Arab Journal of Gastroenterology 14 (2013) 165–168

Contents lists available at ScienceDirect

Arab Journal of Gastroenterology

journal homepage: www.elsevier.com/locate/ajg

Original Article

Zero inflated statistical count models for analysing the costs imposed by GERD and dyspepsia

Alireza Akbarzadeh Baghban [a,*], Asma Pourhoseingholi [b], Farid Zayeri [c], Sara Ashtari [d], Mohammad Reza Zali [d]

[a] Department of Basic Sciences, School of Rehabilitation, Shahid Beheshti University of Medical Science, Tehran, Iran
[b] Department of Biostatistics, School of Paramedical Science, Shahid Beheshti University of Medical Science, Tehran, Iran
[c] Proteomics Research Center, School of Paramedical Science, Shahid Beheshti University of Medical Science, Tehran, Iran
[d] Research Center for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Science, Tehran, Iran

ARTICLE INFO

Article history:
Received 2 May 2012
Accepted 29 September 2013

Keywords:
GERD
Dyspepsia
Cost analysis
Count models
Zero inflated models

ABSTRACT

Background and study aims: Recent studies have shown that the high prevalence and the various clinical presentations of gastro-oesophageal reflux disease (GERD) and dyspepsia impose an enormous economic burden on society. Economic cost data have unique characteristics: they are counts, and they have zero inflation. Therefore, these data require special models. Poisson regression (PR), negative binomial regression (NB), zero inflated Poisson (ZIP) and zero inflated negative binomial (ZINB) regression are the models used for analysing cost data in this paper.

Patients and methods: In this study, a cross-sectional household survey was distributed to a random sample of individuals between May 2006 and December 2007 in the Tehran province of Iran to determine the prevalence of gastrointestinal symptoms and disorders and their related factors. The cost associated with each item was calculated. PR, NB, ZIP and ZINB models were used to analyse the data. The likelihood ratio test and the Voung test were used to conduct pairwise comparisons of the models. The log likelihood, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) were used to compare the performances of the models.

Results: According to the likelihood ratio test and the Voung test and all three criteria used to compare the performance of the models, ZINB regression was identified as the best model for analysing the cost data. Sex, age, smoking status, BMI, insurance status and education were significant predictors.

Conclusion: Because the NB model demonstrated a better fit than the PR and ZIP models, over-dispersion was clearly only due to unobserved heterogeneity. In contrast, according to the likelihood ratio test, the ZINB model was more appropriate than the ZIP model.

Conclusion: The ZINB model for the cost data was more appropriate than the other models.

© 2013 *Arab Journal of Gastroenterology*. Published by Elsevier B.V. All rights reserved.

Introduction

Gastro-oesophageal reflux disease (GERD) is one of the most common gastrointestinal disorders and is characterised by heartburn and/or acid regurgitation [1]. Recent studies have shown that the high prevalence and variety of clinical presentations of GERD impose an enormous economic burden on society [2,3]. Another common gastrointestinal disorder is dyspepsia, which refers to a group of upper gastrointestinal symptoms. The Rome criteria are international criteria used for diagnosing functional dyspepsia [4]. Although this disorder is not life-threatening, it has a considerable impact on patients and society [5]. Recently, many studies have analysed the economic burden of these symptoms worldwide [2,3,6–9]. These studies have shown that the direct costs of GERD and dyspepsia range from PPP$172 (purchasing power parity dollars) to PPP$176 per person per year. Given the importance of this matter, some studies have recently been performed in Iran to estimate the economic burden.

Rezailashkajani and Moghimi Dehkordi published papers in 2007 and 2011, respectively [3,9]. Moreover, the continuing rise in healthcare expenses worldwide has increased researchers' interest in estimating the precise costs imposed by these diseases and the impact of relevant treatments on the cost of medical care [10]. Therefore, there has been a heightened interest in examining healthcare costs in recent years [11]. Cost data are uniquely distributed and are thus difficult to describe using standard

* Corresponding author. Address: Department of Basic Sciences, School of Rehabilitation, Shahid Beheshti University of Medical Science, Across Bu-Ali Hospital, Damavand Street, Tehran, Iran. Tel.: +98 2177561721; fax: +98 2177561409.
E-mail address: akbarzad@sbmu.ac.ir (A. Akbarzadeh Baghban).

1687-1979/$ - see front matter © 2013 *Arab Journal of Gastroenterology*. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ajg.2013.09.004

approaches such as ordinary least squares regression [12]. The Poisson model is one of the approaches used for analysing cost data. However, due to over-dispersion (a common problem with Poisson regression that frequently arises with count data), other 19 models such as negative binomial regression are also employed for these data [13].

Negative binomial distributions may not satisfactorily account for over-dispersion because of a new problem that emerges in some count data [14,15]. The problem is that the data contain more zeros than other datasets. In these cases, zero inflated models are recommended [14]. Zero inflated binomials, zero inflated negative binomials, and zero inflated Poisson are different types of zero inflated models. These models and comparisons of these models with other count models have recently increased in the medical and healthcare fields [14,16–22]. In this paper, we use Poisson regression (PR), negative binomial (NB), zero inflated Poisson (ZIP) and zero inflated negative binomial (ZINB) models to analyse the costs of GERD and dyspepsia.

Patients and methods

This cross-sectional household survey was conducted between May 2006 and December 2007 in a random sample of individuals in the Tehran province of Iran to determine the prevalence rates of gastrointestinal symptoms and disorders and related factors. A total of 18,180 adults (age >18) were selected (more details in [9,23–29]). Next, trained healthcare employees from local health centres contacted all of the selected individuals (18,180) and asked them to participate in interviews and to answer certain questions. The research protocol of this study was approved by the ethics committee of the Research Center for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Science. The questionnaire, designed specifically for this study, had acceptable levels of validity and reliability and included questions about the symptoms of dyspepsia [23] (based on the Rome III criteria), GERD (i.e., heartburn and regurgitation) and their frequency in the previous 6 months [27]. In addition to information about the listed symptoms, the frequency of health services/resource utilisation (including doctor visits, drug use, laboratory tests and hospitalisations) and productivity loss due to GERD/dyspepsia symptoms were reported. GERD was defined as heartburn or acid regurgitation experienced weekly for the previous 6 months. Dyspepsia was diagnosed, based on the Rome II criteria, as experiencing 1 or more of the following symptoms for 3 months (with the onset of symptoms at least 6 months prior to diagnosis): bothersome postprandial fullness, early satiety, epigastric pain and an epigastric burning sensation. The cost analysis methodology used in this study was similar to the methods employed in other cost analyses in Iran [3,30]. For the cost analysis and estimates, direct and indirect expenses (including physicians' fees, drugs, laboratory tests, hospitalisation, and sick days) due to GERD- and dyspepsia-related symptoms were considered. The economic estimates were calculated as follows: Direct costs = physician's visit fees + drug expenses + laboratory test costs + hospitalisation expenses; Indirect costs = number of days of total productivity loss + number of days with (at least 30%) reduction in functionality; and Total cost = direct costs + indirect costs.

All of the estimated costs were converted to PPP$ to facilitate international comparisons. PPP$ is an economic technique used to determine the relative values of two currencies [9].

Statistical methods

Poisson regression (PR) belongs to the class of generalised linear models (GLMs), and it describes count outcomes or proportions/rates [13]. This model assumes that the responses have a Poisson distribution. Count data usually exhibit greater variability in the response counts than one would expect if the response distribution were truly Poisson. In such circumstances, the variances are much larger than the means, whereas Poisson distributions have equal means and variances. The presence of greater variability in a dataset than expected for a general linear model is called over-dispersion. A common cause of over-dispersion is subject heterogeneity [13]. The negative binomial (NB) model, another statistical model in the GLM class that can be used as an alternative to the PR model, accounts for over-dispersion caused by unobserved heterogeneity [31]. This model adjusts the standard errors of the regression coefficients and provides a more flexible approach for predicting the count outcome. The NB model might not be appropriate if the over-dispersion is caused by an excessive number of zeros in the outcome. In these cases, alternative models such as zero inflated models are recommended [14]. Zero inflated Poisson (ZIP) models combine the Poisson distribution with a degenerate component of point mass at zero [32]. This type of model assumes that the observation at zero has a probability of P and that other observations follow a Poisson distribution with a probability of $1 - P$ [14]. However, if the non-zero observation does not follow the Poisson model, then the ZINB model is used, which considers the count as a negative binomial distribution [31]. There is a possibility that the ZINB model accounts for the over-dispersion caused by both an excessive number of zeros and unobserved heterogeneity [31,33]. The nested models (e.g., PR versus NB and ZIP, NB versus ZINB) were compared using the Voung test, and the other two models (ZIP versus ZINB) were compared using the likelihood ratio test. To compare the performance of the models, various methods were used: the log likelihood, the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). p-Values less than 5% were considered significant.

Results

A total of 1929 eligible patients were enrolled in this study, of whom 1186 (61.4%) had high costs for gastrointestinal disorders. The mean cost per patient was PPP$124.56, and the standard deviation was PPP$399.707. The median cost was PPP$47. The mean age was 46.22 years (standard deviation 16.55), with a range of 16–98 years. The distribution of the covariates in the analysis is shown in Table 1.

A descriptive comparison revealed that age and smoking status were consistently associated with the outcomes across all of the models. In the PR model, all of the covariates were statistically significant; in the ZIP and ZINB models, all covariates except marital status were statistically significant; and in the NB model, age and smoking status were statistically significant. The Pearson's chi-square goodness of fit (gof) test ($p < 0.001$) and the other model fit indices indicated that the PR model was a poor fit for the cost data. In the NB model, the estimated dispersion statistic (α) was

Table 1
Zero inflated negative binomial model for cost data.

Variable	Negative binomial part		Zero inflated part	
	Adj. OR* (95% CI)	p-value	Adj. OR* (95% I)	p-value
Female	1.15 (1.01, 1.36)	0.03	1.19 (0.79, 1.82)	0.38
Age	1.01 (1.00, 1.02)	<0.0001	0.99 (0.97, 1.00)	0.06
BMI	0.99 (0.98, 1.00)	0.45	1.04 (1.00, 1.09)	0.04
Smoking	1.61 (1.30, 1.99)	<0.0001	1.64 (0.97, 2.80)	0.06
Insurance	1.16 (1.00, 1.34)	0.05	2.22 (1.47, 3.38)	<0.0001
Marriage	1.22 (0.65, 2.27)	0.52	1.05 (0.86, 1.28)	0.60
Education	1.08 (0.95, 1.23)	0.23	1.80 (1.13, 2.82)	0.01

* Adjusted odds ratio.

Download English Version:

https://daneshyari.com/en/article/3280811

Download Persian Version:

https://daneshyari.com/article/3280811

Daneshyari.com