Statistical modelling of Health Related Quality of Life measures for Colorectal Cancer with beta-binomial and quantile regression approaches.

Andrés Jiménez García[1], *Dae-Jin Lee*[2], *Urko Aguirre Larracoechea*[3]

[1]anjigar@alumni.uv.es, Master student, University of Valencia
[2]dlee@bcamath.org, Applied Statistics Group, BCAM - Basque Center for Applied Mathematics
[3]urko.aguirrelarracoechea@osakidetza.eus, Research Unit, REDISSEC: Health Services Research on Chronic Patients Network, Hospital Galdakao-Usansolo

Abstract

In this work we proposed a methodology for modelling scores from Health Related Quality of Life indexes for both generic and specific questionnaires of colorectal diagnosed patients. We use a two stages approach: i) fit a Beta-binomial regression model for variable selection, and ii) estimate a quantile regression model with the selected variables.

Keywords: HRQoL, Colorectal cancer, Beta-binomial regression.

1.   Introduction

Cancer is one of the leading causes of morbidity and the first causes of mortality worl-wide, with 14 million new cases and 8.2 million deaths in 2012 [1]. Among them, colorectal cancer (CRC) is the third most common cancer in males and the second most common for females. In Spain, there were detected 32.240 new cases and 14.700 deaths in 2012 [2].

In the last 20 years, a lot of research has focused on measuring the quality of life in health status of patients particularly in chronic diseases. Health related quality of life (HRQoL) is a multidimensional, dynamic and subjective concept that is used to measure the perception of quality of life (in diferent constructs) in patients with chronic diseases. To asses the HRQoL is common the use of patient-reported questionnaires. The questionnaires are classified in two groups: generic and specific.

With this study we pretend to understand better the situation of the patients of one of the most common cancers in Spain and know where to focus the efforts to improve their quality of life as much as possible.

2.   Data description

In this study we analyzed a subsample of patients diagnosed of CRC from a multicentre longitudinal study. We focused on the baseline data of patients from Hospital Galdakao-Usansolo (Biscay, Basque Country), measured at the moment of the diagnosis of the disease ($n = 166$). The participants of the study were those patients chosen for surgery. The data consists of socio-demographic, clinical and social support variables. Patients completed five different questionnaires as we can see in the table 1. Figure 1 enumerates the HRQoL questionnaires and the information available in the database.

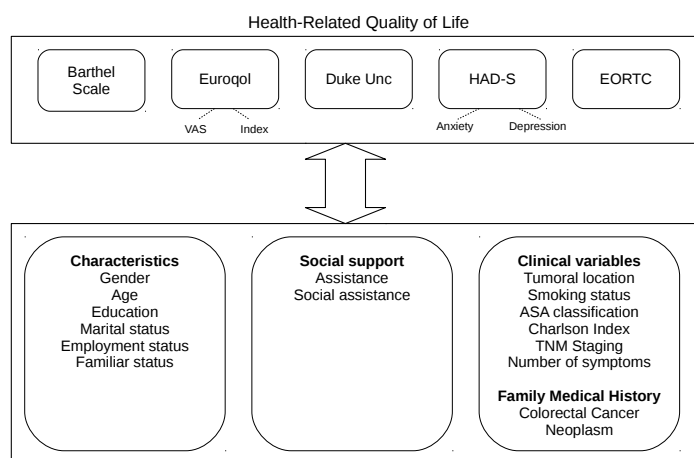| Name of questionnaire	Type	Measures
EQ-5D	Generic	Quality of Life
Barthel score	Generic	Performance in activities of daily living
Duke UNC	Generic	Social support
HADS	Generic	Anxiety and depression
EORTC QLQ-C30	Cancer specific	Quality of life, functional and symptoms scales

Table 1: List of questionnaires

Figure 1: Health status

Due to there were just a few more than a 10% of missing values, we used the multiple imputation to recovery of the missing information. (See [3]).

3.   Regression modelling of HRQoL measures

In this section, we propose a two stages approach. First we used a regression model for variable selection. Due to the characteristics of the scores (discrete responses bounded on a range), we follow the approach by [4, 5, 6] and choose the Beta-binomial regression model. In the abstract only figures some of the results due to the limitation of the space. Figure 2 shows the variables selected from the generic (EQ-5D) and specific questionnaires (EORTC: Emotional functioning, Quality of Life, fatigue, nausea and vomiting and pain).

A second step is to use the variables selected in the previous step to estimate a quantile regression model. Cavrini [7] proposed a quantile regression model for EQ-5D questionnaire but does not compare with a specific questionnaires. The quantile regression does not depend on the nature of the variable (i.e. no distributional assumptions). Using this approach we are able to

compare both generic and specific questionnaires for CRC patients in our study considering scores conditioned to predetermined percentils.

	Gr	ASA	Loc	CRC	Neo	TNM	Ed	As	SA	Ch	Red	Mel	AWL	Ab	Cn
EQ-5D															
Index				√				√						√	
VAS	√		√			√	√	√		√			√	√	
EORTC															
Emotional				√				√			√		√	√	
QoL								√					√	√	
Fatigue	√	√	√					√		√		√	√	√	√
Nausea							√	√	√					√	√
Pain	√								√			√		√	

Gr: Gender, ASA: ASA classification, Loc: Tumoral location, CRC: CR Cancer in the FMH, Neo: Neoplasm in the FMH, TNM: TNM staging, Ed: Education, As: Assistance, SA: Social assistance, Ch: Charlson index, Red: Red blood in the stool, Mel: Melaena, AWL: Apettite or weight loss, Ab: Abdominal pain, Cn: Constipation

Table 2: Significative explanatory variables

EQ-5D Index																	
Neoplasm FMH																	
Assistance																	
Abdominal pain	√	√	√	√	√												
EORTC: Emotional																	
CR cancer FMH				√	√	√	√	√	√	√	√	√					
Assistance	√	√	√	√	√	√	√	√	√	√	√						
Red blood in stool										√	√	√	√	√	√	√	√
Apettite/Weight loss	√	√	√	√	√	√											
Abdominal pain	√	√	√	√	√	√	√	√	√	√	√	√	√				
EORTC: QoL																	
Assistance	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	
Apettite/Weight loss	√																
Abdominal pain	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√	√
Percentiles	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90

Table 3: Significance of explanatory variables by percentiles for *EQ-5D Index, EORTC:Emotional functioning and EORTC:QoL.*

4. Conclusion

In this paper, we proposed a statistical methodology to model scores from HRQoL questionnaires of CRC diagnosed patients. We proposed a two-stage approach based on a Beta-binomial regression model for the response variable and the significant variables of the study and then a quantile regression fit. This approach contributes to determine the most important factors of interest in the study of patients

It should be pointed out that abdominal pain is an important symptom as we can see in Table 2. In general, when a patient suffers of abdominal pain, he shows a HRQoL perception lower. Another important factor is the assistance. In our study, we distinguish between two kinds of assistance: normal assistance and social assistance. Usually when a patient needs assistance it is because is in a worse status than other who does not need assistance. We conclude here that the assistance is focusing in the patients with worse HRQoL, so the resources are making the most.

Gender is significant for fatigue and pain. In [8] they that obtained women show differences in physical problems and pains. Women have more perception of fatigue and pain than men.

Table 3 shows the significance of the explanatory variables of the fitted quantile regression models by percentiles for some of the scores (the rest of the scores were ommitted due to the limited space). For instance, for abdominal pain we showed that it was significant in Table 2 for all the scores shown. However, using a quantile regression model we found that for those patients with low EQ-5D index score ($<$ percentil 30) abdominal pain is significant. For EORTC:QoL, abdominal pain is significant for all the percentiles, but it scope is specific to cancer patients.

5.   Acknowledgments

6.   Bibliography

[1] Ferlay, J., et al. (2013). GLOBOCAN 2012 v1. 0. *Cancer incidence and mortality worldwide: IARC CancerBase 11.*

[2] Sociedad Española de Oncología Médica (2014). *Las Cifras del Cáncer en España.*

[3] Stef van Buuren, Karin Groothuis-Oudshoorn (2011). *mice: Multivariate Imputation by Chained Equations in R.* Journal of Statistical Software, 45(3), 1-67. URL http://www.jstatsoft.org/v45/i03/.

[4] Arostegui, I., Núñez-Antón, V., Quintana, J. M. (2007). *Analysis of the short form-36 (SF-36): the beta-binomial distribution approach.* Statistics in medicine, 26(6), 1318-1342.

[5] Arostegui, I., Núñez-Antón, V., Quintana, J. M. (2012). *Statistical approaches to analyse patient-reported outcomes as response variables: An application to health-related quality of life.* Statistical methods in medical research, 21(2), 189-214.

[6] Khan, I., Morris, S. (2014). *A non-linear beta-binomial regression model for mapping EORTC QLQ-C30 to the EQ-5D-3L in lung cancer patients: a comparison with existing approaches.* Health and quality of life outcomes, 12(1), 163.

[7] Cavrini, G. (2010). *A quantile regression approach for modelling a health-related quality of life measure.* Statistica, 70(3), 273-291.

[8] Krouse, R. S., Herrinton, L. J., Grant, M., Wendel, C. S., Green, S. B., Mohler, M. J., Hornbrook, M. C. (2009). *Health-related quality of life among long-term rectal cancer survivors with an ostomy: manifestations by sex.* Journal of Clinical Oncology, 27(28), 4664-4670.