcome(s): Improvement of Gastroparesis Cardinal Symptom Index after endoscopic intervention in patients with severe gastroparesis
Duration of follow-up: 36M

Protocol version:

Issue date: 29.08.2019
Protocol Amendment Number: 02
Author(s): Prof. J.M., Prof T.R

Date	
….. 29.09.2017	…… Version 3
….. 30.05.2018	…… Amendment 1
….. 11.11.2019	…… Amendment 2
….. 24.09.2020	…… Amendment 3

Sources and types of financial, material, and other support.

There is no official founding for this multicenter international study and each center is responsible for its cost related to this study.

Table 1:

#	Institution, Department	Address, email	Main investigators
1	Department of Hepatogastroenterology, Institute for Clinical and Experimental Medicine, Prague, Czech Republic	Vídeňská 1958/9, 140 21 Prague 4 jan.martinek@volny.cz, vackova.zuz@gmail.com tomas.hucl@ikem.cz jusp@ikem.cz	Jan Martinek Zuzana Vackova Tomas Hucl Julius Spicak
2	University Medical Center Hamburg- Eppendorf | UKE Department for Interdisciplinary Endoscopy, Germany	Neues Klinikum O10, Martinistr 52 20246 Hamburg t.roesch@uke.de	Thomas Rösch
3	Translational Research in GastroIntestinal Disorders, Leuven, Belgium	O&N I Herestraat 49 - box 701, 3000 Leuven jan.tack@kuleuven.be raf.bisschops@uzleuven.be tim.vanuytsel@uzleuven.be	Jan Tack Raf Bishop Tim Vanuytsel
4	King's Institute of Therapeutic Endoscopy, London, UK	Denmark Hill, London, SE5 9RS amynhaji@nhs.net	Amyn Haji
5	III. Medizinische Klinik, Medical Center/Klinikum Augsburg, Germany	Stenglinstr. 2, 86156, Augsburg, helmut.messmann@klinikum-augsburg.de	Helmut Messmann
6	Department of Hepatogastroenterology at Cliniques universitaires St-Luc, Brussels, Belgium	Ave Hippocrate 10, 1200 Brussels, Belgium, pdeprez@uclouvain.be hubert.piessevaux@uclouvain.be	Pierre H. Deprez Hubert Piessevaux
7	Department of Surgical Gastroenterology, Karolinska University Hospital, Stockholm, Sweden	171 76 Solna Stockholm, Karolinska Universitetssjukhuset koshi.kumagai@gmail.com jon.tsai@ki.se	Koshi Kumaga Jon A. Tsai
8	Department of Gastroenterology and Hepatology, Amsterdam, The Netherlands	University of Amsterdam, Meibergdreef 9, 1105 AZ p.fockens@amc.uva.nl a.j.bredenoord@amc.uva.nl	Paul Fockens Arjan Bredenoord
9	Department of Internal Medicine, University Hospital Trnava, Slovak Republic	A Zarnova 11, 91701 Trnava Slovak Republic rhustak@gmail.com jan.usak@gmail.com	Jan Martinek Rastislav Hustak Jan Usak
10	The Department of Surgical Gastroenterology L, Denmark	Nřrrebrogade 44, 8000 Aarhus C rasti.kunda@aarhus.rm.dk	Rastislav Kunda
11	Jeesenius Faculty of Medicine in Martin, Clinic of Gastroenterological Internal Medicine, Slovak Republic	Kollárova 2, 036 59 Martin, michaldemeter@gmail.com hyrdel@jfmed.uniba.sk martin.duricek@gmail.com	Michal Demeter Rudolf Hyrdel Martin Duricek Peter Banovcin

		pbanovcin@gmail.com	
12	Center for Endoscopic and Therapeutics Research, The University of Chicago, USA	5841 S. Maryland Avenue Chicago, IL 60637 iwaxman@uchicago.edu	Irving Waxman
13	Regional Institute of Gastroenterology, Cluj-Napoca, Romania	Strada General Traian Moșoiu 38A, Cluj-Napoca, matantau@gmail.com	Marcel Tantau
14	Department 2nd Dept. of Internal Medicine – Gastroenterology, University Hospital in Hradec Kralove	Sokolska 581, 500 05 Hradec Kralove, Czech Republic tacheci@gmail.com rejchrt@lfhk.cuni.cz marcela.kopacova@fnhk.cz	Ilja Tachecí Stanislav Rejchrt Marcela Kopacová
15	Klinik für Gastroenterologie und Hepatologie, University of Zurich	Rämistrasse 100, 8091 Zürich, Switzerland Daniel.Pohl@usz.ch	Daniel Pohl (EndoFLIP technology supervisor)
16			

2. Abbreviations:

AE – Adverse Event
sAE – serious Adverse Events
ANMS – American Nuclear Medicine Society
BI, BII – Billroth
BMI – Body Mass Index
CRP – C-Reactive Protein
DM – Diabetes Mellitus
DSMB – Data and Safety Monitoring Board
ESD – Endoscopic Submucosal Dissection
GCSI – Gastroparesis Cardinal Symptom Index
GEBT – Gastric Emptying Breath Test
GET – Gastric Emptying Study
GE – Gastric Emptying
GI – Gastrointestinal
GIST – Gastrointestinal Stromal Tumor
GLP – 1 - Glucagon Like Peptide -1
GP – Gastroparesis
G-POEM – per oral endoscopic pyloromyotomy
IE – Interim Event Report
IRB – Institutional Review Boards
PAGI-SYM - Patient Assessment of Upper Gastrointestinal Symptom severity index
PAGI-QoL – Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life
POEM – Per-Oral Endoscopic Myotomy
Pyloromyotomy
POD – Post Operative Day
PPI – Proton Pump Inhibitors
RBC, WBC – Red Blood Cells, White Blood Cells
QoL – Quality of life

3. Study Summary

Gastroparesis is a disorder triggered by numerous causes and it is defined by symptoms and with an objective evidence of delayed gastric emptying in the absence of obstruction (albeit pyloric spasms may play a role in a subset of patients). Gastroparesis may be a consequence of medication, surgery or diabetes but in approximately one third of patients, the cause remains unknown and the patients are diagnosed with idiopathic gastroparesis. Effective treatment for gastroparesis is challenging especially in patients with severe symptoms. The efficacy of prokinetics is dubious since they have not proven real clinical efficacy in placebo-controlled trials. In refractory gastroparesis, endoscopic or surgical treatments may therefore be considered. Endoscopic treatments include intrapyloric injection of botulinum toxin and transpyloric insertion of a metallic stent. Surgical options involve implantation of a gastric „pacemaker“ (gastric stimulation), pyloroplasty and subtotal gastrectomy. The partial effectiveness of botulinum toxin injection, stents and pyloroplasty suggests that disruption of the pyloric muscle may lead to a decreased intrapyloric tone and consequently to a symptomatic improvement in some patients with refractory gastroparesis.

Recently, a new endoscopic technique, gastric endoscopic per oral pyloromyotomy (G-POEM) has been introduced with promising preliminary results. Uncontrolled studies with so far limited number of patients have demonstrated a significant symptomatic improvement in approximately 70% of patients and improved or normalized of gastric emptying in more than a half of patients after G-POEM. A prospective uncontrolled study suggested that patients with idiopathic or post-surgical gastroparesis experiences higher success rate after G-POEM (70-80%) compared to patients with diabetic gastroparesis (50%).

G-POEM is, in principle, adaptation of POEM (per-oral endoscopic myotomy) in the stomach. POEM is now considered a standard treatment for esophageal achalasia and it has been shown to be safe and effective. In contrast to achalasia, pathophysiology of pyloric function in patients with gastroparesis is less understood and the explanation of how and why G-POEM should work is some-how hypothetical. For example, presumed pylorospasm has not been demonstrated as the predictive factor for treatment success of G-POEM yet. Refractory gastroparesis is often accompanied by psychological or even psychiatric disturbances and hence a placebo“ effect of G-POEM cannot be ruled out. Therefore, the real clinical efficacy of G-POEM can only be demonstrated in a clinical randomized sham-controlled trial.

To assess the severity of gastroparesis-related symptoms, the Gastroparesis Cardinal Symptom Index (GCSI) has been developed for this item. The GCSI is part of a larger questionnaire PAGI-SYM (Patient Assessment of Upper Gastrointestinal Symptom severity

index) established for assessment of patient-reported symptoms in gastroparesis (dyspepsia and gastroesophageal reflux) [1]. PAGI-SYM as well as GCSI subscale scores varied significantly by global disease severity, with higher (worse) scores observed in those subjects who rated their gastroparesis as moderate to severe.

The aim of this prospective, sham-controlled, cross-over study (cross-over for patients randomized to the sham arm) is to compare short and long-term efficacy and safety of G-POEM in patients with refractory gastroparesis. Symptoms and objective parameters of gastric emptying will be the main outcome criteria. The reason of using a sham protocol is to control for the potential confounders (therapeutic effects of touch and belief, which are components of the placebo effect**).**

We plan to randomize 86 patients (43 in the active arm, ratio 1:1 active vs. sham). Sample size is calculated based on expected therapeutic success of G-POEM in 50% of patients vs. 20% in the sham group; significance level 0,05; study power 0,8; beta error 0,2; adjustment for 15% expected drop out.

Patients will be randomized in blocks of 6, stratified according to the etiologies: (idiopathic, diabetic, and post-surgical; patients after esophagectomy with gastric pull-through will not be included). Control visits will be scheduled at 3, 6, 12, 24, and 36 months. The primary outcome will be the proportion of patients with treatment success in the active group vs. sham group at 6 months after the procedure. Several secondary outcomes will also be assessed, including procedure-related parameters and safety parameters and change in Gastric Emptying Study (GET) after G-POEM vs. sham. After 6 months, patients randomized to the sham group will be offered G-POEM procedure and further followed up (cross-over part of the study) providing that they did not have a therapeutic effect of the sham procedure.

Martinek J, *et al. Gut* 2022;0:1–9. doi: 10.1136/gutjnl-2022-326904

4. Introduction

Gastroparesis (GP) is relatively common gastrointestinal (GI) motility disorder, defined as epigastric symptoms associated with delayed gastric emptying (GE) in the absence of a mechanical obstruction. The prevalence of gastroparesis is unknown due to the difficulties inherent of undertaking true population-based studies. In a large study, the age-adjusted incidence of gastroparesis was 2.4 per 100,000 person-years for men and 9.8 per 100,000 person-years for women [2]. Women are more commonly affected than men [3]. In clinical practice, idiopathic and diabetic gastroparesis are the most common causes, each accounting for about one third of the patients [4]. While traditionally gastroparesis has been mainly associated with type 1 DM, the global rise of obesity-related diabetes has resulted in a high proportion of gastroparesis patients with type 2 DM. The cumulative incidence of developing gastroparesis in type 1 DM is 5.2% over 10 years and 1% in type 2 DM. Gastroparesis may also develop as a consequence of gastric or abdominal surgery or may accompany some other neurological, infectious, and infiltrative disorders [4-7]. Symptoms are not specific and may be mild, moderate or severe and include nausea, vomiting, dyspeptic symptoms, regurgitation, weight loss and poor nutritional status.

Diagnosis of gastroparesis should be confirmed by an objective gastric emptying study. Gastric scintigraphy has been considered as the gold standard for the evaluation of gastric emptying. The most reliable parameters for diagnosis of gastroparesis is gastric retention of solids at 4 h after standardized food ingestion and a half-time (T1/2) of gastric evacuation [8]. More recently, Gastric Emptying Breath Test (GEBT) has been validated for the diagnosis of delayed gastric emptying, and has gained increasing acceptance now that FDA has approved gastric emptying breath test [9].

Effective treatment for gastroparesis is a real clinical challenge especially in patients with severe symptoms. Dietary measures and drugs (prokinetics, antiemetics etc.) have limited efficacy [10, 11]. If conservative measures do not help (= refractory gastroparesis), endoscopic or surgical therapies may be considered with the main aim to decrease the tonus of pyloric sphincter. Two endoscopic methods have been studied: (1) intrapyloric botulinum toxin injection is only partially effective [12, 13], and a systematic review did not confirm its clinical effectiveness compared to placebo [14]; (2) trans-pyloric stent placement may be effective but it provides only a short-term effect (stent must be removed and eventually re-inserted), and there is a considerable risk of migration [15, 16].

The surgical method of choice for treatment of refractory gastroparesis is a laparoscopic pyloroplasty according to Heineke-Miculicz [17, 18]. Two studies reported symptomatic

improvement in more than 80% of patients; however there is a risk of dumping syndrome after pyloromyotomy [16]. A longstanding experimental approach for treatment of refractory gastroparesis represents gastric stimulation, but in spite of almost 2 decades of research, the benefit of this method is still controversial [19, 20].

Traditionally, gastroparesis has been considered as a disorder caused by gastric hypomotility and the role of pyloric muscle might have been underestimated. However, recent studies have shown that pyloric pressure is elevated in a subset of patients with gastroparesis and, therefore, a pylorospasm may be an underlying cause (or an additional pathophysiological factor) of delayed gastric emptying [21]. Treatments targeting the pyloric muscle leading to its decreased tone may therefore provide a therapeutic effect.

In 2007 Pasricha et al. published experimental endoscopic esophageal myotomy by using a submucosal tunnelling technique [22]. In 2008, prof. Inoue performed the first human per-oral endoscopic myotomy (POEM) in a patient with achalasia. At present, POEM is considered as a standard treatment modality for esophageal achalasia [23-25].

Based on favourable experiences with POEM, it is conceivable that "gastric modification of POEM", so called G-POEM (gastric per-oral endoscopic pyloromyotomy), may be beneficial in patient with refractory gastroparesis. Khashab et al. performed the first human G-POEM in one patient with severe gastroparesis with a significant symptomatic improvement [26]. A French group reported promising results of G-POEM in 23 patients. In this study, G-POEM was effective (symptomatic improvement in 70% of subjects) and safe (no serious adverse events). Patients with responded to G-POEM seemed to be more effective in patients with idiopathic or post-surgical gastroparesis compared to patients with diabetic gastroparesis. (). Another multi-centre analysis of 30 patients also showed promising results of G-POEM in patients with refractory gastroparesis [26]. Thus, endoscopic pyloromyotomy seems promising mini-invasive method for the treatment of (at least some) patients with severe refractory gastroparesis. However, larger studies comparing this new method with other treatment modalities or with a "sham" procedure are necessary to establish a real potential of G-POEM for treatment of this disease.

Martinek J, *et al. Gut* 2022;0:1–9. doi: 10.1136/gutjnl-2022-326904

5. Hypothesis & specific aims

This study intends to assess the clinical efficacy and safety of G-POEM in patients with refractory gastroparesis in a randomized, cross-over, sham controlled trial.

The null hypothesis to be tested (refused): G-POEM has the comparable efficacy to the sham procedure in patients with refractory gastroparesis.

6. Methods

6.1 Patient recruitment, in- and exclusion criteria

Patients will be prospectively recruited from all participating centres. Inclusion and exclusion criteria are listed in Table 2. All main etiologies of gastroparesis are eligible for enrolment (e.g. idiopathic, diabetic and post-surgical).

Table 2: In- and exclusion criteria

Inclusion criteria:	
1	**Refractory** (> 6 months) and severe (based on a validated total GSCI = Gastroparesis Cardinal Symptom Index) gastroparesis, with confirmed gastric emptying based on a gastric emptying study: standardized protocol of **scintigraphy in all patients** (performed less than 6 months prior to enrolment), or confirmed by a **validated gastric emptying breath test** [27]. The total GSCI score must be >2.3 [28]. • Abnormal gastric emptying is defined as retention of Tc-99 m >60% at 2 h and/or ≥10% of residual activity at 4 h on a standardized sulphur colloid solid-phase gastric emptying study. • Radiolabelled liquids emptying study will be reserved as alternative technique for patients with poor tolerance of solids during scintigraphy. Abnormal gastric emptying will represent >50% retention of radiolabelled content (e.g. In-111) at 1 hour. • Abnormal gastric empyting breath test based on a solid normal range determination for the test used (e.g. T1/2 > 109 min)
2	Severe refractory disease is defined as GCSI >2.3 and failure or recurrence in patients who received available optimal pharmacological therapies.
3	Persons 18 years or older at the time of signing the informed consent
4	Signed informed consent

Exclusion criteria	
1	No previous attempt with at least one prokinetic drug
2	No previous attempt to withdraw anticholinergic agents and glucagon like peptide -1 (GLP-1) and amylin analogues* in patients treated with these substances
3	Active treatment with opioids or a history of treatment with opioids within 12 months before enrolment.
4	Previous gastric surgery BI or II, esophagectomy, gastric pull-through
5	Previous pyloromyotomy or pyloroplasty
6	Known eosinophilic gastroenteritis
7	Organic pyloric (or intestinal) obstruction (fibrotic stricture, etc.)
8	Sever coagulopathy
9	Esophageal or gastric varices and /or portal gastropathy
10	Advanced liver cirrhosis (Child B or Child C)
11	Active peptic ulcer disease
12	Pregnancy or puerperium
13	Malignant or pre-malignant gastric diseases (dysplasia, gastric cancer, GIST): patients with a history of such disease after its cure are eligible for enrolment
14	Any other condition, which in the opinion of the investigator would interfere with study requirements
15	Uncontrolled diabetes mellitus
16	Diagnosis of rumination syndrome or "eating" disorder (mental anorexia, bulimia nervosa) **
17	Severe constipation without using laxatives
18	Inability to obtain informed consent

* Attempts to normalize glycaemic control using amylin analogues (e.g., pramlintide) or GLP-1 analogues (e.g., exenatide) may result in delayed gastric emptying [8].

***** The presence of a rumination syndrome or eating disorders (anorexia nervosa, bulimia) is an exclusion criterion. In case of doubts, a psychiatric examination should be performed***

6.2 Questionnaire(s) (Appendix 2)

Patients will be asked to complete validated questionnaires throughout the study to assess severity of symptoms related to gastroparesis.

6.2.1 GCSI score

The GCSI consists of nine items and three subscales to measure symptoms related to gastroparesis [1]. **The nausea/vomiting subscale** consists of the following three items: nausea, retching, and vomiting. **The postprandial fullness/early satiety subscale** consists of the following four items: stomach fullness, inability to finish a normal-sized meal, feeling excessively full after meals, and loss of appetite. **The bloating subscale consists of the**

following items: bloating and stomach or belly visibly larger. The GCSI total score is constructed as the average of the three symptom subscales.

6.2.2 PAGI-SYM score

Questionnaire was developed to measure specific symptoms of patients with upper gastrointestinal disorders. It records 20 symptoms (6 subscales) and assesses their severity within the 2 weeks prior to the test. Subscale scores are calculated by averaging across items comprising the subscale; scores vary from 0 (none or absent) to 5 (very severe). The PAGI-SYM subscale scores have good internal consistency and test-retest reliability [29].

6.3 Pre-procedure tests and process

6.3.1. Prior to randomization

- Detailed history and physical examination, checking for inclusion and exclusion criteria, baseline GCSI score, PAGI-SYM and score. There is no washout period for prokinetics or antiemetics prior to G-POEM (allowed drugs). All prokinetics should be discontinued at least 48 – 72 h before gastric emptying study. Patients with a previous attempt(s) of pyloric balloon dilatation, temporary stenting or botulinum toxin injection are eligible for inclusion, but there must be a wash-out period of at least 6 months prior to randomization.
- Upper GI endoscopy with gastric and duodenal biopsies (diagnosis of *H. pylori* and exclusion of eosinophilic gastroenteritis) (less than 4 months prior to randomization). *In patients tested positive for H.pylori, its treatment will be discussed with the patient individually. Treatment of H.pylori is not necessary before enrolment unless there is an absolute indication for its treatment.*
- Gastric emptying study (**Appendix 3**): scintigraphy protocol in all patients – (protocol endorsed by ANMS: American Nuclear Medicine Society, 2009); (less than 6 months prior to randomization). *Test will begin with patients under fasting conditions for a minimum of 6 hours. A radiolabelled meal will be prepared by adding 0.75 mCi ^{99m}Tc-sulfur colloid into 2 the liquid egg whites.Eggs will be cooked in a microwave or on a hot nonstick skillet, the egges will be stirred once or twice during cooking until firm – to the consistency of an omelette. Then, the bread will be toasted and jelly spread on the toasted bread. Gamma camera images will be obtained immediately after meal ingestion and then at 1, 2, 3 and 4 hours. The geometric mean of delay-corrected counts will be used to estimate the proportion of ^{99m}Tc emptied at each time point. Diagnostic criterion for gastroparesis is defined as the percentage of gastric retention >60% at 2 h and equal to or greater than 10% at 4 h or both. Half-time (T1/2) emptying time will also be calculated. In case of poor tolerance of solids during gastric scintigraphy, radiolabelled liquids will be used (see inclusion criteria). At least 72 hours before gastric emptying test, narcotics and other medications that can delay gastric emptying should be discontinued*.

Other alternative meals may be used for patients with egg allergies or egg´s intolerance according to the local principles.

- (Optional) Gastric Emptying breath test (GEBT) (**Appendix 3**). *The kinetics of appearance of ^{13}C in breath CO2 reflects the rate of gastric emptying of the solid phase of a meal. A dose of 100 mg Octanoic Acid is administered orally in a solid test meal. The test meal is standardized and consists of one scrambled egg with two slices of white bread and 5g of margarine, together with 150 ml water (swallowed immediately after ingestion of the meal). The total caloric content is 250 kcal. The half emptying time and the lag phase time are calculated as well as the gastric emptying coefficient (GEC) (Tab 3 and 4).*

Table 3: Summary of GEBT

	Dose	Samples	
Adults	100 mg (1-^{13}C)-Octanoic Acid	2	*Before administration*
		6	*Every 5 minutes for the first 30 minutes after administration (0.5.h)*
		14	*Every 15 minutes for the next 210 minutes after administration (3.5 h)*

Normal/abnormal values of gastric empyting breath test will be based on a solid normal range determination for a test, which will be used (for example, one criterion might be a gastric emptying half-time – T1/2 > 109 min)

Table 4: FDA accepted **r**eference range cut-off points in healthy population for GEBT (https://www.accessdata.fda.gov/cdrh_docs/pdf11/P110015c.pdf)

Time point	kPCD (min^{-1})
45 min	12.9
90 min	26.9
120 min	**34.4**
150 min	39.5
180 min	43.0
240 min	**35.0**

([i]kPCD - a metric which expresses a subject's $^{13}CO2$ excretion rate at each measurement time)

6.3.2. Randomization

Patients fulfilling all inclusion criteria (and without exclusion criteria) will be asked to sign an informed consent form. Then, participants will be randomly allocated into the two groups:

A. Group that will receive G-POEM (active arm), or
B. Group that will receive sham procedure ("placebo" arm)

The patients will be stratified by the etiology of gastroparesis and gender. Patients will be randomised in blocks of 6. After signing an informed consent form, each centre will e-mail patient initials, birth date, patient´s gender and etiology of gastroparesis (1= idiopathic or other, 2= diabetic; 3 = post-surgical) to a study nurse in Prague and she will send back as soon as possible the treatment allocation with patient´s number. Patients will be randomized in 1:1 allocation ratio and patients must not be informed about their assignment.

6.3.3. Post-randomization tests (before the procedure)

- PAGI - QoL (**appendix 4**)
- Body weight/BMI, ASA physical status
- Current lab values (Haemoglobin, WBC, RBC, Thrombocytes, CRP, Quick, Haemoglobin A1c in patients with DM)
- Upper GI endoscopy 1 day prior to the procedure (optional – for removing food residues)
- EndoFLIP measurement of pyloric distensibility, cross-sectional area and diameter – under general anesthesia before G-POEM/sham

Patients will be admitted to the hospital one day prior to the procedure or the day of the procedure. The day before the procedure, patients will be allowed to drink until 20.00. In the morning (day of the procedure), proton pump inhibitor (PPI) (e.g. omeprazole 40 mg) will be administered intravenously to patients allocated to the active (G-POEM) group; patients in the sham arm will be given placebo (normal saline) (**Appendix 5**).

7. EndoFLIP

EndoFLIP (= Functional Lumen Imaging Probe) is a new method using a principle of impedance planimetry allowing to measure distensibility of a hollow organ, ideally sphincter. In the stomach, it could help to understand the contributing pathophysiology of an impaired function of a pyloric sphincter in patients with gastroparesis. Furthermore, it could help to select appropriate patients for pylorus-directed therapies in patients with gastroparesis, as EndoFLIP provides real-time and dynamic information on pyloric distention, cross-sectional area and diameter.

EndoFLIP measurement will be performed three times: before the procedure (during sham procedure or just before G-POEM), immediately after G-POEM and 3-9 months after G-POEM, always in sedated patients (3-9 months) or in patients under general anesthesia (prior to and after G-POEM). After the calibration, the balloon, equipped with pressure and impedance sensors, will be introduced into the esophagus and under endoscopic control will be passed through the pylorus into the correct position. If necessary, special accessories (snare, grasper) will be used. After that, measurement of various parameters will be performed and the following values will be recorded: distensibility, balloon pressure, cross sectional area, diameter under different balloon volumes (30, 40 and 50 mL). EndoFLIP procedure prolongs the endoscopic examination by approximately 10-15 minutes.

G-POEM procedure

The procedure consists of the following steps (Table 5):

Table 5:

1) Mucosal incision at the greater curvature 3-5 cm from the pylorus
2) Submucosal tunnelling
3) Finding pyloric sphincter
4) Myotomy (2-3 cm) of the pyloric muscle
5) Incision closure (endoclips or suture device)

All procedures will be performed by an experienced endoscopist under general anaesthesia with a high-definition endoscope, fitted with a plastic distal attachment. Exclusively CO^2 will be used for insufflation. Submucosal tunnel will be created by choosing an entry point (usually at 5-6 o´clock) in the antrum at the greater curvature approximately 3-5 cm proximal to the pylorus. After a mucosal incision (1-2 cm), a submucosal tunnel towards pyloric muscle will be created. After finding pyloric arc, the muscle will be myotomised (at 6 o´clock position, complete myotomy to the serosa, length 2-3 cm). For the whole procedure, TT knife or IT knife (Olympus) will be used. For mucosal incision, endocut mode will be used; for tunnelling and myotomy, spray or swift coagulation will be used. Coag-grasper will be used for haemostasis. At the end of the procedure, the mucosal incision will be closed by using endoclips, alternatively, suturing device (Apollo® OverStitch), OTSC clip or KING closure (endoloop + clips) may be used at the direction of an investigator. Inadvertent mucosal injuries will be closed by endoclips if necessary. All procedure-related instruments are listed in Table 6.

Table 6: G-POEM baseline instruments

Erbe Vio 300D with presets: Endocut Q, Spray Coagulation 40-60W, Effect 2 (incision, dissection and myotomy) or Swift Coagulation. Endocut I, Spray or Forced Coagulation 40W, Effect 2 (bleeding control with coag grasper)
CO_2 Unit, low flow CO_2
Waterjet pump with sterile fluid for flushing
High definition endoscope
Single use endotherapeutic instruments: - IT and/or TT knife (with jet function if available) - Coagrasper for haemostasis - Injector - Clips or endoscopic suturing device or endo-loop device or OVESCO clip - e.g. MH-588 distal attachment (Olympus or Fuji)

Experienced endoscopist to be eligible to perform G-POEM in this study:

- At least 4 G-POEMs and one of following (Table 7)

Table 7:

A) >35 POEMs or
B) >30 ESD procedures

8. SHAM procedure

Patients randomized into the sham group will undergo general anaesthesia (or deep sedation with propofol) and a standard upper GI endoscopy with a high definition endoscope will be performed. EndoFLIP measurement will be performed during the sham procedure. No G-POEM will be effectuated and patient will be awakened after 30-60 minutes. All other post-procedure tests will be done in the same fashion as in patients with G-POEM arm.

All "trialists" including the "paramedics" staff will be asked not to inform the patient about the treatment allocation.

9. Perioperative and post-operative management and follow-up

a. Perioperative management

- Sixty to fifteen minutes before the procedure, the patients (active group) will be administered antibiotics i.v.: Ceftriaxone 2 gr (or similar antibiotics) plus Metronidazole 500 mg. Patients in the sham group will be given placebo (normal saline).
- If necessary, pneumoperitoneum will be decompressed by using a Verres needle or venous cannula.

b. Postoperative management - day of the procedure:

- Recovery from general anaesthesia/sedation
- Analgesic and anti-emetic as needed
- i.v. omeprazole 3x40 mg (or other IPP) only in active arm, placebo (normal saline) in the sham arm
- Nothing per mouth until POD 1 (both arms)
- Thorough monitoring until POD 1

c. Post-operative day 1 (POD 1):

- Morning: last dose of i.v. omeprazole 40 mg i.v. (or similar PPI) – placebo in the sham group. Then esomeprazole 2x40 mg (or other PPI) for at least 3 weeks (both arms).
- Ceftriaxone 2 gr i.v. or similar ATB, placebo in sham arm
- Blood (Haemoglobin, RBC, WBC, CRP, etc.)
- Mucosal integrity will be confirmed with either endoscopy or X-ray with water soluble contrast or both at the discretion of an investigator. *If necessary, additional intervention to close mucosal incision will be used.*
- If no leak is detected, patients will be allowed to drink clear fluids and begin re-alimentation.
- Discharge possible on POD 1 (or POD 2)

d. Follow-up visits:

Follow up visits are scheduled at 3 and at 6 months. At 6 months, all patients will be informed about their treatment allocation and patients in the sham group will be offered G-POEM in case of their persisting symptomatology (no or minor benefit of the sham procedure). The decision must be made in next 3 months and these patients undergo the G-POEM procedure no longer than 6 months after the previous follow-up visit. Further follow-up visits for patients randomized to the active arm are scheduled at 12M, 24M and 36M.

Patients originally randomized into the sham arm, who will undergo G-POEM, will be followed like patients after G-POEM (f-u visits at 3M, 6M, 12M, 24M and 36M) (**Appendix 6**) except of scintigraphy/GEBT at 12M (Table 8B).

Patients originally randomized into the sham arm who had treatment success or didn´t want to undergo G-POEM will be followed at 12M and 24M to be sure there is none treatment recurence.

For both, active and SHAM groups, diet modification, nutrition support, prokinetics, antiemetics are allowed during the follow-up. No interventions such as pyloric balloon dilatations, transpyloric stent placement, botox application or surgery) are allowed during the follow-up.

Unscheduled visits or telephone contacts may occur as needed. No time windows or minimum time separations are imposed for such visits or contacts. Data collection forms are

not required at interim visits. If gastroparesis symptom exacerbation occurs between scheduled visits, complete the Interim Event Report (IE) form. The visit code for the form will be "n" (*not applicable in this situation)*

3M visit (±3 weeks):

- Symptom questionnaires (PAGI-QoL, GSCI, PAGI-SYM)
- Body weight/BMI
- Endoscopy
- Gastric emptying study (scintigraphy) in all pts
- Gastric emptying breath test (optional)
- EndoFLIP measurement (only after G-POEM, may be performed 3-9 months after G-POEM), patients from sham group will undergo the "sham EndoFLIP measurement"

6M visit (± 1 month):

- Symptom questionnaires (PAGI-QoL, GSCI, PAGI-SYM)
- Body weight/BMI

12M visit (± 1 month):

- Symptom questionnaires (PAGI-QoL, GSCI, PAGI-SYM)
- Body weight/BMI
- Endoscopy (optional)
- Gastric emptying study (scintigraphy) in the active arm only (Tab. 8A)
- Gastric emptying breath test (optional)
- No GES (scintigraphy) in the sham group after unblinded allocation to G-POEM (Tab. 8B)

24M and 36M visit (±2 months):

- Symptom questionnaires (PAGI-QoL, GSCI, PAGI-SYM)
- Body weight/BMI
- Gastric emptying study (scintigraphy) at 36M (optional)

Table 8A – study design for patients in the active (G-POEM) group

	Patients in G-POEM group at the beginning of randomisation							
	Baseline	*POD 0 – day of GPOEM*	*POD 1*	*3M*	*6M*	*12M*	*24M*	*36M*
Scinti / GEBT	☻			☻	-	☻		☻ (optional)
Endoscopy	☻	☻	☻ (optional)	☻	-	☻ (optional)		
GCSI + PAGI-SYM + PAGI-QoL	☻			☻	☻	☻	☻	☻
Blood tests	☻		☻					
EndoFLIP		☻ *Before and after GPOEM*		☻				

GEBT (Gastric emptying breath test), QoL = Quality of Life, GCSI = Gastroparesis Cardinal Symptoms Index, PAGI-SYM = Patient Assessment of Upper Gastrointestinal Disorders Symptoms, PAGI-QoL = Patient Assessment of Upper Gastrointestinal Disorders-Quality of Life, POD – Post Operative Day.

Table 8B – study design for patients randomized in the sham group

	Patients in sham group / allocation to G-POEM procedure*											
	Baseline	*POD 0*	*POD 1*	*3M*	***6M ****	*POD 0/ **G-POEM***	***POD1***	*3M*	*6M*	*12M*	*24M*	*36M*
Scinti / GEBT	☻			☻	-			☻				☻ (optional)
Endoscopy	☻	☻	☻ (optional)	☻	-		☻ (optional)	☻		☻ (optional)		
GCSI + PAGI-SYM + QoL, PAGI-QoL	☻			☻	☻			☻	☻	☻	☻	☻
Blood tests	☻		☻				☻					
EndoFLIP		☻ *Before*[⊥]		☻ *sham*		☻ *After GPOEM*		☻				

** At 6M, the patients in sham arm will be offered to undergo G-POEM (if their symptoms persist).*

[⊥] In patients having undergone EndoFLIP during the sham procedure, no EndoFLIP measurement will be repeated prior to G-POEM.

♦ sham measurement

10. Study outcomes

a. Main outcome

Main outcome is the proportion of patients with treatment success at 6 months after the procedure.

Treatment success is defined as a decrease of a total GSCI symptom score at least **50%** (see Table 9 for GSCI values to define treatment success according to a baseline GSCI value).

Table 9: GSCI values to define treatment success according to the baseline GSCI value

Mean GCSI score total – baseline	Maximal GSCI score for treatment success (decrease of GSCI about 50%)
2.31*	1.15
2.4	1.20
2.5	1.25
2.6	1.30
2.7	1.35
2.8	1.40
2.9	1.45
3.0	1.50
3.1	1.55 (see description below the table)
3.2	1.60
3.3	1.65
3.4	1.70
3.5	1.75
3.6	1.80
3.7	1.85
3.8	1.90
3.9	1.95
4.0	2.00
4.1	2.05
4.2	2.10
4.3	2.15
4.4	2.20
4.5	2.25
4.6	2.30
4.7	2.35
4.8	2.40
4.9	2.45
5.0	2.50

NOTE: GSCI above 2.3 is an inclusion criterion

Example: If baseline GSCI value was 3.1, a patient will have treatment success if post-treatment GSCI score will be 1.55 or lower.

10.2. Secondary outcomes:

1. Proportion of patients with treatment success in the active arm at 3M, 12M, 24M and 36M.
2. Proportion of patients with treatment success in the sham group at 3M.
3. Change in GSCI and PAGI-SYM before and after G-POEM at 3, 6, 12, 24 and 36M and before vs. after sham procedure at 3M and 6M; comparison of the change of the scores between the active and sham groups.
4. Proportion of patients (randomized into the sham group and undergoing G-POEM after 6M) with treatment success after the sham procedure (6M) and after G-POEM (6M) – cross over part
5. Change in GSCI and PAGI-SYM before and after G-POEM vs. sham procedure in patients randomized to the sham group at 3 and 6M (cross over part)
6. **Subgroup post-hoc analyses of the treatment success and change in symptomatic scores according to etiology of gastroparesis.**
7. Change in gastric emptying (scintigraphy) study, EndoFLIP values and/or gastric emptying breath test before and after both G-POEM and sham procedure; comparison of the mean change of these parameters between active and sham groups.
8. Procedure details (length of the procedure, technical success, perioperative adverse events).
9. Short- and long-term adverse events**.**

11. Sample size calculation

A total of 86 patients will be randomized.

- 43 patients will be randomized into the G-POEM group
- 43 patients will be randomized into sham group

Sample size calculation is based on expected therapeutic success of G-POEM in 50% of patients vs. 20% in the sham group; significance level 0,05; study power 0,8; beta error 0,2; adjustment for 15% expected drop out.

12. Statistical analysis

Data will be analyzed for both the intention to treat (ITT – a cohort for primary end point analysis) and the per-protocol population (sensitivity analysis). The per-protocol analysis will include only patients who will complete the entire follow-up. The efficacy of G-POEM at month 6 (main outcome) will be evaluated by Poisson regression with robust standard errors or similar regresion based approach. Baseline characteristics will be compared by using chi-squared tests for categorical variables and t-tests for continuous data.

The efficacy of treatment at other time points (3M, 12M, 24M and 36M) as well as the pooled efficacy cross-over + active groups will be evaluated similarly.

Mean changes in the GCSI score and gastric emptying time in the active vs. control group will be analysed by ANCOVA. A p value less than .05 will be considered statistically significant. **An interim-analysis** will be performed when **40%** of patients will have completed the 6 months follow-up.

Statistical analysis plan

To prevent possible bias caused by the choice of statistical methods, this extended statistical analysis plan aims to adhere to the plan presented in the original version of the protocol while adding details of the planned analysis approaches.

The cohort for the primary analysis will be the intention to treat (ITT) population including all randomized patients according to their original allocation regardless of the actual treatment received or follow-up adherence to the protocol exhibited. Missing data in the ITT population will be imputed using the multiple imputation method. Further, the primary outcome will be analyzed on the per-protocol population to asses sensitivity.

The difference in efficacy of G-POEM versus sham at month 6 (main outcome) will be evaluated by logistic regression, same as the differences in treatment success at other time points (3M, 12M, 24M and 36M). Logistic regression will also be used to search for predictors of treatment success among other variables (age, sex, baseline values of scores, baseline GES, and EndoFLIP measurements).

Further, 95% confidence intervals for the point estimates of treatment success rates will be constructed using the Wilson method for G-POEM, sham, cross-over, pooled original + cross-over G-POEM groups and also etiology sub-groups at all time points.

Continuous secondary outcome variables will be presented as means or medians with 95% confidence intervals in dependence on the results of normality test for particular variables. The confidence intervals will be constructed using a bootstrapping method, which can be

conveniently combined with the multiple imputation approach even for estimation of the median, where normality is not assumed. Between group differences for the secondary outcomes will be tested using t-tests, possibly after data transformation for highly non-normal data.

The difference in treatment success between G-POEM and sham groups at 6 months is the only confirmatory hypothesis test of the study and the 5% p-value threshold for statistical significance will be applied. All the other hypotheses and secondary outcomes are considered exploratory and p-values from the corresponding tests will be presented without any multiple testing correction without the aim to keep the overall false positive error rate at 5% across all results.

Baseline characteristics will be compared only using descriptive statistics as any potential statistically significant difference between the two study groups would be due to chance by design in a randomized trial.

An interim-analysis will be performed when 40% of patients will have completed the 6 months follow-up. The interim analysis will be performed on available data basis investigating the primary hypothesis with the same tools as planned for the final analysis. Since no rules for interim stopping of the trial based on adjusted thresholds for p-values was specified in the study design, the Haybittle–Peto boundary will be used for potential stopping of the trial for early confirmation of the treatment effect. On the other hand, enrollment would be stopped for futility if the interim conditional power of the study assuming the observed effect sizes for the remaining patients was below 20%. Also, analysis of adverse events could result in study termination.

13. Study monitoring, data and safety monitoring board

The study will be monitored by an independent certified agency (to check for reliability of data and ethical standards). A standard Data and Safety Monitoring Board (DSMB) will be created in the coordinating centre in Prague and its membership will include: 4 gastroenterologists, 1 independent specialist (MD), 1 statistician, 1 lawyer and 1 independent person (not MD). The main aim of DSMB will be: to review the research protocols, informed consent documents, and plans for data and safety monitoring, including all proposed revisions; to evaluate the progress of studies, including periodic assessments of data quality and timeliness, participant recruitment; to protect the safety of the study participants; to report on the safety of the study participants and progress of the trial; to consider factors external to the study when relevant information becomes available, such as scientific or therapeutic

developments that may have an impact on the safety of the participants or the ethics of the study; to make recommendations.

14. Study termination

Individual case:

- *Severe peri-procedural complications as bleedings or perforations requiring a surgical intervention.*
- *Technically unsuccessful/unfinished G-POEM.*
- *The patient's request to terminate the participation in the study.*
- *Severe symptomatology requiring intervention not allowing to finish 6 months of the follow up, especially in patients in the sham group (pyloric botulinum toxin injection, balloon dilatation, laparoscopic surgery, transpyloric stenting).*

15. Ethical and administrative aspects

15.1. Ethical considerations

The study will be conducted according to the Declaration of Helsinki, adopted by the 18th WMA General Assembly, Helsinki, Finland, June 1964, and amended by the 59th WMA General Assembly, Seoul, October 2008. Approval of the medical ethical committee or Institutional Review Board of all participating centers will be obtained. The DSMB in Prague (see above) will function as independent safety monitoring board and will receive written study reports of study outcomes and follow-up. All patients are required to sign a written informed consent form prior to randomization. Study patients can leave the study at any time for any reason without consequences. The study protocol, patient information and informed consent form and all other necessary documents (study amendments) will be submitted together with an appropriate request form to Ethical Committees (or IRB) for approval at each participating centre.

15.2. Patient Information and Informed Consent

Each patient will be informed adequately about content, consequences and risks of the study. This will happen by a standardized consent form, as well as verbally by a study investigator at each centre. It is the responsibility of the principal investigator or his assigned co-investigators to obtain signed informed consent according to the international standards

(including data management and security) from each patient before inclusion into the study. Patients will be given the opportunity to ask any questions that might arise. The informed consent form will be filled single-handed by the patient. The original remains at the study site, the patient will be handed a copy.

15.3. Data security

Only anonymous data will appear in any publication. Regulations of data security of the countries of the participating centres will be adhered to. Patients will be informed that their data will be pseudonymized according to documentation obligations and notification duties by §§ 12 and 13 GCP enactment. Patients who will not agree with these regulations will not be enrolled. Patients will be provided with a code during the randomization process. For data management, only pseudonymized data will be provided to the main study centre (Prague).

15.4. Administrative Aspects and Adverse Events recording

Patients will be coded using a numeric randomization code (anonymized) and only anonymized data will be submitted to the PI site (Prague) for data management. Adverse events (AE) and serious adverse events (SAE) (**Appendix 8**) will be reported to the coordinating study centre and there to the DSMB and the study steering committee (see below) to control safety issues and to discuss intervention or protocol amendments accordingly. An adverse event is any undesirable event that occurs whether or not considered to the participation in the study. Therefore, AEs can be any undesirable, unintentional, or unanticipated outcome or symptom or any disease in a timely relation to the study participation, independent of a suspected relation.

AE will be documented on designated CRF forms. A serious adverse event is defined as any event within the study timeframe fulfilling at least one of the following criteria:

- Death
- Live-threatening event
- Hospitalisation (required or prolonged)
- Event that results in disability
- Any event that an intervention to prevent one of the points above....

According to ICH-GCP guidelines, SAE(s) have to be reported within 24hours by email (gaps@ikem.cz, copy: jan.martinek@volny.cz). SAE have to be reported to the local ethics committee of the study site.

CRFs as well as all study related documents will be kept at least 10 years after study termination. Any study subject is allowed to withdraw her / his consent for study participation at any time for any given reason.

16. References

1. Revicki, D.A., et al., *Gastroparesis Cardinal Symptom Index (GCSI): development and validation of a patient reported assessment of severity of gastroparesis symptoms.* Qual Life Res, 2004. 13(4): p. 833-44.
2. Jung, H.K., et al., *The incidence, prevalence, and outcomes of patients with gastroparesis in Olmsted County, Minnesota, from 1996 to 2006.* Gastroenterology, 2009. 136(4): p. 1225-33.
3. Dickman, R., et al., *Gender aspects suggestive of gastroparesis in patients with diabetes mellitus: a cross-sectional survey.* BMC Gastroenterol, 2014. 14: p. 34.
4. Langworthy, J., H.P. Parkman, and R. Schey, *Emerging strategies for the treatment of gastroparesis.* Expert Rev Gastroenterol Hepatol, 2016. 10(7): p. 817-25.
5. Davis, B.R., et al., *The Long-Term Efficacy and Safety of Pyloroplasty Combined with Gastric Electrical Stimulation Therapy in Gastroparesis.* J Gastrointest Surg, 2016.
6. Heetun, Z.S. and E.M. Quigley, *Gastroparesis and Parkinson's disease: a systematic review.* Parkinsonism Relat Disord, 2012. 18(5): p. 433-40.
7. Hasler, W.L., *Gastroparesis--current concepts and considerations.* Medscape J Med, 2008. 10(1): p. 16.
8. Camilleri, M., et al., *Clinical guideline: management of gastroparesis.* Am J Gastroenterol, 2013. 108(1): p. 18-37; quiz 38.
9. Szarka, L.A., et al., *A stable isotope breath test with a standard meal for abnormal gastric emptying of solids in the clinic and in research.* Clin Gastroenterol Hepatol, 2008. **6**(6): p. 635-643 e1.
10. Quigley, E.M., *Prokinetics in the Management of Functional Gastrointestinal Disorders.* J Neurogastroenterol Motil, 2015. 21(3): p. 330-6.
11. Stacher, G., et al., *Cisapride versus placebo for 8 weeks on glycemic control and gastric emptying in insulin-dependent diabetes: a double blind cross-over trial.* J Clin Endocrinol Metab, 1999. 84(7): p. 2357-62.
12. Arts, J., et al., *Clinical trial: a randomized-controlled crossover study of intrapyloric injection of botulinum toxin in gastroparesis.* Aliment Pharmacol Ther, 2007. 26(9): p. 1251-8.
13. Lacy, B.E., et al., *The treatment of diabetic gastroparesis with botulinum toxin injection of the pylorus.* Diabetes Care, 2004. 27(10): p. 2341-7.
14. Ukleja, A., et al., *Endoscopic botox injections in therapy of refractory gastroparesis.* World J Gastrointest Endosc, 2015. **7**(8): p. 790-8.
15. Khashab, M.A., et al., *Refractory gastroparesis can be successfully managed with endoscopic transpyloric stent placement and fixation (with video).* Gastrointest Endosc, 2015. 82(6): p. 1106-9.
16. Clarke, J.O. and W.J. Snape, Jr., *Pyloric sphincter therapy: botulinum toxin, stents, and pyloromyotomy.* Gastroenterol Clin North Am, 2015. 44(1): p. 127-36.
17. Hibbard, M.L., C.M. Dunst, and L.L. Swanstrom, *Laparoscopic and endoscopic pyloroplasty for gastroparesis results in sustained symptom improvement.* J Gastrointest Surg, 2011. 15(9): p. 1513-9.
18. Borrazzo, E.C., *Surgical management of gastroparesis: gastrostomy/jejunostomy tubes, gastrectomy, pyloroplasty, gastric electrical stimulation.* J Gastrointest Surg, 2013. 17(9): p. 1559-61.
19. Chu, H., et al., *Treatment of high-frequency gastric electrical stimulation for gastroparesis.* J Gastroenterol Hepatol, 2012. 27(6): p. 1017-26.

20. Abell, T., et al., *Gastric electrical stimulation for medically refractory gastroparesis.* Gastroenterology, 2003. 125(2): p. 421-8.
21. Nguyen, L.A. and W.J. Snape, Jr., *Clinical presentation and pathophysiology of gastroparesis.* Gastroenterol Clin North Am, 2015. 44(1): p. 21-30.
22. Pasricha, P.J., et al., *Submucosal endoscopic esophageal myotomy: a novel experimental approach for the treatment of achalasia.* Endoscopy, 2007. 39(9): p. 761-4.
23. Ward, M.A., et al., *Outcomes in patients with over 1-year follow-up after peroral endoscopic myotomy (POEM).* Surg Endosc, 2016.
24. Shiwaku, H., et al., *Peroral endoscopic myotomy for esophageal achalasia: outcomes of the first over 100 patients with short-term follow-up.* Surg Endosc, 2016. 30(11): p. 4817-4826.
25. Kristensen, H.O., et al., *Long-term outcome of peroral endoscopic myotomy for esophageal achalasia in patients with previous Heller myotomy.* Surg Endosc, 2016.
26. Khashab, M.A., et al., *Gastric per-oral endoscopic myotomy for refractory gastroparesis: results from the first multicenter study on endoscopic pyloromyotomy (with video).* Gastrointest Endosc, 2016.
27. Bromer, M.Q., et al., *Simultaneous measurement of gastric emptying with a simple muffin meal using [13C]octanoate breath test and scintigraphy in normal subjects and patients with dyspeptic symptoms.* Dig Dis Sci, 2002. 47(7): p. 1657-63.
28. Gonzalez J.M. et al., *Gastric per-oral endoscopic myotomy with antropyloromyotomy in the treatment of refractory gastroparesis: clinical experience with follow-up and scintigraphic evaluation (with video),* GIE, 2017. 85(1): p. 132-139
29. Rentz, A.M., et al., *Development and psychometric evaluation of the patient assessment of upper gastrointestinal symptom severity index (PAGI-SYM) in patients with upper gastrointestinal disorders.* Qual Life Res, 2004. 13(10): p. 1737-49.

Appendix 1:

Steering Committees (SC)
All lead investigators will be steering committee members. One lead investigator per country will be nominated as national coordinator

Data and Safety Monitoring Board (DSMB)
Organisation of steering committee meetings
Reporting SAEs (*Serious adverse events)*
Assistance with international review, board/independent ethics committee applications
Data verification
Randomisation, unblinded analysis of result

Data Manager
Maintenance of trial IT system:

a) Mgr. Jan Mareš (IKEM)
b) M.D. Martin Janicko Ph.D, 1 Department of Internal medicine, Pavol Jozef Safarik University in Kosice, Louis Pasteur University hospital, 04001 Kosice, Slovak Republic

Data entry and verification:

a) M.D. Rastislav Hustak, Gastroenterology Department, University Hospital of Trnava and FZsPTU, SlovakRepublic
b) M.D. Martin Janicko Ph.D

Appendix 2.
Questionnaires

GSCI score

		none	Very mild	Mild	Moderate	Severe	Very severe
1.	Nausea	0	1	2	3	4	5
2.	Retching	0	1	2	3	4	5
3.	Vomiting	0	1	2	3	4	5
4.	Stomach fullness	0	1	2	3	4	5
5.	Not able to finish a normal sized meal	0	1	2	3	4	5
6.	Feeling extensively full after meals	0	1	2	3	4	5
7.	Loss of appetite	0	1	2	3	4	5
8.	Bloating	0	1	2	3	4	5
9.	Stomach or belly visibly larger	0	1	2	3	4	5

1-3 = nausea/vomiting
4-7 = post-prandial fullness/early satiety
8-9 = bloating

Calculation:
Total GSCI score = arithmetic mean of the three symptom subscales
Subscores = arithmetic means of (1-3), (4-7) and (8-9)

PAGI-SYM score

	None	Very mild	Mild	Moderate	Severe	Very severe
1. Heartburn (burning pain rising in your chest or throat) during the day	0	1	2	3	4	5
2. Regurgitation or reflux (fluid or liquid from your stomach coming up into your throat) during the day	0	1	2	3	4	5
3. Heartburn (burning pain rising in your chest or throat) when lying down	0	1	2	3	4	5
4. Regurgitation or reflux (fluid or liquid from your stomach coming up into your throat) when lying down	0	1	2	3	4	5
5. Feeling of discomfort inside your chest during the day	0	1	2	3	4	5
6. Bitter, acid or sour taste in your mouth	0	1	2	3	4	5
7. Feeling of discomfort inside your chest at night (during sleep time)	0	1	2	3	4	5
8. Vomiting	0	1	2	3	4	5
9. Nausea (feeling sick to your stomach as if you were going to vomit or throw up)	0	1	2	3	4	5
10. Retching (heaving as if to vomit, but nothing comes up)	0	1	2	3	4	5
11. Stomach fullness	0	1	2	3	4	5
12. Not able to finish a normal-sized meal	0	1	2	3	4	5
13. Feeling excessively full after meals	0	1	2	3	4	5
14. Loss of appetite	0	1	2	3	4	5
15. Bloating (feeling like you need to loosen your clothes)	0	1	2	3	4	5
16. Stomach or belly visibly larger	0	1	2	3	4	5
17. Upper abdominal (above the navel) discomfort	0	1	2	3	4	5
18. Upper abdominal (above the navel) pain	0	1	2	3	4	5
19. Lower abdominal (below the navel) pain	0	1	2	3	4	5
20. Lower abdominal (below navel) discomfort	0	1	2	3	4	5

1 - 7 = heartburn/regurgitation
8 - 10 = nausea/vomiting
11 - 14 = post-prandial fullness/early satiety
15 - 16 = bloating
17 - 18 = upper abdominal pain
19 - 20 = lower abdominal pain

Calculation:

Subscale scores are calculated by averaging across items comprising the subscale; scores vary from 0 (none or absent) to 5 (very severe). The half-scale rule is applied for missing data (i.e., the subscale score is calculated by using the mean of non-missing items; when more than 50% of items are missing, the score is set to missing).
A total score is calculated by averaging all subscale scores.

Appendix 3:

Egg Beaters Gastric Emptying Scintigraphy

Items needed for Egg Beaters Gastric Emptying Scintigraphy:
118 mL of liquid egg whites (Egg Beaters; egg substitute): 99% real eggs, cholesterol free, fat free, low calorie (120 g Egg Beater, 60 kcal, ***approx. two large eggs****), 2 slices of wheat bread (120 kcal), Strawberry jam (30 g, 74 kcal) Water (120 ml). Technetium-99m 0.75 mCi*

To prepare the meal, 0.75 Ci of 99Tc sulfur – colloid is mixed with liquid egg whites, the mixture is cooked in a microwave or on a hot nonstick skillet. The Egg Beater mixture is stirred once or twice during cooking and is cooked until it has the consistency of an omelet (3-5 min). The bread is toasted. Jelly is spread on the bread, and a sandwich is made of the jellied bread and cooked egg mixture. The subject completes the sandwich meal quickly, within max. 10 minutes.

Gastric emptying studies are generally performed in the morning. Patient should be fasting overnight or for at least 6 hours. Patients should generally stop prokinetic agents, and anticholinergic agents that can affect gastric emptying for 3 days prior to the test. should have a reasonable glucose level for the test. Generally, the fasting glucose in diabetic patients should be between 75 and 275 mg/dL (4.2 to 15.3 mmol/l). Diabetic patients should self-administer their insulin with meal ingestion, generally ½ what they take normally. The nutritional composition of the meal is 69-72% carbohydrate, 22-24% protein, 2% fat and 2% fiber. Other alternative meals may also be useful for patients with egg allergies or intolerance to eggs, and patients with gluten-sensitive enteropathy according local principles.

Gastric Emptying of Solids 13C-Octanoic Acid Breath Test (Leuven Model)

The test is performed after an overnight fast.

A dose of 100 mg (1-^{13}C)-Octanoic Acid is administered orally in a solid test meal. The test meal is standardized and consists of one scrambled egg with two slices of white bread and 5 g of margarine, together with 150 ml water (swallowed immediately after ingestion of the meal). The total caloric content is 250 kcal. The egg yolk is doped with 100 mg (1-^{13}C)-Octanoic Acid and fried separately from the egg white. The meal is consumed within 10 minutes.

Breath samples are collected before (2x), every 5 minutes during the first 30 minutes (0.5 h) and every 15 minutes for the next 210 minutes (3.5 h) after the ingestion of the (1-^{13}C)-Octanoic Acid. ^{13}C enrichment in breath CO_2 is determined by Isotope Ratio Mass Spectrometry (IRMS). The equation of the breath test results is obtained by 2 non-linear regression curves fitting the % dose ^{13}C recovered in breath per minute and the cumulative % dose recovered in breath. From this equation the half emptying time and the lag phase time are calculated as well as the gastric emptying coefficient (GEC).

Appendix 4

The PAGI-QOL (Quality of Life Questionnaire)

The following questions ask about how some of the gastrointestinal problems you may be experiencing (such as pain, discomfort or other problems) may have affected your overall quality of life and well-being in the past 2 weeks.
Please answer every question by *circling the number* that best represents your opinion. There are no right or wrong answers.

During the past 2 weeks, because of your gastrointestinal problems, how often…	*None of the time*	***A little of the time***	***Some of the time***	***A good bit of the time***	***Most of the time***	***All of the time***
1. have you had to depend on others to do your daily activities?	0	1	2	3	4	5
2. have you avoided performing your daily activities?	0	1	2	3	4	5
3. have you had difficulty concentrating?	0	1	2	3	4	5
4. has it taken you longer than usual to perform your daily activities?	0	1	2	3	4	5
5. have you felt tired?	0	1	2	3	4	5
6. have you lost the desire to participate in social activities such as visiting friends or relatives?	0	1	2	3	4	5
7. have you been worried about having stomach symptoms in public?	0	1	2	3	4	5
8. have you avoided performing physical activities or sports?	0	1	2	3	4	5
9. have you avoided traveling?	0	1	2	3	4	5
10. have you felt frustrated about not being able to do what you wanted to do?	0	1	2	3	4	5
11. have you felt constricted in the clothes you wear?	0	1	2	3	4	5
12. have you felt frustrated about not being able to dress as you wanted to?	0	1	2	3	4	5
13. have you felt concerned about what you can and cannot eat?	0	1	2	3	4	5
14. have you avoided certain types of foods?	0	1	2	3	4	5
15. have you restricted eating at restaurant or at someone's home?	0	1	2	3	4	5
16. have you felt less enjoyment in food than usual?	0	1	2	3	4	5
17. have you felt concerned that a change in your food habits could trigger your symptoms?	0	1	2	3	4	5
18. have you felt frustrated about not being able to choose the food you wanted to?	0	1	2	3	4	5
19. have you left frustrated about not being able to choose the type of beverage you wanted to?	0	1	2	3	4	5
20. has your relationship with your spouse or partner been disrupted?	0	1	2	3	4	5
21. has your relationship with your children or relatives been disrupted?	0	1	2	3	4	5
22. has your relationship with your friends been disrupted?	0	1	2	3	4	5
23. have you been in a bad mood?	0	1	2	3	4	5
24. have you felt depressed?	0	1	2	3	4	5
25. have you felt anxious?	0	1	2	3	4	5
26. have you felt angry?	0	1	2	3	4	5
27. have you felt irritable?	0	1	2	3	4	5
28. have you felt discouraged?	0	1	2	3	4	5
29. have you been stressed?	0	1	2	3	4	5
30. have you felt helpless?	0	1	2	3	4	5

The PAGI-QOL contains 30 items with five subscales:

(1) daily activities (1 – 10)
(2) clothing (11 – 12)
(3) diet/food habits (13 – 19)
(4) relationship (20 – 22)
(5) psychological well-being and distress (23 – 30)

The PAGI-QoL questionare contains of 30 items with five subscales: (1) daily activities; (2) clothing; (3) diet/food habits; (4) rela- tionship; and (5) psychological well-being and distress. Each items are scored on a 6-point Likert scale, with response options ranging from 0 (none) to 5 (severe problem all of the time). Subscale scores are calculated by averaging the item responses. A total score is calculated by averaging subscale scores.

Appendix 5
Perioperative protocol

The day before G-POEM (POD-1)

- Fasting for 24 to 48 hours (according the investigator)
- Current weight, QoL, ASA physical status
- Endoscopic examination with event. removal of the remaining food (optional)
- Anesteshiologist´s examination
- Blood sampling for: Haemoglobin, WBC, RBC, Thrombocytes, CRP, Quick, Haemoglobin A1c in patients with DM. Other blood tests possible. - – may also be done in the morning prior to the procedure ("day 0").
- Liquids up to 20:00 hours, then nil per os
- Supportive infusions according to an attending physician

Day „ 0 " (procedure)

- Omeprazole 40 mg iv (or similar PPI) will be given to patients at 6.00 – 7.00 o´clock and then 3xdaily until the morning on POD 1 (saline only will be given in the sham group). Please, keep patients and hospital staff blinded as much possible.
- 15 to 60 minutes before the procedure, the patients (active group) will be administered intravenous antibiotics: e.g. Ceftriaxone 2 gr (or similar antibiotics) and Metronidazole 500mg. Patients in the sham group will be given placebo (normal saline).
- After the procedure, nil per os for 24 hours (both group).
- Analgesics and anti-emetics as needed.

Day after G-POEM (POD1)

- Morning (6:00 am): last dose of i.v. omeprazole 40 mg in active arm (saline in placebo group), then esomeprazole 2x40 mg (or other PPI) for at least 3 weeks and then on demand for all patients.
- Ceftriaxone 2 gr i.v. (or similar antibiotic), placebo (normal saline) in the sham arm.
- Mucosal integrity will be confirmed with either endoscopy or X-ray with water soluble contrast or both at the discretion of an investigator. *If necessary, additional intervention to close mucosal incision will be used.*
- If no leak or other problems are detected, patients will be allowed to drink clear fluids and begin re-alimentation.
- Analgesics if necessary
- Blood sampling: blood count, coagulation, creatinine, ions, CRP, glycemia
- Discharge from a hospital according to the clinical condition

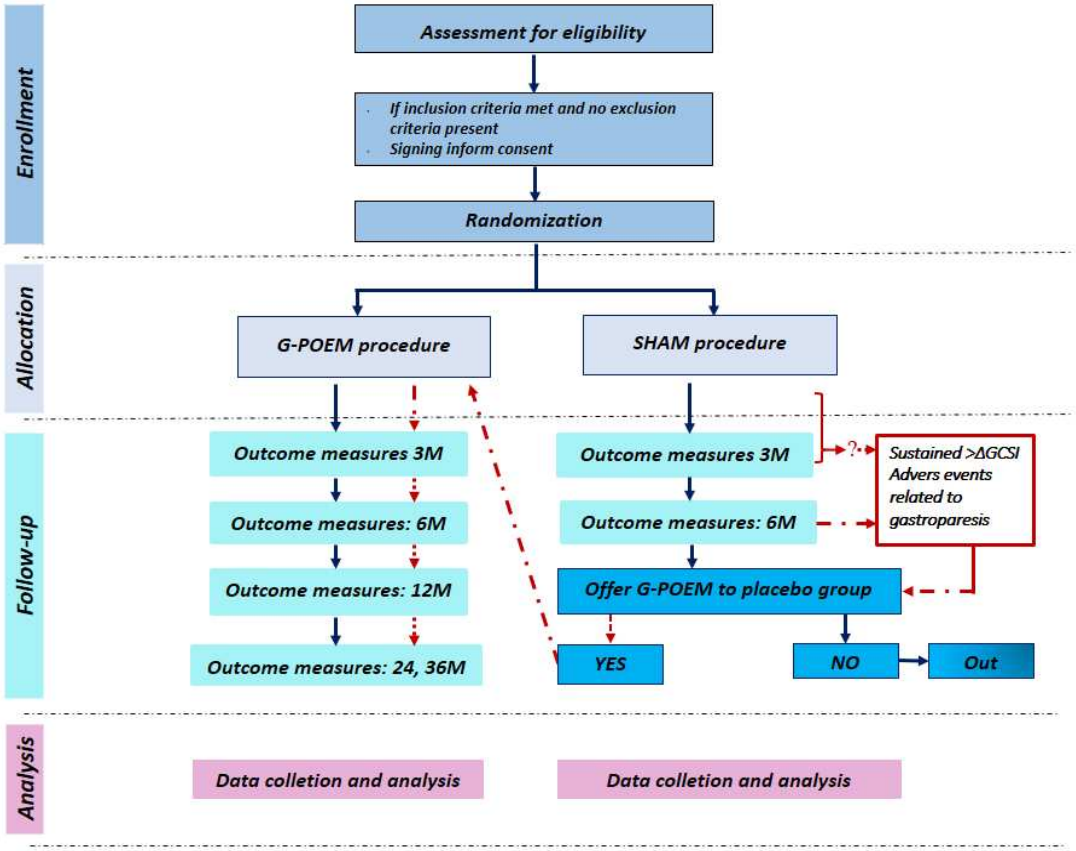
Second day after the procedure (POD2)

- PPI orally – all patients
- ATB only, if necessary

Recommendation: After patient´s discharge, it is advisable to keep patient´s allocation at the Clinic in a closed envelope. I case of need (e.g. urgent visit), the envelope can be opened and a physician on duty has an immediate access to the treatment´s allocation.

Appendix 6

Study design



Appendix 7

Center ID and Patient IDs

Center	Center ID	Patients code
1. Department of Hepatogastroenterology, IKEM, Prague, Czech Republic	***01***	*1 - 50*
2. University Medical Center Hamburg- Eppendorf, Germany	***02***	*1 – 50*
3. Translational Research in GastroIntestinal Disorders, Leuven, Belgium	***03***	*1 – 50*
4. King's Institute of Therapeutic Endoscopy, London, UK	***04***	*1 - 50*
5. III. Medizinische Klinik, Medical Center/Klinikum Augsburg, Germany	***05***	*1 – 50*
6. Department of Hepatogastroenterology at Cliniques universitaires St-Luc, Brussels, Belgium	***06***	*1 – 50*
7. Department of Surgical Gastroenterology, Karolinska University Hospital, Stockholm, Sweden	***07***	*1 – 50*
8. Department of Gastroenterology and Hepatology, Amsterdam, The Netherlands	***08***	*1 – 50*
9. Department of Internal Medicine, University Hospital Trnava, Slovak Republic	***09***	*1 – 50*
10. The Department of Surgical Gastroenterology L, Denmark	***10***	*1 – 50*
11. Jeesenius Faculty of Medicine in Martin, Clinic of Gastroenterological Internal Medicine, Slovak Republic	***11***	*1 - 50*
12. Center for Endoscopic and Therapeutics Research, The University of Chicago, USA	***12***	*1 - 50*
13. Regional Institute of Gastroenterlogy, Cluj-Napoca, Romania	***13***	*1 - 50*
14. Department 2nd Dept. of Internal Medicine – Gastroenterology, University Hospital in Hradec Kralove, Czech republic	*14*	*1 - 50*
15.		
16.		
17.		
18.		
19.		

1. **Codification (for CRF)**

Name : _____
Birth date: _____

Patient ID ☐☐ – ☐☐
Center ID ***Patient code***

Examples:

Name : XXX
Birth date: DD.MM.YY

0 1 – 0 1
Center ID ***Patient code***
(IKEM) *(Patient No 1)*

Name : XXX
Birth date: DD.MM.YY

0 2 – 1 1
Center ID ***Patient code***
(Hamburg) *(Patient No 11)*

Appendix 8:

Adverse event (AE) / Serious Adverse Event (SAE) Report Form

Definition: An adverse event is any undesirable, unintentional or unanticipated event that occurs during use of the investigational device, whether or not considered related to the therapy. A serious adverse event (SAE) is an event that is: fatal, life-threatening, results in persistent or significant disability/incapacity, requires or prolongs inpatient hospitalization. SAE must be reported within 24 hours to the Prague study center (gapc@ikem.cz, copy: jan.martinek@volny.cz) and the Ethics Committees/IRB if applicable.

Hospital visits due to follow up visits are not considered to be SAE.

□ Initial report

□ Consecutive report

Date AE start: _____ / _____ / _____ (DDMMYY)
□ expected event □ unexpected event

Event related to G-POEM / SHAM procedure
□ No □ Possibly □ Yes

Complication: □ Perforation □ Bleeding □ Infection □ Other

Please describe complication:
_____
_____

Intervention required: □ No □ Yes
please describe intervention:
_____
_____

medication required: □ No □ Yes

medication(s):
_____
_____
_____
_____

Report of a Serious Adverse Event

Hospitalization or prolongation of hospital stay required (SAE):
□ Yes □ No

If yes, please report within 24 hours to the Prague study center and Ethics Committee/IRB if applicable!

Date of hospitalisation/ - prolongation_____ (DDMMYY)

Date hospital discharge_____ (DDMMYY)

- -
- -

□ event resolved □ event ongoing

□ long term sequela □death □ unknown

Description/ comment: _____

Date AE stop: _____ / _____ / _____ (DDMMYY)

Data and Safety Monitoring Board

IKEM, Department of hepatogastroenterology

8th of February 2021, Prague

Study: A Randomized, Sham and Cross-Over-Controlled Trial of per-oral endoscopic pyloromyotomy (G-POEM) in patients with refractory gastroparesis (GREG)

ClinicalTrials.gov identifier: NCT03356067

Grant number: 17-28797A

To the principal investigator

Prof. Jan Martínek

Department of hepatogastroenterology

IKEM

Vídeňská 1958/9, Prague

The DSMB met at 29th of January 2021 and reviewed interim results after 41 subjects have been randomized and 33 patients completed the M6 visit. At present, there is a significant difference regarding the main endpoint (treatment success at 6 months) in favor of the active group (vs. sham). Also, 100 % of patients, who had undergone cross-over G-POEM after they underwent a sham procedure, experienced a treatment success. Based on these results, the Committee recommends to stop enrolment now because:

1. It would be highly unlikely that during the further course of the study, the results would change qualitatively. The p-value of the logistic regression analysis comparing the treatment success between G-POEM and sham groups is equal to p=0.003. Since no interim stopping scheme with pre-specified corrected alpha thresholds was specified, we primarily consider the Haybittle–Peto boundary. This boundary states the critical value of p=0.001 at all interim analyses and sometimes p=0.002 is used if only one interim analysis is performed. Our obtained p-value does not directly reach this level of significance, but taking into account the heuristic nature of the Haybittle–Peto approach and the fact that this approach is considered very conservative for stopping a trial, we believe that the evidence in favor of G-POEM efficacy is sufficient for stopping further enrollment.
2. Even if the strict statistical threshold for early trial termination is not met, ethical reasons should be considered. In particular, the risks related to general anesthesia in the sham procedure can not be further justified in the light of the current results. The Board consider it unethical to continue with study enrolment.

There should be further studies assessing the key question of how to select patients who will benefit from G-POEM, but we believe that there is no need for further inclusion of sham procedures.

Data and Safety Monitoring Committee members:

Chairperson:

MUDr. Alice Štrosová – gastroenterologist

Head physician

MEDIENDO s.r.o.

Thámova 13, 186 00 Praha 8-Karlín

MUDr. Jana Krajčíová – gastroenterologist

MUDr. Klára Chmelová – hepatologist

MUDr. Denisa Erhartová – gastroenterologist and hepatologist

MUDr. Martin Jerie – neurologist, independent specialist

Mgr. Barbora Bučková – statistician

MUC. Dagmar Dražilová – medical student

JUDr. Klára Daňková – lawyer

In Prague, 8th of February 2021

MUDr. Alice Štrosová

Head physician

MEDIENDO s.r.o.

Thámova 13, 186 00 Praha 8-Karlín