Gastroparesis: Clinical Evaluation of Drugs for Treatment

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact Maureen Dewey at 301-796-0845.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

August 2019
Clinical/Medical
Revision 1

Gastroparesis: Clinical Evaluation of Drugs for Treatment

Guidance for Industry

Additional copies are available from:

Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353; Email: druginfo@fda.hhs.gov
https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

August 2019
Clinical/Medical
Revision 1

TABLE OF CONTENTS

Gastroparesis: Clinical Evaluation of Drugs for Treatment Guidance for Industry[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

I. INTRODUCTION

The purpose of this guidance is to assist sponsors in the clinical development of drugs for treating idiopathic and diabetic gastroparesis.[2] Specifically, this guidance addresses FDA's current recommendations regarding clinical trial designs and clinical endpoint assessments to support developing gastroparesis drugs.

This draft guidance is intended to serve as a focus for continued discussions among the responsible FDA divisions in the Office of New Drugs, pharmaceutical sponsors, the academic community, and the public.[3]

This guidance revises the draft guidance for industry of the same name issued in July 2015. Changes from the previous draft reflect FDA's current thinking about developing clinical outcome assessment tools and statistical considerations for using those tools to assess primary and secondary efficacy endpoints.

This guidance does not address detailed patient-reported outcome (PRO) instrument development and validation; these topics are addressed in the guidance for industry *Patient-*

[1] This guidance has been prepared by the Division of Gastroenterology and Inborn Errors Products in the Center for Drug Evaluation and Research at the Food and Drug Administration.

[2] For the purposes of this guidance, all references to *drugs* include both human drugs and therapeutic biological products unless otherwise specified.

[3] In addition to consulting guidances, sponsors are encouraged to contact the division to discuss specific issues that arise during the development of drugs to treat gastroparesis.

Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims (December 2009).[4]

More details regarding statistical analysis and clinical trial design are addressed in the ICH guidances for industry *E9 Statistical Principles for Clinical Trials* (September 1998) and *E10 Choice of Control Group and Related Issues in Clinical Trials* (May 2001), respectively.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

Gastroparesis is a disorder characterized by delayed gastric emptying (DGE) in the absence of mechanical obstruction. Symptoms are chronic with episodic exacerbation (Parkman et al. 2004). The idiopathic form of the disorder, which accounts for the greatest number of cases (Karamanolis et al. 2007), predominantly affects young adult females. Gastroparesis is also frequently associated with diabetes (diabetic gastroparesis), which likely occurs because of impaired neural control of gastric motility (Parkman et al. 2004). In addition, acute hyperglycemia has the potential to slow gastric emptying and decrease the effects of prokinetic drugs (Camilleri 2010).

The core signs and symptoms of gastroparesis are nausea (92 to 96 percent), vomiting (68 to 88 percent), postprandial fullness (54 to 77 percent), early satiety (42 to 86 percent), and upper abdominal pain (36 to 85 percent) (Hoogerwerf et al. 1999; Anaparthy et al. 2009). Patients may experience any combination of signs and symptoms with varying degrees of severity. Pain is more prevalent in patients with idiopathic gastroparesis than it is in patients with diabetic gastroparesis. Patients with diabetic gastroparesis may experience further derangement of glucose control because of unpredictable gastric emptying and altered absorption of orally administered hypoglycemic drugs.

Because the signs and symptoms of gastroparesis overlap with other gastrointestinal conditions, gastroparesis may be incorrectly diagnosed as bowel obstruction, functional dyspepsia, irritable bowel syndrome, or peptic ulcer disease. In a patient with signs and symptoms suggestive of gastroparesis, a finding of DGE in the absence of an obstruction or alternative diagnosis provides critical support for the diagnosis of gastroparesis and can be assessed using either gastric emptying scintigraphy, the gastric emptying breath test, or the SmartPill motility testing system.

There is an urgent medical need for development of safe and effective therapies to treat patients with gastroparesis.

[4] We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents.

III. ENDPOINTS AND TRIAL DESIGN FOR GASTROPARESIS CLINICAL TRIALS

Primary efficacy assessments for adequate and well controlled trials must be well defined and reliable.[5] Because gastroparesis is a symptomatic condition, a well-defined and reliable PRO instrument that measures all the clinically important signs and symptoms of gastroparesis would be the ideal primary efficacy-assessment tool in clinical trials used to support labeling claims for treating gastroparesis.[6] However, we are currently not aware of such a measure. Until an appropriate PRO instrument for gastroparesis becomes available, sponsors should consider the strategies discussed in the following sections when designing gastroparesis clinical trials. Sponsors may also wish to review FDA's Center for Drug Evaluation and Research Clinical Outcome Assessment (COA) Drug Development Tool Qualification Program web page for information on qualified tools or tools currently under development.

Sponsors may wish to include and evaluate well-defined PRO instruments assessing the relevant and important signs and symptoms in early drug development — and evaluate the results in dose-ranging phase 2 trials or stand-alone noninterventional studies — to support their future use in phase 3 trials. We encourage early and regular discussions with FDA regarding the development of these PRO instruments.

Because gastroparesis manifests as more than one core sign or symptom, the effect of new drugs intended to treat gastroparesis on each core sign and symptom should be assessed. Early phase trials should help inform which of the core signs and/or symptoms should be included as prespecified endpoints intended to support labeling claims, based on which signs or symptoms the treatment is likely to improve. It is important to show that even drugs intended to treat only a subset of the core signs or symptoms, based on the mechanism of the drug, do not worsen the remaining signs or symptoms of gastroparesis. For example, a drug may be expected to improve gastroparesis-related nausea and vomiting but not abdominal pain based on its mechanism of action. In this scenario, clinical studies should demonstrate that nausea and vomiting improved and that the treatment did not worsen the symptoms of abdominal pain, postprandial fullness, and early satiety.

The following sections provide recommendations regarding trial design, trial populations, outcome assessment measures, trial endpoints, and statistical considerations.

A. Trial Design

In general, the trial design should consist of a randomized, double-blind, placebo-controlled trial and should include a 1- to 2-week screening period. The screening period can be used for investigators to establish the presence and persistence of trial-entry criteria and for patients to gain experience completing the PRO instruments employed in the trial and demonstrate adequate

[5] 21 CFR 314.126.

[6] See the guidance for industry *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims.*

understanding of and compliance with completing these instruments. The screening period assessments of gastroparesis signs and symptoms can serve as the baseline values used in the analyses of the primary endpoint (see section III. D., Trial Endpoints). FDA recommends a baseline assessment period of at least 7 days. The primary endpoint should measure the change in signs and symptoms from baseline over a treatment period of at least 12 weeks' duration.

Trial designs should address the need for maintenance treatment to prevent recurrence of signs or symptoms.

Endpoint assessment should be based on patients' daily reporting to avoid recall error, and the protocol should state whether rescue medication (i.e., protocol-specified therapy for continued exacerbation of symptoms that is standardized across study sites) is allowed. Daily diaries should be collected throughout the entire trial.

In addition, we recommend a randomized, controlled, long-term safety study of 12 months' duration, with appropriate prespecified provisions for rescue medications, which should be conducted before submitting a new drug application.

B. Trial Populations

Idiopathic and diabetic gastroparesis patients should be studied in separate clinical trials. In general, diabetic gastroparesis patients experience the same core signs and symptoms as patients with idiopathic gastroparesis, but individual signs and symptoms may occur more often or with greater severity in one population compared with the other, and the degree of diabetic control can also confound results. To fully describe safety and efficacy in each population, we recommend separate trials. Because idiopathic and diabetic gastroparesis are closely related conditions, a single phase 3 trial in each population with demonstration of reliable and clinically meaningful results may support approval for both indications.[7]

We recommend that trial-entry criteria include the following:

- The trial populations should have a clinical diagnosis of idiopathic or diabetic gastroparesis (for the individual trials) based on a documented history of gastroparesis symptoms, exclusion of other potential etiologies, and DGE (Abell et al. 2008; Parkman et al. 2004). To optimize the ability to demonstrate a treatment effect, the trial should enroll patients with higher symptom severity (moderate to severe). Because there are currently no accepted definitions of gastroparesis severity, the sponsor should provide a justification for the severity index selected, including what defines moderate and severe symptoms.

- Diabetic gastroparesis patients should have controlled and stable blood glucose levels. Patients prone to acute hyperglycemic events may confound interpretation of the therapeutic effect of the drug.

[7] See the guidance for industry *Providing Clinical Evidence of Effectiveness for Human Drug and Biological Product*s (May 1998).

- Patients on opioids should be excluded because opioid use may affect gastrointestinal motility and potentially confound results.

C. Approach for Outcome Assessment Measures

Until a well-defined and reliable PRO instrument that measures all the clinically important signs and symptoms of gastroparesis is available, we recommend that the five core signs and symptoms of gastroparesis — nausea, vomiting, postprandial fullness, early satiety, and abdominal pain — be included as endpoints in well-controlled clinical trials (Karamanolis et al. 2007; Hoogerwerf et al. 1999; Anaparthy et al. 2009). Sponsors should identify and empirically justify the questionnaire items (and their wording) used to assess signs or symptoms of gastroparesis that will be included in the trial.[8]

Each sponsor should propose a primary endpoint definition (see section III. D. Trial Endpoints) and a method for measuring each of the five signs and symptoms as described below. Piloting the proposed instrument(s) in phase 2 trials can provide an opportunity to evaluate the ability of the instrument(s) to detect change, provide guidelines for interpretation of clinically meaningful within-patient change, and confirm the endpoint definition. Pilot results can further inform plans for implementation of the proposed instrument(s) in the phase 3 trials. Wording of the questionnaire should be carefully thought out so the questions or requests do not overlap in their measurement concepts (e.g., postprandial fullness and early satiety), and the concepts should be well-defined so that they are interpreted in a consistent way by patients (i.e., the questionnaire should include definitions for postprandial fullness, early satiety, or other terms that may vary in their interpretation among patients). Each core sign and symptom should be separately measured and documented in the clinical trial.

The sponsor should also specify the mode of data collection that will be used by patients to record their daily signs and symptoms (e.g., electronic diary).

All signs and symptoms except vomiting should be rated by severity. For example, question or request item responses can range from 0 for no symptom to 4 for the most severe symptom (0=none; 1=mild; 2=moderate; 3=severe; and 4=very severe) or have a numerical rating scale from 0 to 10, where 0 reflects the absence of the symptom and 10 reflects the worst possible symptom experience. When possible, the rating scale should be consistent across the core signs and symptoms. We recommend that reporting of vomiting in a daily symptom diary be measured by frequency rather than severity. Frequency should be reported as the exact number of times over a 24-hour period, and a clear definition of what is considered "one time" of vomiting should be provided to patients to ensure consistency both within and between patients in reporting the number of times vomiting has occurred. The severity of nausea, postprandial fullness, early satiety, and abdominal pain should be recorded based on the patient's worst experience over a 24-hour period.

[8] See the guidance for industry *Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims*.

D. Trial Endpoints

1. Primary Endpoints

Changes in patient-reported signs and symptom scores should form the basis of the primary efficacy assessment in therapeutic trials for idiopathic and diabetic gastroparesis. The primary endpoint should be based on patients' core signs and symptoms or a subset of them. Gastric emptying time should not be used as a primary efficacy endpoint because changes in gastric emptying time are not associated with the changes in the clinically important signs and symptoms in patients with gastroparesis.

The primary endpoint should measure change in signs and symptoms from baseline. The analysis plan should include an evaluation of treatment effect throughout the 12-week study period.

We recommend the use of an endpoint that is based on core signs and symptoms. This may be based on prespecified core signs and symptoms or a symptom severity summary score (excluding vomiting) and vomiting frequency (collected as a continuous variable). The primary endpoint should not be limited to a single sign or symptom. If sponsors propose a summary score, they should evaluate question-level (or request-level) responses to determine whether individual questions (or requests) overly influence the total score. Currently, we do not have evidence to recommend one approach over the other. Scores based on severity should be analyzed separately from those based on frequency (e.g., vomiting).

2. Secondary/Other Endpoints

FDA recommends that changes from baseline in the individual signs and symptoms that are not assessed as part of the primary endpoint be measured as secondary endpoints. Therefore, the primary and secondary endpoints should include an evaluation of all five core signs and symptoms. Change in gastric emptying time can be measured as a secondary endpoint if desired (Abell et al. 2008). The prespecified plan should address an analysis of the remaining core signs or symptoms that are not included in the primary endpoint.

3. Defining Clinically Meaningful Within-Patient Changes in Sign and Symptom Scores

To aid in the interpretation of the results, sponsors should determine the amount of change that is meaningful to patients, in a total summary score or in individual sign and symptom scores. Ideally, this should be based on actual data and established in advance of phase 3 trials, so clinically meaningful within-patient change thresholds may be prespecified. There are two clinically meaningful change thresholds of interest: one for a clinically important improvement from baseline and one for a clinically important deterioration from baseline. Depending on the proposed mechanism of action of the drug and trial objectives, a proposed threshold can specify some level of improvement in each of the five core signs and symptoms, or it can specify some level of improvement in a subset of those core signs and symptoms. Worsening of core signs

and symptoms would be inconsistent with the expected clinical benefit and taken into account when evaluating benefit and risk.

We recommend the use of an anchor-based approach, typically using phase 2 trial data, to estimate clinically meaningful change. For this approach, we recommend including in phase 2 and 3 trials multiple anchor scales, such as patient global impression of severity (PGIS) and patient global impression of change (PGIC) scales, with the intent of providing accumulated evidence to help interpret a clinically meaningful within-patient score change. In contrast to a PGIC scale, a PGIS scale is not subject to recall error and can also be used to assess change from baseline data. The PGIS scale is the preferred anchor scale over the PGIC scale; however, there is no perfect anchor scale, and it is helpful to include multiple anchor scales for anchor-based analyses.

The following item, which could be asked of patients (following the assessment schedule and recall period of the prespecified endpoint) and at baseline, is an example of a PGIS scale:

> "Please choose the response below that best describes the severity of your gastroparesis symptoms over the [insert appropriate recall period here]."

Sponsors can consider the following response options to this item: 0=none; 1=mild; 2=moderate; 3=severe; and 4=very severe.

The following item, which could be asked weekly of patients, is an example of a PGIC scale:

> "Please choose the response below that best describes the overall change in your gastroparesis symptoms since you started taking the study medication."

Sponsors can consider the following response options to this item: much better, a little better, no change, a little worse, much worse.

Sponsors should determine the clinically meaningful within-patient change threshold range using anchor-based methods (e.g., patient global impression scale as an anchor), supplemented with empirical cumulative distribution functions (eCDFs) of within-patient score change. Separate eCDF curves should be generated for each meaningful anchor category (e.g., improved, no change, worsened) using data pooled across trial arms.

E. Statistical Considerations

To evaluate daily diary assessments created during a trial, an adequate number of the assessments should be available. The sponsor can determine this number based on evidence derived from the particular PRO assessment used in the trial. For example, if a weekly summary score is used, in general, the sponsor should provide assessments from at least 4 of the 7 days. However, evidence from a particular PRO assessment may support the need for data from a higher number of days for that instrument to provide reliable results.

The efficacy analysis plan should clearly define how patients who take rescue medication will be considered in the final analysis. Sponsors also should propose methods for handling missing data, including missing rescue medication data and missing PRO data at both question or request and instrument levels, in the analysis plan. Sponsors should consider different approaches before the trial is initiated and the properties of these approaches should be evaluated.

We recommend that sponsors analyze the primary and secondary endpoints as continuous or ordinal variables; we do not recommend the use of percentage change. In general, a traditional responder analysis would not be appropriate unless the targeted response is complete resolution of signs and symptoms. In addition, we encourage the use of baseline values and other covariates to improve the efficiency of primary and secondary endpoint analyses.

Additionally, sponsors should submit supportive descriptive analyses (i.e., graphs of eCDFs of within-patient change from baseline for primary and secondary endpoints by treatment arm) for FDA review.

REFERENCES

Abell TL, RK Bernstein, T Cutts, G Farrugia, J Forster, WL Hasler, RW McCallum, KW Olden, HP Parkman, CR Parrish, PJ Pasricha, CM Prather, EE Soffer, R Twillman, and AI Vinik, 2006, Treatment of Gastroparesis: A Multidisciplinary Clinical Review, Neurogastroenterol Motil 18:263–283.

Abell TL, M Camilleri, K Donohoe, WL Hasler, HC Lin, AH Nauer, RW McCallum, T Nowak, ML Nusynowitz, HP Parkman, P Shreve, LA Szarka, WJ Snape, and HA Ziessman, 2008, Consensus Recommendations for Gastric Emptying Scintigraphy: A Joint Report of the American Neurogastroenterology and Motility Society and the Society of Nuclear Medicine, Am J Gastroenterol, 103:753–763.

Anaparthy R, N Pehlivanov, J Grady, H Yimei, and PJ Pasricha, 2009, Gastroparesis and Gastroparesis-Like Syndrome: Response to Therapy and Its Predictors, Dig Dis Sci, 54(5):1003–1010.

Hoogerwerf WA, PJ Pasricha, AN Kalloo, and MM Schuster, 1999, Pain: The Overlooked Symptom in Gastroparesis, Am J Gastroenterol, 94:1029–1033.

Karamanolis G, P Caenepeel, J Arts, and J Tack, 2007, Determinants of Symptom Pattern in Idiopathic Severely Delayed Gastric Emptying: Gastric Emptying Rate or Proximal Stomach Dysfunction? Gut, 56:29–36.

Parkman HP, WL Hasler, and RS Fisher, 2004, American Gastroenterological Association Technical Review on the Diagnosis and Treatment of Gastroparesis, Gastroenterology, 127(5):1592–1622.

Parkman HP, K Yates, WL Hasler, L Nguyen, PJ Pasricha, WJ Snape, G Farrugia, KL Koch, J Calles, TL Abell, RW McCallum, L Lee, A Unalp-Arida, J Tonascia, and F Hamilton, 2011, Similarities and Differences Between Diabetic and Idiopathic Gastroparesis, Clin Gastroenterol Hepatol, 9:1056–1064.