Reviewer's report

Title: Cost and burden of gastroesophageal reflux disease among patients with persistent symptoms despite proton pump inhibitor therapy: an observational study in France

Version: 1 Date: 10 October 2012

Reviewer: Wieslaw Bochenek

Reviewer's report:

Cost burden of gastroesophageal reflux disease among patients with persistent symptoms despite proton pump inhibitor therapy: observational study in France Stanislas Bruley des Varannes et al.

This observational 12 months study was designed and undertaken to address specific issue of cost and burden of GERD in population of patients who did not respond adequately to treatment with PPI agents. Total of 262 patients were enrolled and followed at various intervals using a number of assessment instruments that included RESQ-7 at baseline, 3, 6, 9 and 12 months; HADS at baseline and 12 months; HRQL using SF-36v2 and EQ-5D at baseline, 6 and 12 months; and WPAI-GERD at baseline 3, 6, 9 and 12 months.

Following are my comments and suggestions related to the study design, evaluation of results and discussion.

Patients for the study were recruited from primary care physician (PCP) and gastroenterology practices. It is likely that the diagnosis and treatment used by these two types of practices are different with gastroenterologists evaluating patients more thoroughly, likely to scope patients and then more carefully applying treatment. Patient visits reported for PCPs were almost double compared to patient visits for gastroenterology sites. Some 76% of patients had endoscopy within 6 month prior to entering the study and only 50% were reported as having characteristics of reflux esophagitis and another 16% with Barrett's esophagus – most likely related to reflux. Is it then certain that the objective of studying patients with GERD symptoms, as implied by study title was met, if only 66% were documented to have endoscopically confirmed GERD?

Information on PPI treatment states that the majority of patients (86%) reached maximum dose. No explanation is given what constitutes maximum. Is it the labeled use, is it the off-label use, is it once daily use or twice daily. For 29% of patients the treatment was not optimized. Does this mean no maximum dose was reached?

The treatment used is stated as essentially remaining unchanged over the 12 month period. Was this information collected or presumed and if collected how reliable the information is?
Study information was collected using CRFs completed at all visits. Was this paper CRF or CRF filled out electronically? Who checked the CRFs for completeness of information and how complete were the CRFs and the survey forms. How the two relate? How much information presented in the paper come from CRFs and how much from survey forms?

Considerable effort was made to use questionnaires and other assessment instruments yet no information is provided on the completeness of the information collected. A recall method was used that is considered very unreliable. The HADS was collected but no results are presented nor discussed.

Work productivity was assessed yet 51% of patients were not employed. It is uncertain how reliable is the information and no details are provided which patients were included in this assessment and how the data were analyzed. Was productivity assessed in unemployed/retired, if not this should be stated?

The most interesting is the information on the cost per patient. The authors did not discuss why the cost 1-12 months of the study is so much different from the pre-study period. Is it because pre-study collected information for 6 month while on the study information covers 12 month. If so, it might be more meaningful averaging cost per month/patient for the two periods.

In the discussion, the authors state that the strength of the study relates to the large sample size. This is doubtful given the demographic diversity of the study population where the subgroups, if analyzed, would constitute rather small numbers. This compares unfavorably with some other similar studies where much larger number of patients participated. The recall method used in collecting the questionnaire information is very unreliable and no information is provided on the number of missing data and how the missing data were treated. Similarly no information is provided if patients who dropped from the study were considered in the overall assessment.

It is my opinion that the study has significant methodological issues aside from the fact that the study population and its size do not allow for meaningful analysis of the data. Novelty of the information is questionable as well and for these reasons I recommend that the paper should not be accepted for publication.