Published in final edited form as:

Am Surg. 2015 May ; 81(5): 467–471.

Gastric Electrical Stimulation for Refractory Gastroparesis: Predictors of Response and Redefining a Successful Outcome

Bryan Richmond, M.D., M.B.A., Benny Chong, M.D., Asmita Modak, M.S., Mary Emmett, Ph.D., Kimball Knackstedt, D.O., Benjamin Dyer, M.D., and Zachary Aburahma, D.O.

Department of Surgery, West Virginia University/Charleston Division, Charleston, West Virginia

Abstract

Predictors of a favorable response and measures of success with gastric electrical stimulation (GES) for gastroparesis remain elusive. Published results remain inconsistent with respect to patient perceived benefit, despite statistical improvements in objective measures of symptom severity. We performed a retrospective analysis of 56 patients with gastroparesis who underwent insertion of a gastric electrical stimulator during the study period. Data included demographics, symptoms, total symptom severity score (TSS, range 0–24, initial and most recent), and gastric emptying times. TSS were grouped into four severity categories (0–10, 11–14, 15–18, 19–24). TSS improvement was defined as movement to a lower severity category. Perception of improvement was compared with that of TSS score improvement using χ^2 test. Etiology as a predictor of improvement was measured using logistic regression. Initial mean TSS was 21, and post-treatment TSS was 13.5. Improvement was significant for individual symptoms and in reduction of TSS for both diabetic/idiopathic etiologies ($P = 0.001$). No correlation was noted between likelihood of success/failure and gastric emptying times ($P = 0.32$). Thirty-eight improved (moved to lower TSS category), whereas 18 failed (remained in same category) ($P = 0.001$), which correlated with perception of improvement. Of 18 failures, 14 (77.7%) were idiopathic. On logistic regression, diabetics were more likely than idiopathic patients to move to a lower TSS category (odds ratio 14, $P = 0.003$) and even more likely to improve based on patient perception (odds ratio 45, $P = 0.005$). GES produces far more consistent improvement in diabetics. Further study of GES in idiopathic gastroparesis is needed. Application of the proposed TSS severity categories allowed differentiation of small, statistically significant (but clinically insignificant) reductions in TSS from larger, clinically significant reductions, thereby permitting more reliable application of TSS to the evaluation of GES efficacy.

Gastroparesis is a syndrome of delayed gastric emptying in the absence of mechanical obstruction.¹ Symptoms include early satiety, postprandial fullness, nausea, vomiting, bloating, and abdominal pain. Gastroparesis significantly affects patients' quality of life, increases health care costs through hospitalizations and treatment costs, and is associated with significant morbidity for affected patients. Common etiologies include diabetes (29%), idiopathic (36%), and postsurgical causes (13%).²

High-frequency, low-intensity gastric electrical stimulation (GES) (Enterra™; Medtronic, Inc., Minneapolis, MN) has been shown in previously published literature to improve symptoms and quality of life in patients with medically refractory gastroparesis.^{3–8} However, the definition of what constitutes “improvement” in the literature is ill-defined and has lacked standardization in published studies to date. As a result, a significant amount of skepticism exists regarding the usefulness of GES in the treatment of gastroparesis.^{9, 10} We sought to define the degree of improvement in various subgroups of patients based on the degree of change in total symptom severity score (TSS) as well as patients’ perception of improvement, and to compare its effectiveness (as measured by these two benchmarks of treatment success) in patients with diabetic and idiopathic etiologies.

Methods

We conducted a retrospective analysis of 56 patients with medically refractory gastroparesis who underwent insertion of a gastric electrical stimulator (Enterra™) during the study period. All patients met criteria for treatment using GES under the guidelines suggested by the Food and Drug Administration in its classification of the Enterra™ gastric electrical stimulator as a humanitarian use device.¹¹ These included, at a minimum: (i) Symptoms compatible with a diagnosis of gastroparesis, (ii) An abnormal nuclear medicine gastric emptying study, and (iii) The exclusion of mechanical obstruction as a cause. In addition, all patients considered for GES were refractory to a trial of a minimum of six months of medical therapy. Before the study, approval was obtained from the Charleston Area Medical Center/West Virginia University Institutional Review Board.

Descriptive statistics (mean, median, range, and standard deviation) were used to examine age, race, comorbidities, preoperative diagnostic testing, and any preoperative medical and surgical treatments used before the institution of GES. Total symptom severity (TSS) scores, as initially described by Lin et al.,⁴ was used to evaluate response to therapy. Briefly, the TSS uses a five-point Likert scale to evaluate the severity of six symptoms of gastroparesis: epigastric pain, nausea, vomiting, bloating, early satiety, and postprandial fullness (0–4 points for each symptom, maximum score 24).⁴ The TSS was assessed both at initial presentation and at the last follow-up visit. The scores were then analyzed to determine median change from baseline using the Wilcoxon signed rank test. Initial and repeat TSS were subsequently grouped into the following four severity categories: 0–10, 11–14, 15–18, and 19–24.

For the purpose of showing improvement based on change in TSS score, these four diagnostic severity categories were devised based on both initial and repeat TSS scores. Before creating categories, quartiles were generated for initial and repeat TSS data. For repeat TSS, the 50th percentile was found to be a TSS of 12. Therefore, based on quartile ranges ($Q_1 = 10$ and $Q_3 = 18$) the categories were made as 0–10, 11–14, 15–18, and 19–24 so that approximately the same number of patients fell into each category for repeat TSS. The same categories were also used to categorize the patients based on their initial TSS as well.

An improvement in TSS was defined as a shift to a lower category from the initial TSS to that obtained at the last follow-up visit after institution of GES. Patients' perception of improvement was measured by asking the patient if they felt "much improved, somewhat improved, or not improved." Patients' perceptions were then compared with TSS results using χ^2 test. A *t* test compared demographics and symptoms among patients who showed improvement or non-improvement based on both TSS and patient perception. Symptom etiology as a predictor of improvement was measured using logistic regression. A *P* value of 0.05 was considered statistically significant. Statistical analysis was performed using SAS 9.3 (SAS Institute, Cary, NC).

Results

Fifty-six patients were treated using GES. The mean age was 48 years. In all, 75 per cent of patients in the series were male. Symptoms at diagnosis included vomiting (93%), nausea (100%), satiety (64%), bloating (64%), postprandial fullness (61%), pain (79%), weight loss (52%), and constipation (41%). Diabetes was the most frequent comorbidity (68%). Medical treatment was attempted in all patients prior to consideration of GES therapy and all patients were evaluated using prior to GES therapy via diagnostic endoscopy. Mean follow-up for the series was 37.8 months (range 2.1–97.4 months).

The initial median TSS for all 56 patients in the series was 21. The median TSS after GES treatment was significantly reduced to 13.5 ($P < 0.0001$). In addition, the median symptom change from baseline to follow-up was documented for all patients and found to be significant for each individual symptom ($P < 0.0001$). No correlation was noted between likelihood of success or failure and gastric emptying times ($P = 0.32$). The overall reduction in TSS for both diabetic and idiopathic patient groups was also statistically significant ($P < 0.0001$) (Table 1).

TSS scores, both initial and at follow-up, were categorized into the four proposed severity categories (0–10, 11–14, 15–18, and 19–24) and each patient's score analyzed to determine if improvement occurred, with the definition of improvement being a shift to a lower severity category. If the patient dropped to a lower severity category at the time of the repeat assessment of TSS, then that patient was considered to be improved. If the patient remained in the same severity category then that patient was considered to have failed to improve (Table 2).

In addition, patients' overall perception of improvement *versus* no improvement was recorded as a logical measure of treatment success. This perception of improvement was compared with the observed change in the TSS. A perception of improvement occurred in 38 patients who moved from a higher TSS category to a lower category, whereas 18 patients perceived no improvement ($P < 0.001$), thus establishing a statistically significant association between movement to a lower TSS severity category and patients' perception of improvement.

There were four observed exceptions. One of the 38 patients who shifted to a lower TSS category and showed improvement did not report a change based on perception. This patient

had multiple significant comorbid conditions, including psychosis, so the accuracy of self-reported subjective measures of success must be questioned. Three out of the 18 patients who did not improve based on a shift to a lower TSS category did have an improvement based on their own perception. The change for these patients from their initial TSS to the repeat TSS was too small to result in a shift to a lower category (Table 2).

Our interpretation of our data changed radically with the employment of these two metrics as measures of success. The diabetic population was still statistically likely to improve with GES, as seen in 28 of 32 patients ($P < 0.0001$) (Table 3) when improvement was defined as movement to a lower symptom severity category. Similarly, 30 of 32 diabetic patients improved when patient perception of improvement was used ($P < 0.0001$) (Table 3).

Equally impressive were the results when the idiopathic population was reanalyzed using these two metrics of improvement. In both groups (those measured by a shift to a lower category and those measured by patient perception of improvement) only 10 of 24 patients were considered improved ($P = 0.0003$ for movement to lower severity category, $P < 0.0001$ for improvement based on patient perception) (Table 3). These data in fact demonstrated that patients with idiopathic gastroparesis were statistically more likely *not* to improve with GES than to respond to GES.

Multivariate logistic regression was used to furthermore assess the likelihood of improvement (using the criteria of both movement to a lower symptom severity category and in patients' overall perception of improvement) with the two different etiologies—diabetic and idiopathic. Based on these data, diabetic patients were 14 times more likely to improve based on TSS scores (movement to lower category) than idiopathic patients [odds ratio (OR) 14.0, 3.34–58.77, $P = 0.0003$], and 45 times more likely to improve based on patients' perception of improvement than idiopathic patients (OR 45, 5.26–385.17, $P = 0.0005$). These data demonstrate that patients with idiopathic gastroparesis were statistically more likely *not* to improve with GES than to respond to GES. These results were highly significant, but were based on a small sample size, which was likely responsible for the wide confidence intervals observed (Table 3).

The overall morbidity in the series was 3.6 per cent, occurring in two of the 56 cases. One patient, a brittle diabetic, developed erosion of the skin over the device resulting in exposure of the generator, which required reoperation and replacement at an alternative site. A second patient suffered erosion of one of the leads through the full thickness of the gastric wall. This manifested as mild hematemesis and was diagnosed on upper endoscopy. His device was explanted fully and was not reinserted as he had not responded to GES.

Five additional patients who suffered no overt complications also underwent device explantation—three by request due to treatment failure and two who had the device removed during subtotal gastrectomy due to failure to respond and intractability of symptoms. All of these six explantations occurred in the idiopathic group.

Discussion

Numerous studies have been published in support of GES for the treatment of medically refractory gastroparesis.^{3–8, 12} Despite this fact, the device remains available only under a humanitarian device exemption due to lack of conclusive proof of effectiveness.¹¹ The majority of published series to date have been open-label observational studies, lacking in appropriate controls, involving small series of patients, and have been heterogeneous in terms of methodology and in the method of assessment of successful response to GES.^{7, 8, 10, 11} Subsequently, the role of GES in clinical practice remains controversial and is not universally accepted as the standard of care for the treatment of medically refractory gastroparesis.^{9, 10}

Measures of success in previously published studies have included reduction in the need for nutritional support, change in weight, improvement in gastric emptying, and symptom improvement.^{3–8, 12, 13} Symptom improvement (logically) has been the most commonly assessed measure of a successful response. Several methods of assessment of symptom improvement have been used, with no clear consensus as to which method, if any, is superior. Symptom-based criteria for assessment of disease severity and response have included improvements in TSS, vomiting severity score, nausea severity score, quality of life scores (SF-36), and the Gastroparesis Cardinal Symptom Index (GCSI).^{3–8, 12, 13} Although the GCSI has been previously validated and the TSS has not (although it has correlated well when compared with the GCSI and has been referenced more often than GCSI in the GES literature),^{4, 6, 14} we chose to use TSS as our preferred metric for symptom improvement. The reason for this choice was that the symptom of abdominal pain, which is both a significant symptom of gastroparesis and a previously reported predictor of GES failure, is measured as a component of TSS but not GCSI.^{4, 6, 14} For this reason, TSS was felt to be a superior means for assessing disease severity and response to therapy.

Several previously published studies report statistically significant reductions in symptom scores after institution of GES.^{3–8, 12} What has not been determined however is exactly how *much* of decrease in TSS is clinically significant. Previous studies have reported the mean reductions in TSS that were observed,³ but have not analyzed where the cutoff point lies with respect to clinical success *versus* failure. Other authors have arbitrarily set a number for reduction in TSS (three points) for success *versus* failure of GES,⁶ but this lacks an evidence-based rationale. In our series, with univariate analysis, statistically significant reductions in TSS from baseline to last follow-up were observed in the series as a whole, as well with each individual symptom. When this data were then furthermore interpreted by dividing the groups into diabetic and idiopathic etiologies, we demonstrated similar results in terms of improvement. Although the changes were less dramatic in the idiopathic group, they remained highly statistically significant ($P < 0.0001$). However, when the clinical response was assessed via the two clinical metrics—movement to a lower TSS symptom severity category or assessment of improvement based on patients' perception—no benefit was seen in the idiopathic group. In fact, based on the results of both univariate and multivariate logistic regression analysis, idiopathic patients were statistically more likely to fail than to improve with GES (Table 3). These data demonstrate that GES patients may experience reductions in TSS that reach *statistical* significance across a series, but ultimately

do not derive significant *clinical* improvement with GES. Therefore, validated criteria for clinical improvement may represent more useful tools for assessing GES efficacy.

Limitations of this study include its retrospective design and its relatively small sample size, although the sample size is consistent with or larger than many of the published series to date. Also, we acknowledge the need for furthermore prospective validation of the use of our proposed symptom severity categories and the movement from a higher to a lower severity category as a valid means to assess clinical improvement.

Conclusions

This study makes several useful points for consideration. These data demonstrate that GES patients may experience reductions in TSS that reach statistical significance across a series, but ultimately do not derive clinical improvement with GES. This represents a potential major shortcoming in analyzing and applying the results of previously published studies to date. As demonstrated in previously published studies, the benefits of GES in the diabetic population were impressive even with the application of more strict criteria for improvement. However, the finding that idiopathic patients were more likely *not* to improve than to improve with GES when more strict clinical criteria were used to define improvement suggests that furthermore prospective study is needed into the role of GES in idiopathic patients. Also, we acknowledge the need for furthermore prospective validation of the use of our proposed symptom severity categories and the movement from a higher to a lower severity category as a valid means to assess clinical improvement.

REFERENCES

1. Hornbuckle K, Barnett J. The diagnosis and work-up of the patient with gastroparesis. *J Clin Gastroenterol.* 2000; 30:117–124. [PubMed: 10730917]
2. Jones M, Maganti K. A systematic review of surgical therapy for gastroparesis. *Am J Gastroenterol.* 2003; 98:2122–2129. [PubMed: 14572555]
3. Abell T, McCallum R, Hocking M, et al. Gastric electrical stimulation for medically refractory gastroparesis. *Gastroenterology.* 2003; 125:421–428. [PubMed: 12891544]
4. Lin Z, Forster J, Sarosiek I, et al. Treatment of diabetic gastroparesis by high frequency gastric electrical stimulation. *Diabetes Care.* 2004; 27:1071–1076. [PubMed: 15111523]
5. Brody F, Vaziri K, Saddler A, et al. Gastric electrical stimulation for gastroparesis. *J Am Coll Surg.* 2008; 207:533–538. [PubMed: 18926455]
6. McKenna D, Beverstein G, Reicheldelderfer M, et al. Gastric electrical stimulation is an effective and safe treatment for medically refractory gastroparesis. *Surgery.* 2008; 144:566–574. [PubMed: 18847640]
7. Chu H, Lin Z, Zhong L, et al. Treatment of high-frequency gastric electrical stimulation for gastroparesis. *J Gastroenterol Hepatol.* 2012; 27:1017–1026. [PubMed: 22128901]
8. O'Grady G, Egbuji J, Du P, et al. High-frequency electrical stimulation for the treatment of gastroparesis: a meta analysis. *World J Surg.* 2009; 33:1693–1701. [PubMed: 19506941]
9. Stanghellini V. Unfulfilled wishes by gastric electrical stimulation. *Clin Gastroenterol Hepatol.* 2011; 9:447–448. [PubMed: 21238610]
10. Sofer E. Gastric electrical stimulation for gastroparesis. *J Neurogastroenterol Motil.* 2012; 18:131–137. [PubMed: 22523722]
11. Camilleri M, Parkman H, Shafi M, et al. Clinical guideline: management of gastroparesis. *Am J Gastroenterol.* 2013; 108:18–38. [PubMed: 23147521]

12. Abell T, Lou J, Tabbaa M, et al. Gastric electrical stimulation for gastroparesis improves nutritional parameters at short, intermediate, and long-term follow up. *J Parenter Enteral Nutr.* 2003; 27:277–281.
13. O'Loughlin P, Gilliam A, Shaban F, et al. Pre-operative gastric emptying time correlates with clinical response to gastric electrical stimulation in the treatment of gastroparesis. *Surgeon.* 2013; 11:134–140. [PubMed: 23206591]
14. Abidi N, Starkebaum W, Abell T. An energy algorithm improves symptoms in some patients with gastroparesis and treated with gastric electrical stimulation. *Neurogastroenterol Motil.* 2006; 18:334–338. [PubMed: 16553589]

Table 1

Comparison of TSS Before and After Initiation of GES

	Baseline	Last Follow-up	P Value
Total TSS (entire series)	21(13–24)	13.5(0–23)	<0.0001
Vomiting	4(0–4)	2(0–4)	<0.0001
Nausea	4(3–4)	2(0–4)	<0.0001
Early satiety	4 (2–4)	3(0–4)	<0.0001
Bloating	4(0–4)	2(0–4)	<0.0001
Postprandial fullness	4(1–4)	3(0–4)	<0.0001
Pain	4(0–4)	2(0–4)	<0.0001
TSS for Diabetic group	22(13–24)	12(0–21)	<0.0001
TSS for Idiopathic group	21(17–24)	18(1–23)	<0.0001
Diabetic group (n = 32)			
Vomiting	4(0–4)	2(0–4)	<0.0001
Nausea	4(3–4)	2(0–4)	<0.0001
Early satiety	4(2–4)	2(0–4)	<0.0001
Bloating	4(0–4)	2(0–4)	<0.0001
Postprandial fullness	4(1–4)	2(0–4)	<0.0001
Pain	4(0–4)	2(0–4)	<0.0001
Idiopathic group (n = 24)			
Vomiting	4(0–4)	3(0–4)	0.0001
Nausea	4(3–4)	3(0–4)	0.0005
Early satiety	4(3–4)	3(0–4)	0.0156
Bloating	4(0–4)	3(0–4)	0.0039
Postprandial fullness	3.5(2–4)	3(1–1)	0.0156
Pain	4(0–4)	3(0–4)	0.0156

All values use the **median** (range), significance set at $P < 0.05$.

Table 2

Comparing TSS with Patients' Perception of Improvement

	Improvement Based on TSS Scores (Movement to Lower Category)		
	Improved	Not Improved	Total
Much Improved	24 (100.00%)	0 (0%)	24
Somewhat Improved	13 (81.25%)	3 (18.75%)	16
Not Improved	1 (6.25%)	15 (93.75%)	16
Total	38	18	56

Improvement based on patients' perception of improvement.

Table 3**Etiology as a Predictor of Improvement with GES**

Improvement (Defined as Movement to a Lower Category of Symptom Severity)			
Univariate analysis	No improvement (n = 18) (32.1%)	Improvement (n = 38) (67.9%)	P value
Diabetic	4 (22.2%)	28 (73.7%)	<0.0001*
Idiopathic	14 (77.8%)	10 (26.3%)	0.0003**
Logistic regression	OR		P value
Diabetic (compared to idiopathic)		14(3.335–58.768)	0.0003*
Improvement (defined by patients' perception of improvement)			
Univariate analysis	No improvement (n = 16) (28.6%)	Improvement (n = 40) (71.4%)	P value
Diabetic	2 (12.5%)	30 (75.0%)	<0.0001*
Idiopathic	14 (87.5%)	10 (25.0%)	<0.0001**
Logistic regression	OR		P value
Diabetic (compared to idiopathic)		45.0 (5.257–385.170)	0.0005*

* Statistical significance set at *P* value < 0.05; mean follow-up 37.8 months (range 2.1–97.4).

** *P* value indicates that idiopathic patients are statistically significantly likely *not* to improve with GES.