

# RISK REGRESSION ANALYSIS FOR OPENCAST MINE DUMPER OPERATORS

M. K. Singh<sup>1</sup>, U. K. Dey<sup>2</sup>

<sup>1</sup>Collieries Division, Steel Authority of India Ltd, Dhanbad, Jharkhand, India

<sup>2</sup>Dept. of Mining Engineering, Birsa Institute of Technology, Sindri, Dhanbad, Jharkhand, India

## Abstract

This regression analysis studies of dumper drivers were conducted for an opencast mine located in one of mine of Jharia Coal field, Jharkhand, India. Statistical tools have been developed in the light of analytical principles to see whether there is any role of some personal and impersonal factors in the occurrence of coal miners' injuries. With the progress in exploitation of minerals, safety of persons employed became a matter of concern. Occupational injuries in mines are attributed to many factors. Raw data on personal and organizational factors based on the questionnaire (samples were collected from 113 dumper drivers) and other methods were collected from the mine and thereafter they were analyzed. A logistic regression model is used in this research work. With this model both the prediction of group membership and the analysis were done in the form of an odds ratio. On the basis of binary logistic regression for both injury vs no injury and reportable vs no injury cases, then on the basis of ordinal logistic regression results the factors which were found to be most important among the initial factors were selected. The variables those show distinct relationship and considered to be significant with the degree of injury are the personal factors, environmental condition, machinery condition, job satisfaction and risk taking behavior and their respective adjusted odds ratios were found 12.17, 5.26, 6.35, 5.04 and 9.85 for significant risk factors in relation to injury vs No injury case. Similarly, 9.54, 8.40, 4.86 and 9.28 respectively for reportable vs no injury cases. In other words the above mentioned high risky significant factors are 12.17, 5.26, 6.35, 5.04 and 9.85 times more susceptible to injury in comparison to reference category. Out of the above 5 factors most highly significant categories are personal factor and risk taking behavior.

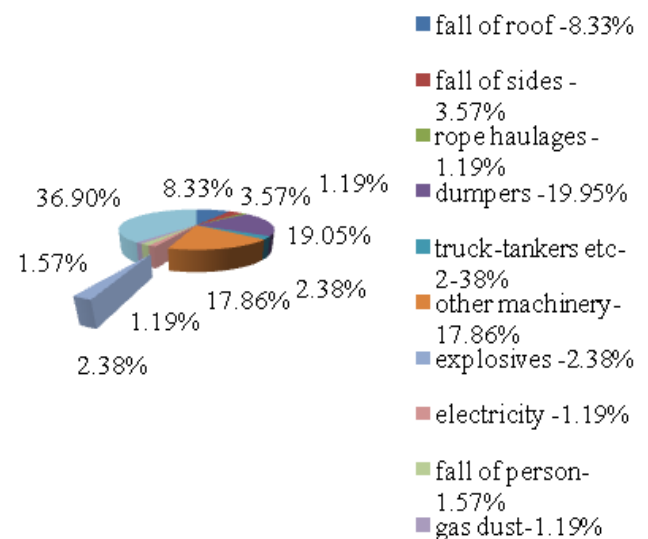
**Key Words:** Risk analysis, Statistical tools, Coal, Occupational Injury, Variables, Personal factors, Impersonal factors, Environment, logistic regression.

\*\*\*

## 1. INTRODUCTION

The majority of the accidents of the opencast mines are caused by the HEMM. 70% to 75 % of the fatal accidents are caused by the transportation machinery. Most of the accidents due to dumpers are caused by the non – compliance of the safety matter and human error. The major accidents by dumpers occurred during reversing operation. However negligence and careless driving, loss of control, defective safety appliances including brakes, unauthorized driving, non-provision of berms, parking brakes, hitting by projectiles due to blasting, etc are the other reasons too. The accidents by other trackless machine (trucks, tankers etc.) are due to the careless and negligence driving, careless reversal operation, unauthorized riding etc. Accidents by the other HEMM are mainly due to poor maintenance, careless and negligence operation, non provision of safety appliances etc. As per DGMS (Tech.) Circulars issued time to time it has been revealed that, root causes of accidents are lack of awareness, lack of adequate training, issues regarding maintenance of dumpers, tippers/trucks, design and maintenance of haul roads [1]. This research work highlights in depth on the occurrences of accidents due to behavior of dumper drivers, organizational factors and environmental conditions. Bar chart below shows the cause wise distribution of fatal accidents in coal- mines in year 2014 [2]. With the help of this Bar chart, it can be brought to

notice that, dumpers (19.95 %) causing more fatal injuries than other causes, independently.



**Chart-1:** Cause wise distribution of fatal accidents in Indian coal mines - Year 2014

**2. METHODOLOGY**

The logistic regression model is used in this exercise [3]. The relationship between the predictor and response variables is not a linear function in logistic regression; instead, the logistic regression function is used which is the logit transformation of q:

$$\theta = \frac{e^{(\alpha + \beta_1 r_1 + \beta_2 r_2 + \dots + \beta_n r_n)}}{1 + e^{(\alpha + \beta_1 r_1 + \beta_2 r_2 + \dots + \beta_n r_n)}} \quad \text{--- (1)}$$

Where  $\alpha$  = constant of the equation and,  $\beta$  = coefficient of the predictor variables.

An alternative form of the logistic regression equation is:

$$\begin{aligned} \text{logit}(p) &= \ln \left[ \frac{p}{1-p} \right] \\ &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \dots + \beta_i x_i. \quad \text{--- (2)} \end{aligned}$$

Where  $\beta_0$  corresponds to  $\alpha$ .

Thus, if  $x_{age}$ ,  $x_{exp}$ ,  $x_{pf}$ ,  $x_{env}$ ,  $x_{mc}$ ,  $x_{hr}$ ,  $x_{js}$ , and  $x_{rt}$  represents the risk factors age, experience, personal factors, environmental condition, machinery condition, haul road condition, job satisfaction and risk taking behavior (independent variables) and  $y$  is a binomial-outcome variable with probability of injury =  $p$ , the multivariate logistic-regression model is given by

$$\begin{aligned} \text{logit}(p) &= \ln \left[ \frac{p}{1-p} \right] \\ &= \beta_0 + \beta_{age} x_{age} + \beta_{exp} x_{exp} + \dots + \beta_{hr} x_{hr} + \dots + \beta_{rt} x_{rt} \dots \quad \text{(3)} \end{aligned}$$

□

$\left[ \frac{p}{1-p} \right]$  is called the logistic transformation and it is used as the dependent variable. The term is known as the odds of risk. There are two main uses of logistic regression. The first is the prediction of group membership. The second is the analysis in the form of an odds ratio since logistic regression calculates the probability or success over the probability of failure [4].

**3. A CASE STUDY**

The opencast mine is located in the Jharia coal field of Jharkhand, India. The mine is producing on an average of 6 lac te Tonne of coal per year. The main equipment consists of shovel –dumper combination for production of coal. However dumpers are also being used for overburden removal. A total of 20 dumpers are being employed in the mines having capacity of 35Tonne. 35 Tonne dumpers are mainly being used for overburden material transport. There are four quarriable seams occurring in this block which are numbered from top to bottom as 14, 13, 13/14 combined, 12, seams. Total mineable reserves are 40 million Tonne provided that 120.64 million m<sup>3</sup> of overburden is removed.

The maximum and average gradient of coal seams are 1in 1.2 and 1in 1.4 respectively.

**4. ANALYSIS OF RAW DATA**

Raw data on personal and organizational factors based on the questioners (samples were collected from 113 dumper drivers) and other methods were collected from the mine and thereafter they were analyzed .The summary statistics of the above mentioned factors are presented in Table-1. Mean, median, mode, standard deviation, minimum and maximum values of individual factors are derived for the said case study mine. The mean, median and mode values of age are 42.42, 42.00, and 40.00 respectively whereas for experience these values are 15.74, 15.00 and 14.00 respectively. The mean, median and mode of personal factor are 9.65, 9.00 and 9.00 respectively. Similarly for other factors too these statistical values are found out which depicted in the Table- 1. Based on the frequency distribution, categorization has been done which is depicted in the Table-2. It shows the description of Total Injuries, Serious and Reportable injuries with respect to various factors.

**5.VARIOUS TESTS FOR SHOWING RELATIONS [5, 6]**

**5.1 Chi Square Test**

The chi-square statistics was used to test the hypothesis showing that the row and column variables are independent. The rows consist of the variable categories and the columns consist of injury group and no injury group. This test was done with the categorized variables and the significance of each factor was observed. The chi-square test reveals that all factors are significantly influencing injury w.r.t. injury (serious, reportable and minor) vs no injury case. Statistical significance was determined at  $p < 0.05$ . The SPSS package was used for the analysis of the raw scores.

**5.2 Odds Ratio**

The odds ratio provides a measure of the relative risk of injury (i.e. relative risk to be in the case group) of a category of a variable compared to the reference category. Of the various epidemiologic measures of association, the odds ratio (OR) seems to strike people as the most “mysterious”. Table-4 presents the crude odds ratios (ORs). Significant ORs were found for all factors for the injury vs no injury case. The study, results based on odds ratios, revealed that the older age drivers, drivers having less experience, those who are on the high risky side of personal factors, environmental conditions, machinery conditions, haul road conditions, job satisfaction and risk taking behavior played marked role in driver’s injuries.

**6. VARIABLES CODING**

The dependent variable has been coded as ‘0’ for no injury and ‘1’ for injury in both the reportable or serious injury cases. The independent variables are coded as dummy variables. These variables include age, experience, personal

factors, environmental condition, machinery condition, haul road condition, job satisfaction, and risk taking behavior. A description of the categories with indicator coding scheme has been shown in Table 4.

**7. GOODNESS -OF-FIT TEST (INJURY VS NO INJURY)**

The goodness-of-fit statistics for the logistic model run for injury vs no injury case is shown in the next paragraph. For the model with all of the independent variables the value of -2LL is 57.033, which is smaller than the -2LL for the model containing only the constant, 128.710. Here the model chi-square is the difference between -2LL for the model with only a constant and -2LL for the current model (128.710 - 57.033 = 71.677).

**Goodness-Of-Fit Statistics For Model Containing Only The Constant**

Dependent Variable: injury  
Beginning Block Number 0. Initial Log Likelihood Function

-2 Log Likelihood 128.710 Constant is included in the model. No terms in the model

Thus, the model chi-square tests the null hypothesis that the coefficients for all the terms in the current model, except the constant, are zero. The small observed significance level (.000) for the model chi-square value indicates that the null hypothesis can be rejected. Hence it can be inferred that the proposed model is a good fit to the data set. Here Hosmer and Lemeshow test is showing no difference between observed and expected value.

**Goodness-of-fit statistics for model containing the independent variables**

-2 Log Likelihood (-2LL)	57.033			
Chi-Square		df	Significance	
Model Chi-Square		71.677	10	.000
Improvement		71.677	10	.000
<b>Hosmer and Lemeshow test</b>				
Chi-square		df	Significance	
4.094		8	.849	

**8. ADJUSTED ODDS RATIO – INJURY VS NO INJURY**

In a multiple logistic regression model where the response variable is the presence or absence of a disease, an odds ratio for a binomial exposure variable is an adjusted odds ratio for the levels of all other risk factors included in a multivariable model. An adjusted odds ratio results from the comparison of two strata similar at all variables except

exposure (or the marker of interest).

Adjusted odds ratios (Table-5) reveal that out of 8 factors 5 factors are significantly associated with drivers' injuries. They are personal factors, environmental condition, machinery condition, job satisfaction, and risk taking behavior. On the other hand, factors like AGE2(OR 3.93, 95%CI .624 to 24.64), AGE3(OR 3.19, 95%CI .584 to 17.38), EXP1(OR 2.33, 95%CI .438 to 12.46), EXP2(OR 5.55, 95%CI .699 to 44.06 ), and HRC(OR 2.17, 95%CI .493 to 9.60), are not associated with injuries [7].

**9. LOGISTIC COEFFICIENTS- b**

It presents logistic coefficients  $\hat{a}$  for the case specification, as injury vs no injury. It also shows p-values for Wald chi square statistic. p-values less than 0.05 shows that null hypothesis is rejected, thus factors are significant. The significant parameters are indicated by an asterisk at 0.05 probability level of significance. From this, it is seen that none of the age categories are statistically significant for injury vs no injury relationship. Similarly, none of the experience categories are significant. Highly significant categories are personal factor and risk taking behavior.

	AGE(YR)	EXPERIENCE(YR)	PERS L FAC	ENV IN	MAC HI NAR	HAU L ROAD	JOBS ATIS	RISK T AKB
N	113	113	113	113	113	113	113	113
Missing	0	0	0	0	0	0	0	0
Mean	42.42	15.74	9.65	9.85	6.92	6.82	11.08	20.20
Median	42.00	15.00	9.00	10.00	7.00	9.00	11.00	18.00
Mode	40	14	9	12	9	9	13	13
Std. Dev.	3.95	4.27	1.81	2.48	2.13	2.53	2.84	3.44
Minimum	34	4	7	4	3	3	5	14
Maximum	54	27	15	12	9	9	15	28

**Table-1: Statistics**

**Note :**

N= Number of cases, AGE(YR) = Age of the dumper driver in years, EXPERIENCE(YR)= Experience of dumper drivers in years, PERSLFAC = Personal factor  
ENVIN=Environmental factor, MACHINAR=Machinery condition,  
HAULROAD =Haul-road condition , JOBSATIS=Job satisfaction  
RISKTAKB=Risk taking behavior. Valid= Valid cases, Missing = Missing cases ( In this case study there is no such missing case)

**Table-2: Categorization of variables**

Variables	categories	Name
AGE	3	AGE1<=40, 40<AGE2<44, AGE3>=44
EXPERIENCE	3	EXP1<15, 15<=EXP2<17, EXP3>=17
PERS FACTORS	2	PERF1<10, PERF2>=10
ENVIRON COND	2	ENVC1<11, ENVC2>=11
MACH COND	2	M/CC1<8, M/CC2>=8
HAUL RD COND	2	HRC1<9, HRC2>=9
JOB SATIS	2	JS1<11, JS2>=11
RISK TK BEH	2	RTB1<19, RTB2>=19

**Table-3:** Crude Odds Ratios With 95% Confidence Int.

	Crude odds ratio	Confidence interval
AGE2vs AGE1	5.03	1.50 – 16.92
AGE3 vs AGE1	8.09	2.47 – 26.50
EXP1 vs EXP3	1.64	0.638 – 4.197
EXP2 vs EXP3	5.80	1.50 – 22.37
PERF2 vs PERF1	7.87	2.21 – 28.03
ENVC2 vs ENVC1	8.33	2.66 – 26.06
M/CC2 vs M/CC1	6.25	2.00 – 19.52
HRC2 vs HRC1	3.67	1.46 – 9.25
JS2 vs JS1	3.25	1.34 – 7.87
RTB2 vs RTB1	7.42	2.57 – 21.37

**Table-4:** Description of variables

Variables and Category	Indicator Coding	
	Options Used in Log. Reg.	
AGE1<=40	X <sub>1</sub>	X <sub>2</sub>
40<AGE2<44	0	0
AGE3>=44	1	0
	0	1
EXP1<15	X <sub>3</sub>	X <sub>4</sub>
15<=EXP2<17	1	0
EXP3>=17	0	1
	0	0
PERF1<10	X <sub>5</sub>	
PERF2>=10	0	
	1	
ENVC1<11	X <sub>6</sub>	
ENVC2>=11	0	
	1	
M/CC1<8	X <sub>7</sub>	
M/CC2>=8	0	
	1	
HRC1<9	X <sub>8</sub>	
HRC2>=9	0	
	1	
JS1<11	X <sub>9</sub>	
JS2>=11	0	
	1	
RTB1<19	X <sub>10</sub>	
RTB2>=19	0	
	1	

Here the scaling describes the procedures of assigning

numbers to various degrees of opinion, attitude and other concepts. This can be done in two ways viz., (i) making a judgment about some characteristics of an individual and then placing him directly on a scale that has been defined in terms of those characteristics and (ii) constructing questionnaires in such a way that the source of individuals responses assigns him a place on a scale. The rating scale as mentioned here, involves qualitative description of a limited number of aspects of a thing or of traits of a person. Considering scale properties, the scales may be classified as nominal, ordinal, interval and ratio scales. Variable were categorized on the basis of various case studies and response of quantity of persons to be studied.

**Table-5:** Relationships between various factors and drivers' injury (Injury vs No Injury) adjusted odds ratios and their 95% confidence intervals

	Adjusted odds ratio and 95% confidence interval	
AGE2 vs AGE1	3.93	.626 – 24.64
AGE3 vs AGE1	3.18	.584 – 17.38
EXP1 vs EXP3	2.33	.438 – 12.46
EXP2 vs EXP3	5.55	.699 – 44.06
PERF2 vs PERF1	12.17	1.99 – 74.49
ENVC2 vs ENVC1	5.26	1.06 – 26.07
M/CC2 vs M/CC1	6.35	1.01 – 39.97
HRC2 vs HRC1	2.17	.493 – 9.604
JS2 vs JS1	5.04	1.19 – 21.43
RTB2 vs RTB1	9.85	1.91 – 50.85

## 10. ADJUSTED ODDS RATIOS – REPORTABLE INJURY VS NO INJURY CASE.

In a multiple logistic regression model where the response variable is the presence or absence of a disease, an odds ratio for a binomial exposure variable is an adjusted odds ratio for the levels of all other risk factors included in a multivariable model. It is also possible to calculate the adjusted odds ratio for a continuous exposure variable. An adjusted odds ratio resulted from the comparison to two strata similar at all variables except exposure ( or the marker of interest). It can be calculated when stratified data are available as contingency tables by Mantel-Haenszel test. Adjusted odds ratios reveal that out of 8 factors 4 factors are significantly associated with driver's reportable injuries. They are personal factors, machinery condition, job satisfaction and risk taking behaviors. On the other hand, factors like AGE2, AGE3, EXP1, EXP2, ENVC, HRC are not associated with reportable injuries.

## 11. Ordinal Logistic Rsgression Model Fitting

When the dependent variable has more than two levels, then we use either multinomial or ordinal logistic regression for predicting the dependent variable. When dependent variable categories are in ordered form, then we prefer ordinal logistic regression. Like logistic regression, ordinal logistic regression uses maximum likelihood methods, and finds the best set of regression coefficients to predict values of the logit-transformed probability that the dependent variable falls into one category rather than other. Polytomous universal model can estimate a variety of ordinal regression models including the ordered logit and ordered probit. The ordered logit model, also known as cumulative logit model, estimates the effects of independent variables on the log odds of having lower rather than higher scores on the dependent variables. While nominal logistic regression can be used to analyze ordinal response variables, it is often not the choice because it ignores the ordering of the categories. This model results showing initial -2 Log Likelihood value first, when there is intercept only. Then -2 Log Likelihood value from final model. Difference is giving chi-square value. The value of 65.729 with 10 df is the more relevant value here. This is the likelihood ratio test that all coefficient for all independent variables are equal to zero. This null hypothesis is rejected since the test is highly significant. Next is goodness of fit test which is not significant, indicating that the model fits the data well. However, this test is not informative because of the large number of zero frequencies in the three way table of the variables in use. This information is really only relevant if the small number of categorical independent variables is used. Here PseudoR-square measures indicate that the model is performing fairly well. As we know that PseudoR-square measures confound goodness of fit and explanatory power of the model. This is done so that positive coefficient indicates increased likelihood of higher scores on the dependent variable.

## 12. CONCLUSION

An attempt has been made in this paper to analyze reasons of accidents of dumper drivers due to personal and organizational factors. The chance of surviving of an accident is greater when a seat belt is worn by the drivers and helpers. Drugs and alcohol dull senses. They reduce response time, attention span, and person's ability to identify an impending problem. A person under the influence of drugs or alcohol is more likely to cause a serious injury or death to himself and a co-worker, not to mention damage to company property. Risk taking behavior must be controlled by any driver. Before operation of any dumper, driver must take very precautionary measures. It is important to determine what factors are causing injuries in mines so that appropriate preventive measures can be taken in coal mining industry. The latter must be thoroughly checked by the driver regarding its healthiness. In case of bad weather including the unsafe road condition. The variables those show distinct relationship and considered to be significant with the degree of injury are the personal factors, environmental condition, machinery condition, job satisfaction and risk taking behavior and their respective adjusted odds ratios were found 12.17, 5.26, 6.35, 5.04 and

9.85 for significant risk factors in relation to Reportable vs No Injury case. In other words the above mentioned high risky significant factors are 12.17, 5.26, 6.35, 5.04 and 9.85 times more susceptible to injury in comparison to reference category. Out of the above 5 factors, most highly significant categories are personal factor and the risk taking behavior. Training must be included in the knowledge of a dumper's mechanical system and that system's safe operating ranges, how to conduct a thorough preoperational inspection, and how to get on and off the equipment. At last we come to conclusion that starting from top to bottom level of management, to drivers all are responsible for accident of injury occurrence. It is thus recommended that increased management attention should be given at the mine level to reduce dumper accidents and thus driver's injury.

## ACKNOWLEDGEMENTS

The authors are thankful to the case study mine, research institutes and former colleagues of first author for providing necessary help for preparing this paper.

## REFERENCES

- [1] DGMS Tech. circular 03 of 2013, Ministry of labour and employment, Govt. of India
- [2] DGMS Standard Note - 2015 .
- [3] Kleibaum, Kupper, Muller, Nizam, Applied Regression analysis and other Multivariate Methods, 3<sup>rd</sup> edition
- [4] Singh et al, (2005), A Study of Effect of Personal and Organizational Characteristics on Coal Mine Dumper Drivers Injury Using Logistic Regression Model, National proceeding on Mine safety, IIT Kharagpur
- [5] Minerals Industry Safety Handbook, Edition 1, July, 2002
- [6] Bhattacharjee A, Ray SC, Mukhopadhyay SK, Saran MB, Ghosh AK (2000). An investigation into personal and impersonal variates in coal mine accidents.
- [7] In: Proceedings of the International Conference on Technology Management for Mining, Processing and Environment, 1-3rd Dec., 2000, IIT, Kharagpur: 297
- [8] Judge TA, Bono JE, Locke EA (2000). Personality and job satisfaction: the mediating role of job characteristics. Journal of Applied Psychology 85 (2): 237-249
- [9] Das Samir Kumar ,(2002) , Mines Safety and Legislation, Lovely Prakashan, Dhanbad

## BIOGRAPHIES



**M. K. Singh<sup>1</sup>**, B.E and M.Tech. Mining Engineering. Dy. Manager (Mining), Collieries Division, Steel Authority of India Limited, Dhanbad, Jharkhand, India. (1 Corresponding Author +91-9470194935, Email- [mksingh4a@gmail.com](mailto:mksingh4a@gmail.com))



**Prof (Dr) U.K. Dey<sup>2</sup>** Professor and Head, Department of Mining Engineering, BIT Sindri, Dhanbad, Jharkhand, India