Available online at www.elixirpublishers.com (Elixir International Journal)

Statistics

Elixir Statistics 97 (2016) 42142-42145

ISSN: 2229-712X

Ordinal logistic model for finding the risk factors of HIV testing in injecting drug users

Rakesh Kumar Saroj[1,*], Dr.K.H.H.V.S.S.Narasimha Murthy[1] and Avadhesh Kumar[2]

[1]Department of Kayachikitsa,IMS- Banaras Hindu University, Varanasi, India.

[2]Department of Community Medicine,IMS- Banaras Hindu University, Varanasi, India

ARTICLE INFO

Article history:

Received: 3 May 2016;

Received in revised form:

5 August 2016;

Accepted: 10 August 2016;

Keywords

Ordinal logistic regression,
HIV testing,
proportional odds model,
IDUs.

ABSTRACT

The ordinal regression is a method that is used to robust the model when dependent variable is ordinal and Independent variables may be dichotomous, polytomous, and continuous or combination of these. Ordinal logistic regression is used to predict the "odds" of having a lower or a higher value for dependent variable (y), based on independent variable (x). In practice, the frequently used type of model is a proportional odds model in ordinal logistic regression. HIV testing is necessary for preventing and reducing the HIV transmission. However, there are various Socio-demographic and HIV related behavior factors contribute the high or low HIV testing in general population and high risk groups. Intend of this study find out the important factors of the HIV testing in Injecting drug users (IDUs) patients. The ordinal logistic regression model makes assumptions about the nature of the relationship between the order response variable HIV testing Methods: Total 139 IDUs patients' collect the information for this research based on specific questioner from the district Kamur in Bihar. In study, Ordinal logistic regression analysis to determine the factors which are considered to be a significant contributor in HIV testing. The ordinal logistic regression model was used to build models for dependent variable HIV testing and independent variables which are Age, Marital Status, Education, Occupation, Stigma, Income, STI/STD problems, Needle injecting sharing and HIV information. Results: In this research apply the proportional odds model for confirm the applicability of the ordinal logistic model. We determine the all parameter the significant of the model. We found that Needle sharing, Abscess problem, Abuse, Heard about STI, HIV, Income, HIV knowledge, HIV transmission through multiple partners shows significant contribution to IDUs patient for HIV testing. Conclusion: This study has made an attempt to recognize the predictors of HIV testing for injecting drug users by developing an ordinal logistic regression model.

Introduction

In 2013, HIV prevalence in India was an estimated 0.3%. Overall, India's HIV epidemic is slowing down, with a 19% decline in new HIV infections (130,000 in 2013), and a 38% decline in AIDS-related deaths between 2005 and 2013. Despite, this 51% of deaths in Asia are in India. The five states with the highest HIV prevalence (Nagaland, Mizoram, Manipur, Andhra Pradesh and Karnataka) are in the south or east of the country. Some states in the north and northeast of the country, report rising HIV prevalence.HIV prevalence among injecting drug users (IDUs) in India has remained largely unchanged since 2007, at around 7% of the 177,000 people in this population. 30% of IDUs reside in north-eastern states where injecting drug use is the major route of HIV transmission. However, HIV prevention efforts in this region have been effective in reducing the number of new infections. This is despite national HIV prevention activities coverage of 80.7% and 150,000 needles and syringes distributed per IDUs per year. Research has emphasized the need for early interventions for IDUs in India. Indeed, many embark on a 'drug career' in their early teens using widely available substances such as tobacco and alcohol before progressing on to illegal drugs through a non-injecting route (e.g. orally or smoking) and eventually using shared needles and syringes putting them at risk of HIV transmission. It is only at this point that IDUs are typically reached by harm reduction services.

Ordinal Logistic Regression (OLR)

Ordinal logistic regression always use if our outcome would be in order way. This model relies on the cumulative logit, in this model the predicted probability of multiple outcomes. This is also known as the "proportional odds model". In ordinal logistic regression, the event of interest is observing a particular score or less. For the rating of judges, model the following odds:

$$\theta 1 = prob(score\ of\ 1)/prob(score\ greater\ than\ 1)$$

$$\theta 2 = prob(score\ of\ 1\ or\ 2)/prob(score\ greater\ than\ 2)$$

$$\theta 1 = prob(score\ of\ 1{,}2\ or\ 3)/prob(score\ greater\ than\ 3)$$

The last category doesn't have an odds associated with it since the probability of scoring up to and including the last score is 1. All of the odds are of the form:

$$\theta j = prob\ (score \leq j)/\ prob\ (score > j).$$

We can also write the equation as θj = prob (score $\leq$ j) / (1 – prob (score $\leq$ j), since the probability of a score greater than j

Tele: 9454196475

E-mail address: rakeshsaro@gmail.com

is 1 – probability of a score less than or equal to j. Having n observations with ordinal response variable, Y, an ordinal logistic regression model relates the probability of an event occurs to predictor variables x'= (x_1, x_2... x_k) can be written as cumulative logits which is defined as follows:

$$logit\left[\frac{P(Y \le j)}{x}\right] = log\ \frac{P(Y \le j/x)}{1 - P(Y \le j/x)}$$

$$= \frac{log\ \pi1(x) + \pi2(x) + \pi3(x) + \cdots \pi j(x)}{\pi j + 1(x) + \pi j + 2(x) + \cdots \pi J(x)}$$

$$= \alpha j + \beta' x,\ \ j = 1,2, \ldots \ldots, j - 1$$

Error! Reference source not found.Error! Reference source not found. In that model, each logit has its own which is called as the entry value and their values do not depend on the values of the independent variable for a particular case. Also, each cumulative logit uses all j response categories. The cumulative logit has its own intercept, but the model has the same effects β for each logit. The expression is called as a proportional odds model and satisfies the following expression:

$$logit\,[P(Y \le j/x1)] - logit\,[P(Y \le j/x2)]$$

$$= log\ \frac{P(Y \le j/x1)/P(Y > j/x2)}{P(Y \le j/x2)/P(Y > j/x1)}$$

In many epidemiological and medical studies, OLR model is frequently used when the response variable is ordinal in nature. In various health/medical studies, ordinal logistic regression to be used but without precise knowledge of ordinal logistic regression interpretation mostly researcher using binary logistic regression. There are several ordinal logistic regression models such as proportional odds model (POM), two versions of the partial proportional odds model-without restrictions (PPOM-UR) and with restrictions (PPOM-R), continuous ratio model (CRM), and stereotype model (SM).The most frequently used ordinal logistic regression model in practice is the controlled cumulative logit model called the proportional odds model. When looking specifically at HIV testing behaviors among IDUs in India, a very rare number of studies were identified. One study discovered that Sexual behavior among men who have sex with women, men, and Hijras in Mumbai. However, those engaging in unprotected sex did not perceive themselves to be at risk of contracting HIV. Another study found to determine and compare the HIV/STI risk perceptions and behavior among men and women with traditionally high-risk and low-risk behavior in Chennai, India. Another study focused The MSWs have high-risk behaviors, low consistent condom use, and high STI/HIV infections.

In an effort to identify studies that specifically IDUs HIV testing status and related attitudes behaviors in the Bihar none one published report was found. Therefore, in an attempt to address the apparent gap in HIV testing status among in IDUs. In this study ordinal logistic regression identifies the predictors of HIV testing status in injecting drug users. The study applies the proportional odds model for the analysis of the HIV testing and it's contributing factors. The aim of this study is to explain the ordinal logistic regression model terminology and identifying the significant predicators affecting the HIV testing status in injecting drug users.

Objective of Research

The overall purpose of the research will be to have better understanding ordinal logistic regression approach for of influencing the socioeconomic, demographic factors, sexual behavior and other relevant covariates in HIV testing of IDUs users.

Material and Methods

The study has utilized the data of need assessment of IDUs patients of year 2013 from Kamur district of Bihar. Total 139 IDUs collected the data using questionnaire where consent was obtained to collect the data. The data consist of several predictors such as Age, Sex, Marital Status, Education, Occupation, Stigma, Income, Residence; Drug used status, History of sex, STI/STD problems, Needle injecting sharing, and HIV information. The original response variable for this study is HIV testing. Several socio-economic and demographic characteristics which are above mention considered were the independent variables to develop the ordinal logististc regression models. In this research HIV testing status was yes, no and don't know three ordinal categories. For the data analysis and model fitting we use the SPSS 20.

Assumption

Before going for ordinal logistic regression approach should follow the below given assumption

1. Dependent variable should be always in ordering way
2. The independent variable should be continuous, ordinal or categorical (including dichotomous variables). However, ordinal independent variables must be treated as being either continuous or categorical.
3. There is no multicollinearity. Multicollinearity occurs when you have two or more independent variables that are highly correlated with each other. This leads to problems with understanding which variable contributes to the explanation of the dependent variable.
4. Proportional odds are a fundamental assumption of the ordinal regression model. The assumption of proportional odds means that each independent variable has an identical effect at each cumulative split of the ordinal dependent variable.

Model Fitting

There are several approaches to be used in ordinal logistic regression namely partial proportional odds model (POM), proportional odds model-without restrictions (PPOM-UR), with restrictions (PPOM-R), continuous ratio model (CRM and stereotype model (SM).Then first we check the model fitting of the ordinal logistic regression, means which type of model we could apply on this data. Test of Parallelism one of the technique to check the proportional odds model (POM) is fit in this data or model assumptions of ordinal regression is that regression coefficients are the same for all categories. If the test of parallel lines or proportional odds assumption is not significant in our data then indicates that proportional odds model (POM) is fit in this data or model assumptions of ordinal regression is that regression coefficients are the same for all categories.

Odds ratio for ordinal data

The odds ratio for every response is taken to be constant across all possible failing of the outcome variable. If the testable assumption is met, the odds ratios in a cumulative logit model are interpreted as the odds of being "higher" or "lower" on the outcome variable, across the whole series of the outcome. The appropriate and spontaneous interpretation of the cumulative logit model are two main reasons for its being measured the most popular model for ordinal logistic regression.

Suppose the target response (Y) on HIV testing has n ordered categories (Yk with j = 1, 2... n) and that two groups (A and B) need to be compared. For the category k, OR is given by:

OR is the ratio between two odds, but now odds is defined in terms of cumulative probabilities. For its interpretation, be sufficient it to recall that the response has been dichotomized, and that the event is to be classified until the category k. If A and B represent, respectively, exposure and non-exposure to a factor, OR quantifies the odds of an individual in the exposed group being classified up to a given category, compared to the odds of the unexposed group. In the context of ordinal data, according to the proportional odds assumption, OR is the same for all categories of the response variable.

Result

The test of parallel lines or proportional odds assumption is found insignificant at 5% level of significance indicating the data satisfy the proportional odds assumption. However, the p-value of the score test is found small (0.774). To confirm the conclusion regarding the assumption of POM. Hence, proportional odds assumption is satisfied for HIV testing status data (Table1). Model-fitting information can ensure that the difference between the two log-likelihoods—the chi square—has an observed significance level of less than 0.05.This means that we can reject the null hypothesis that the model without predictors is as good as the model with the predictors (Table2).Though the proportional odds assumption is suitable for analyzing the ordinal data and its covariates. The Pseudo R-square (Cox and Snell=0.26, Nagelkerke =0.293, and McFadden =0.14) indicates that data perform quite well to identify the factors of HIV testing, the study fitted POM and results of the models are interpreted. The first ordinal logistic regression equation (A) is written as follows using the coefficient of constant, which is -2.48.29 and coefficients value of various response variables:

= -2.48+ [-1.62*Age (31-40 yrs) -1.68* Education (HSC/SSC) -3.02* Income (<100 Rs)-2.83* Abscess Problem +3.01* STI +1.5* STD Problem (No) -3.3* what is HIV -3.8* HIV spread through needle share +3.34* Stigma -1.8* Service provider poor]. ------------------------------- (A)

The second ordinal logistic regression equation (B) is written as follows using the coefficient of constant, which is 0.706 and coefficients value of various response variables:

= 0.706+ [-1.62*Age (31-40 yrs) -1.68* Education (HSC/SSC) -3.02* Income (<100 Rs)-2.83* Abscess Problem +3.01* STI +1.5* STD Problem (No) -3.3* what is HIV -3.8* HIV spread through needle share +3.34* Stigma -1.8* Service provider poor] --------------------------------- (B)

The other result to discuss in this research related to ordinal logistic regression is odd ratio. The Table 3 shows the coefficients for the independent variables. All these coefficients are the log relative risks. The coefficient for age is -1.62, which means that we have a negative association between age and HIV testing status. The indicates that the every one-unit increase in age, the HIV testing decreases. The Odds Ratio (31-40 years = 0.20) means testing status of age group 31-40 yrs 0.20 more compare to other age group. The coefficient for education is -1.68, which means that have a negative association between education and HIV testing status. In other words, for every one-unit increase in education the HIV testing decrease. The odds ratio (education hsc/ssc = 0.19) means testing status of education who have hsc/ssc is 0.19 is more compare to other educated group. The coefficient for income is -3.02, this is also says that have negative association between income and HIV testing status, this is showing that the every one-unit increase in income, the HIV testing decreases.

The odds ratio (income <100 Rs =0.05) indicates that whose income is < 100 RS/monthly is 0.05 more testing is compare to other income group. We have found negative association between abscess problem and HIV testing status because the coefficient of abscess problem is -2.83 and odds ratio is .05.The coefficient of STI is 3.01 means that have positive association between STI and HIV testing status this is showing that every one unit increase in STI, the HIV testing increase and its odds ratio is 20 means those having STI is 20 times more testing compare to non STI patients. Needle sharing is also one of the important factors contributing in the HIV testing. Abuse finds negative association between and HIV testing. The coefficient of heard about STI/HIV is 0.63 and 0.40 respectively means that have positive association between heard about STI/HIV and HIV testing status this is showing that every one unit increase in heard about STI/HIV, the HIV testing increase. The odds for non STD problem to STD problem to improve HIV testing are 4.48. The coefficient of what is HIV,HIV transmission through multiple partners/needle share and service provider is -3.3,-0.37,-3.8,-1.8 respectively which means all are the factors having negative association with HIV testing. The stigma coefficient is 3.34 which shows that have positive association between stigma and HIV testing status this is showing that every one unit increase in stigma, the HIV testing increase and its odds ratio is 28 means those having non-stigma patient is 20 times more testing compare to stigma patients.

Conclusion

At view the proportional odds model seems to be an suitable model for analyzing the considered data since the p-value of chi-squared score test for overall model is insignificant at 5% level of significance indicating proportional odds assumption is not violated. Total calculated variables were found significant in the proportional odds model. Before the most popular among ordinal regression models – the proportional odds model is applied make sure the proportional odds assumption is satisfied. Otherwise, the results cannot be convincing. If the hypothesis is not satisfied the generalized logit models (e.g. partial proportional odds model) should be developed. In this study, the various predictor variables were used to analyze the data using ordinal logistic regression model which are age, income, sex, residence, marital status, education, occupation, stigma, abscess problem, heard about HIV/STD, route of transmission income, STI/STD problems, needle injecting sharing, HIV information and service provider etc. We estimated the parameter and test the significant of the model. We tested the full model which all predicted variables are tested using the same model. As a conclusion, we obtain that the age, education (HSC/SSC), income (<100 Rs), abscess problem, what is HIV,STD problem, STI, Needle sharing, abuse, heard about STI, HIV, HIV knowledge, HIV transmission through multiple, HIV spread through needle share, stigma, service provider are having significant contribution to HIV testing.

Table1.Test of parallel line

Model	-2 log like lihood	Chi-square	Df	sig
Null Hypothesis	175.853			
General	132.684[b]	43.169[c]	51	0.774

Table2. Model Fitting Information (Fitness of good)

Model	-2 log like lihood	Chi-square	Df	sig
Intercept only	289.125			
Final	175.853	113.272	51	.000

Pseudo R-square

Cox and snell	.557
Nagelkerke	.637
Mcfadden	.392

Table 3.Result of the Multiple POM using nutrition status as response three ordered categories.

Covariate	Logistic coefficient (β)	Standard error	P-Value	Odds ratio	95%CI
Iintercept1 (HIV testing 1)	-2.48	2.64			
Iintercept2 (HIV testing 2)	0.706	2.63			
Age (31-40yrs)	-1.62	0.06	0.018	0.20	0.18-0.22
Education (HSC/SSC)	-1.68	0.08	0.042	0.19	0.16-0.22
Income(<100Rs)	-3.02	0.08	0.005	0.05	0.04-0.06
Abscess problem	-2.83	0.07	0.008	0.06	0.05-0.07
STI	3.01	0.13	0.02	20	16-26
Needle sharing	0.23	0.05	0.001	1.26	1.16-1.36
Abuse	-0.19	0.004	0.000	0.82	0.81-0.83
Heard about STI	0.63	0.29	.003	1.9	1.3-2.5
Heard about HIV	0.40	0.70	0.00	1.5	1.35-1.63
STD problem (No)	1.5	0.062	0.017	4.48	3.97-5.06
What is HIV	-3.3	0.138	0.017	0.04	0.03-0.05
HIV transmission through multiple partners	-0.37	0.20	0.00	0.69	0.83-0.55
HIV spread through needle share	-3.8	0.129	0.003	0.02	0.02-0.03
Stigma	3.34	0.103	0.001	28	23-35
Service privider	-1.8	0.081	0.026	0.17	0.14-0.19

References

1.a.b. c. d. e. f. g. h. i. j. UNAIDS (2014) 'The Gap Report'
2.a.b. c. d. e. f. g. h. i. j. k. l. m. n. o. p. q. r. s. t. u. v. w. NACO (2014) 'Annual Report 2013-14.
3.Adhikary, R. et al (2013) 'Spatial distribution and characteristics of injecting drug users (IDU) in five Northeastern states of India' BMC Public Health 11(64)
4. a. b. NACO (2013) 'India: Annual Report 2012-13
5. NACO and Indian HIV/AIDS Alliance (2014) 'Drug Use Patterns among Clients Receiving Services from Targeted Interventions for People Who Inject Drugs'
6. UNODC (2012) 'Association of Drug Use Pattern with Vulnerability and Service Uptake among Injecting Drug Users'
7. Armstrong BG, Sloan M: Ordinal regression models for epidemiologic data. Am J Epidemiol 1989, 129:191-204.
8. Gemeroff MJ: Using the proportional odds model for health-related outcomes: Why, when, and how with various SAS® procedures. Proceedings of the Thirtieth Annual SAS Users Group International Conference: April 10-13, 2005 2005, Paper # 205-30.
9. Abreu M. N. S., Siqueira A. L., Caiaffa W. T., Ordinal logistic regression in epidemiological studies, Rev Saude Publica, 43 (1), 2009.
10. Ananth C. V., Kleinbaum D. G., Regression models for ordinal responses: a review of methods and applications, Int J Epidemiol, 26 (6), pp. 1323–1333, 1997.
11. Bender R., Grouven U., Ordinal logistic regression in medical research, Journal of the Royal College of Physcians of London, 31 (5), pp. 546–551, 1997.
12. Hosmer D. W., Lemeshow S., Applied logistic regression, 2. Ed., New York, John Wiley & Sons, 2000
13. Agresti A., An introduction to Categorical Data Analysis, Wiley, New York
14. McCullagh P: Regression models for ordinal data. J R Stat Soc B 1980, 42:109-142.
15. Anderson JA: Regression and ordered categorical variables. J R Stat Soc B 1984, 46:1-30.
16. McCullagh P, Nelder JA: Generalized Linear Models New York: Chapman and Hall; 1989.
17. Brant R: Assessing proportionality in the proportional odds model for ordinal logistic regression. Biometrics 1990, 46:1171-8.
18. Lee J: Cumulative logit modeling for ordinal response variables: Application in biochemical research. Compt Appl Biosci 1992, 8:555-562.
19. Ananth CV, Kleinbaum DG: Regression models for ordinal responses: A review of methods and applications. Int J Epidemiol 1997, 26:1323-33.
20. Bender R, Grouven U: Ordinal logistic regression in medical research. JR Coll Physicians Lond 1997, 31:546-51.
21. Hendrickx J: Special restrictions in multinomial logistic regression. Stata Technical Bulletin 2000, 56:18-26.
22. Walters SJ, Campbell MJ, Lall R: Design and analysis of trials with quality of life as an outcome: A practical guide. J Biopharm Stat 2001, 11:155-76.
23. Lall R, Campbell MJ, Walters SJ, Morgan K: A review of ordinal regression models applied on health-related quality of life assessments. Stat Methods Med Res 2002, 11:49-67.
24. Hosmer DW, Lemeshow S: Applied Logistic Regression. 2 edition. New York: John Wiley and Sons; 2000.
25. Agresti A: An Introduction to Categorical Analysis New York: John Wiley and Sons Inc; 1996.
26. Pongsapukdee V, Sukgumphaphan S: Goodness of fit of cumulative logit models for ordinal response categories and nominal explanatory variables with two-factor interaction. Silpakorn U Science & Tech J 2007, 1(2):29-38.
27. Sexual behavior among men who have sex with women, men, and Hijras in Mumbai, India--multiple sexual risks AIDS Behave. 2006 Jul;10(4 Suppl):S5-16
28. HIV risk behavior and acceptability of microbicides in Chennai, India, Eastern Journal of Medicine 16 (2011) 160-167.
29. Male sex workers: are we ignoring a risk group in Mumbai, India Indian 2009 Jan-Feb; 75(1):41-51
30. Whitehead J. Sample size calculations for ordered categorical data. Stat Med 1993; 12:2257-71.