Exploring Regional Differences in Cyclist Safety at Roundabouts: A Comparative Study between the UK (based on Northumbria data) and Belgium

Nurten Akgün[a]*, Stijn Daniels[b], Margaret Carol Bell[c], Nina Nuyttens[b], Neil Thorpe[c], Dilum Dissanayake[c]

[a]Faculty of Engineering and Natural Sciences, Bursa Technical University, Bursa 16330, Turkey

[b]Vias Institute Brussels, Belgium

[c]School of Engineering, Cassie Building, Newcastle University, Newcastle upon Tyne NE1 7RU, UK

*Corresponding author

* nurten.akgun@btu.edu.tr

Abstract

The level of safety for cyclists at roundabouts may vary according to national differences not only in the design itself but also sociodemographic, cyclist and driver behaviour as well as environmental factors. This paper investigates the national influence on cyclist casualty severity at roundabouts by comparing the United Kingdom (using Northumbria as a representative sample) and Belgium. The data included speed limits, socio-demographic characteristics, environmental conditions and driver/cyclist behaviour-related contributory factors. First, a logistic regression analysis for the UK data, including 864 cyclist casualties, was carried out. Increasing the speed limit by ten units (for example 30mph to 40mph) increased the probability of a cyclist being killed or seriously injured by 10%. A cyclist casualty was more than three times more likely to be killed or seriously injured (the odds ratio is 3.02) where sudden braking was recorded as a contributory factor. Second, a separate logistic regression analysis for Belgium was conducted. Cyclists ignoring the priority at roundabouts increased the probability of a fatal or seriously injured collisions (the odds ratio is 2.71). Comparing the individual analysis for both countries, the influence of cyclist age was consistent. Each one-year increase in cyclist age increases chance of being killed or seriously injured as opposed to not being killed or seriously injured by 2% (odds ratio is 1.02) in both UK and Belgium. A final comparative analysis was applied considering proxy variables for both countries. Three-way chi-square tests of independence showed that all non-behavioural variables (i.e. sociodemographic characteristics, speed limit, and environmental conditions) were found to be statistically different between UK and Belgium for both slight and killed and seriously injured casualties. This suggests that driver/cyclist interaction and behaviour in the two countries is generally similar whilst speed limits, the sociodemographic characteristics of cyclists and environmental conditions are specific for each country. The third part of the logistic regression analysis suggested that the country residual was highly statistically significant. This indicates that there are some statistically significant differences with respect to the characteristics of the two regional datasets used in the analysis.

Keywords: Cyclist safety; Roundabout; Contributory factors; Three-way chi-square, Multilevel logistic regression

1. Introduction

Roundabouts alter or eliminate conflict points between road users (Montella, 2011) and vehicles are forced to reduce their speed by provided deflection (Gross *et al.*, 2013; Silvano and Linder, 2017). Therefore, they are generally regarded as safer intersections for motor vehicle drivers compared to priority or signalised junctions (Montella, 2011). However, high level of safety does not exist for cyclists. For example, Daniels *et al.* (2008) indicated a 27% increase in collisions involving cyclists after the conversion of intersections into roundabouts. De Brabander and Vereeck (2007) compared safety at roundabouts and signal-controlled intersections and concluded that roundabouts protect vulnerable users less effectively than signalised intersections.

Cycling has journey-specific benefits such as being a fast door to door travel mode in urban areas (Parkin, 2018). However, many people still hesitate to cycle in daily life due to safety concerns (Campisi *et al.*, 2020). This is a barrier to increasing the number of people cycling and realising the benefits of cycling on health, to the environment and economic perspectives. This issue has been a motivation for previous studies (Hels and Orozova-Bekkevold, 2007; Møller and Hels, 2008; Daniels *et al.*, 2009; Sakshaug *et al.*, 2010; Silvano *et al.*, 2015; Jensen, 2017), which provided recommendations for policy makers and design engineers to improve safety for cyclists. However, there is a difference in design dimensions and human behaviour related factors in each country (Tollazzi, 2015). Therefore, a comparative analysis including important contributors (i.e. design, environmental and human related variables) should be investigated in detail to keep encouraging people to cycle (Akgün *et al.*, 2018).

The research reported in this paper identifies country based influence on the relative contribution of the variables (including speed limit, the sociodemographic characteristics of cyclists involved in casualties, environmental conditions and driver behaviour related contributory factors) on the severity of cyclist casualties with the focus on give-way roundabouts with mixed traffic. This comparison is carried out to develop a richer understanding of cyclist casualties by introducing an international dimension. The research questions in this paper are therefore as follows:

1. What are the relative contributions of key variables to the severity of cyclist casualties?
2. To what extent does the national differences with regards to the associated data influence the regression analysis for different countries? and
3. Does the interpretation of the results have a meaning for another country with a different driving/cycling behaviour, roundabout design and environmental conditions?

Section 2 presents a critical state of the art review on cyclist safety at roundabouts followed in Section 3 by the modelling techniques adopted to investigate the National Influence. A wide range of statistical methods used in previous studies are assembled to determine the most appropriate comprehensive set of analytical and prediction methods. Section 4 covers the analytical steps adopted in this research including data collection, descriptive analysis is presented. Next in Section 5 the actual data analysis and the results are presented. Section 6 discusses the results before Section 7 elaborates on the limitations before the final section 8 where conclusions are drawn.

2. A Critical State of the Art Review for Cyclist Safety at Roundabouts

Converting signalised and priority junctions to give-way roundabouts increases the capacity of an intersection generating environmental benefits such as lower emissions due to less queuing and improved safety for vehicle drivers (Akgün, 2019). This has led to a wide range of applications of roundabouts with numbers rising every day. Despite the advantages of roundabouts, they are less safe for cyclists compared to vehicle drivers (De Brabander and Vereeck, 2007; Daniels *et al.*, 2010; Daniels *et al.*, 2011; Polders *et al.*, 2014; Jensen, 2017). Daniels *et al.* (2008) found that fatal and serious injury for cyclists increased by 41-46% after roundabout construction. The concern over the safety of cyclists at roundabouts has led to a considerable amount of research in this area.

The main factors governing higher numbers of collisions and more severe injuries for cyclists are identified as higher speed and speed related geometric design parameters of a roundabout (Akgün *et al.*, 2018). Silva *et al.* (2014) stated that roundabouts have a significant impact on speed reduction by approximately 30% (between 26% and 37%) and the impact area was between 400m and 500m. One of the reasons for the higher collision probability for cyclists is the inconsistency in speed reduction behaviour of drivers on approach and at entry locations of roundabouts (Austroads, 2009). 80% of all collisions occur while circulating and or at the entry locations of roundabouts (Polders *et al.*, 2014).

In addition, the most common type of collision for cyclists at roundabouts occurs when the cyclist is circulating and a motor vehicle is entering the junction (Davies *et al.*, 1997; Rasanen and Summala, 2000). According to Davies *et al.* (1997) the reason for this type of collision is drivers' lack of awareness and subsequent failure to yield to cyclists already on the roundabout. Instead, drivers tend to focus on positioning their vehicle and negotiating other vehicles, taking less notice of the smaller dangers such as cyclists present on the roundabout. This theory is supported by evidence that cyclist collision rates at roundabouts are lower when cyclist volume is low (Davies *et al.*, 1997).

Arnold *et al.* (2010) stated that 32% of cyclists feel uncomfortable at roundabouts and 25% of cyclists reported changing their routes to avoid using them. Davies *et al.* (1997) also found that cyclists avoided using roundabouts due to safety concerns Arnold *et al.* (2010) concluded that traffic considerations were less important than land use, connectivity and directness which means that safety studies also should consider other variables, such as behaviour related contributory factors, in addition to traffic issues. Cyclists avoiding roundabouts means they use less direct routes, with longer journeys and in turn presents a barrier discouraging others to take up cycling. On the other hand, one can argue that this is a good thing because by avoiding less safe locations the collisions figures are kept low.

Previous research has been critically reviewed to endorse the research gaps identified and to inform which variables have been found to be important and therefore need to be considered in this research, the results of which are presented later sections of this manuscript. Studies (Rasanen and Summala, 2000; Silvano *et al.*, 2015) claimed that the *possible* major factors related to cyclist-vehicle collisions might be driver behaviour and yielding problems. 7-15% of drivers were found not to be aware of cyclists approaching from the right (Rasanen and Summala, 2000). The main contributory factor to collisions is that drivers are not looking properly to the right side where cyclists appear unexpectedly (Rasanen and Summala, 2000). Another yielding problem was found to be high approach speed. If a driver's approach speed is higher, their yielding behaviour towards cyclists' decreases. Silvano *et al.* (2015) conversely suggested that cyclist speed has a slight effect on vehicle yielding behaviour. while any increase in vehicle speed causes a sharp decrease in the yielding probability. If the vehicle speed is under 20 km/h, yielding rate is expected to level off. On the other hand, cyclists are often very confident that vehicles will give priority to them; however, this presumption reduces the safety for cyclists. Vehicle driver behaviour has a strong impact on cyclist position at the roundabout Silvano *et al*. (2015).

The studies (Rasanen and Summala, 2000; Møller and Hels, 2008; Sakshaug *et al.*, 2010; Silvano *et al.*, 2015; Silvano and Linder, 2017) addressed the impact of behaviour and focused on the probability of vehicles' yielding to cyclists. These earlier studies tend to investigate the influence on cyclist casualty severity of either the behaviour of the driver or the rider but nor both perspectives in the same study. Such research appears absent from the literature, possibly because this type of data is not commonly collected nor is it straightforward to collect. The research presented in this paper seeks to address this research gap.

A wide range of former studies (Lawton *et al.*, 2003; Hels and Orozova-Bekkevold, 2007; Møller and Hels, 2008; Daniels *et al.*, 2009; Sakshaug *et al.*, 2010; Silvano *et al.*, 2015; Jensen, 2017; Akgün *et al.*, 2018) of cyclist safety at roundabouts in different countries, investigated the impacts of contributory factors on either the number of casualties or their severity. However, considering the valid results for different countries independently may not be appropriate because each study focussed on the requirements specific to the design philosophy and dimensions of their own country. Consequently, Tollazi (2015) even argued that a safe design solution in one country might be very dangerous in another. The lack of across country comparative studies was the motivation for the research presented in this manuscript.

The designs of roundabouts such as radial (continental Europe) and tangential (UK) structures (See Figure 1) may have different impacts on safety. In a radial base design, the arms of the roundabout as their name suggests are defined as radial to the centre. This potentially brings a key advantage in terms of safety from speed reduction due to radial roundabouts having a relatively tight geometry at entry locations compared to tangential roundabouts. However, this has the disadvantage of reduced traffic capacity. Tangential roundabouts are found mainly in the UK, New Zealand and Australia and have arms that are tangential to the centre of the roundabout. Speed reduction is achieved with a deflection at the entry; however, both traffic speed and capacity remain high (Patterson, 2010) as the fundamental principles of the design. Davies *et al.,* (1997) and Lawton *et al.,* (2003) concluded that the tighter geometry in the radial design at approach increases the safety. On the other hand, tangential design, allows a higher capacity but with less safety, particularly for vulnerable users in the circulatory traffic flow, due to the higher entry speed.

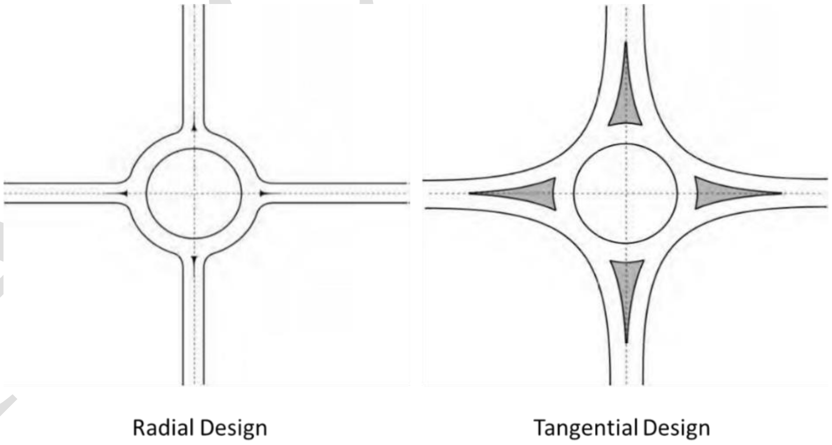

Figure 1 - Radial and Tangential Design of Roundabouts (Patterson, 2010)

In addition to the physical design difference, each region may have different sociodemographic and environmental characteristics. It is seen that there has been no comparative study yet which illustrates the regional influence in a developed cyclist casualty severity analysis which also includes a comprehensive set of other variables including geometric design, as well as environment and sociodemographic. With respect to the literature review, several research gaps were identified, namely cyclist casualty severity analysis, influence of speed limit and driver/rider behaviour on casualty severity and investigating the consistency of casualty modelling including different countries. Previous studies did not consider a comparative analysis including a wide range of variables, such as traffic, sociodemographic, environmental and behaviour related contributory factors, all in one model mainly due to issues of data availability. The interpretation of the logistic regressions was very narrow and detailed analysis, such as calculating the margins, was not considered. It is important to address these gaps because cycling is increasing every year in response to local government policy.

This paper aims not only to address the research gap in cyclist safety at roundabouts across different geometric designs but also to conduct a novel methodology developed specifically to achieve the objectives. Therefore, the structure of the data, limitations and assumptions of the selected analytical methods need a careful investigation before carrying out the analysis. These modelling techniques are presented in the following section.

3. Modelling Techniques for Investigating National Differences

Descriptive statistics were widely used in former studies (Lawton *et al.*, 2003; Møller and Hels, 2008; Daniels *et al.*, 2010; Daniels *et al.*, 2011; Akgün *et. al.*, 2018) because it is essential to understand the basic features of the data before considering any in-depth analysis. The statistics continue with a test for normality, comparing the sample distribution with a normal distribution (Field, 2009). This is important to ensure the assumptions for a specific modelling technique are met (Ghasemi and Zahediasl, 2012) and non-parametric tests are used as appropriate. There is a common statement in the literature that the

Shapiro-Wilk test gives the best result with the highest power compared to other tests (Razali and Wah, 2011; Ghasemi and Zahediasl, 2012).

In addition to descriptive statistics, cyclist casualty severity related former research applied a wide range of statistical methods such as *correlation analysis* and *dimension reduction* (Akgün *et al.*, 2018); *pearson's chi-square* (Møller and Hels, 2008; Polders *et al.*, 2014), *ANOVA* (Rasanen and Summala, 2000), *meta-analysis* (De Brabander and Vereeck, 2007; Daniels *et al.*, 2008; Daniels *et al.*, 2009) and *before and after comparison* (Sakshaug *et al.*, 2010; Arnold *et al.*, 2010). Previous studies also conducted analytical models which were *linear regression* (Møller and Hels, 2008; Daniels *et al.*, 2009), *logistic regression* (Hels and Orozova- Bekkevold, 2007; Daniels *et al.*, 2010; Silvano *et al.*, 2015), *poisson/gamma regressions* (Hels and Orozova-Bekkevold, 2007; Daniels *et al.*, 2011), *hierarchical binomial logistic regression* (Daniels *et al.*, 2010), and *Empirical Bayes* (Daniels *et al.*, 2009).

The analysis of the relationship between two dependent variables is often carried out by testing a null hypothesis such as "A higher speed increases the crash rates". These kinds of studies need basic statistical methods such as correlation analysis and Pearson's chi square test rather than a regression model and it is normally applied when the dataset is limited for fitting into a selected regression model (Harrell, 2001). Pearson's chi-square test is applicable for investigating the impact of each categorical variables individually on cyclist casualties Polders et al. (2014).However, whether the analysis has one or more predictors, if fitting a model is a requirement or the aim of the study is investigating the impacts on an outcome, regression models need to be considered.

The main challenge in data analysis is choosing and using the correct regression model. The appropriate model should analyse the data effectively, align with the whole structure of study aim, be appropriate for further development and be flexible to allow the analysis to be further extended (Harrell, 2001). With respect to the cyclist casualty severity analysis, former research mainly focused on logistic regression and empirical bayes modelling. Logistic regression creates a probability prediction model regarding response and observed variables (Field, 2009), while empirical bayes develops the model by predicting the outcomes by comparing the observed data to prior knowledge in the literature or variation achieved by Monte Carlo Simulation (Efron, 2013). It was observed that the former studies investigated the casualty severity (Daniels *et al.*, 2010; Akgün *et al*., 2018) or yielding probability (Silvano *et al.*, 2015) by applying logistic regression. In logistic regression, the predicted outcome is the probability of an event (Y) occurring given the predictors of X1, X2… Xi. Since the probability of an event should be between 0 and 1, the predicted outcome Y should be in this interval. If the outcome value Y is close to 0 (Probability~0%) it means that Y is unlikely to occur, meanwhile an outcome close to 1 (Probability~100%) means that Y is likely to occur (Field, 2009). The odds ratio, the ratio of odds of success to odds of failure, (Agresti, 2007), is usually used when interpreting the results. The log of odds of the outcome provide the equation of predictors in the logistic regression. There are several types of logistic regression: i) binary logistic regression (two responses 0 and 1, 'yes or no', 'yielded or not yielded' or 'slight or serious'), ii) ordinal logistic regression (with a minimum of three ordered responses) and iii) nominal logistic regression (with multilevel response without ordering) (Field, 2009).

The study in this paper aimed to develop a model to investigate the impact of variables (including sociodemographic characteristic of cyclist, environmental conditions, speed limit, and driver/rider behaviour related contributory factors) on cyclist casualty severity at roundabouts. Regarding this aim, Empirical Bayes could not be an option since the aim was not to develop a prediction model based on prior data. Therefore, a logistic regression should be applied.

With respect to the predictors, there are two main types of logistic regressions. Simple logistic regression is the first form with one response and one predictor and multiple logistic regression is the second with more than one predictor variables. However, there is a limitation for the number of predictor variables in one multiple logistic regression. Agresti (2007) identified that including too many variables causes poor standard errors. The minimum limit of 10 events per variable (EPV) was suggested by the early research (Peduzzi *et al.*, 1996) as a way to restrict the number of variables. Although this limitation was a "*rule of thumb*" in the literature and Ogundimu *et al.* (2016) stated that considering EPV eliminated the bias that occurs with low prevalence of predictors in a regression model. Therefore, it is highly recommended to consider the EPV in regression analysis. If there is a need to reduce the number of variables in the model, (Sperandei, 2013) suggested relaxing p-value criteria. This criterion applies both simple and multiple logistic regressions when selecting the statistically significant variables at 90% confidence level. The selected variables should be included in a final multiple logistic regression.

In some cases, the data is nested in groups and the response variables (i.e. casualty severity) nested in the same groups are more likely to function in the same way than response variables nested in a different group (Sommet and Morselli, 2017) (See Figure 2). For instance, the impact of weather on casualty severity may be different across cities. In that case, the nested cluster impact in the model occurs and multilevel logistic regression should be applied (Sommet and Morselli, 2017).

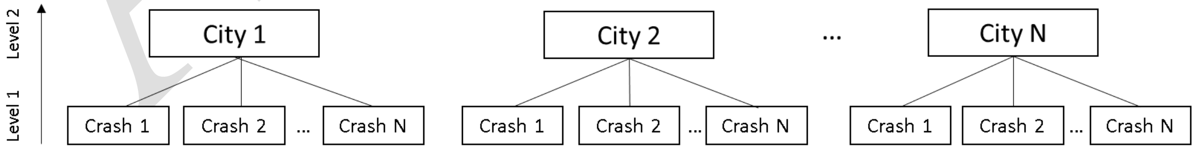

Figure 2 - Multilevel Logistic Regression for Nested Grouped Data

The aim in multilevel logistic regression is to estimate the effect of covariates at a regional (in this study National) level (Li *et al.*, 2011). Previous studies did not consider this type of regression model in their analysis probably because they did not conduct a study which has a large number of nested data in groups as was the case in the research presented in this manuscript. The equation of the multilevel logistic regression (subtype: random intercept model) used to understand the influence of specific variables on cyclist casualties is given below (Steele, 2009):

$$log\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \beta_0 + \beta_1 X_{ij} + u_j$$

$$u_j \sim N(0, \sigma_u^2)$$

where:
β_0 = the log-odds that y = 1 when x = 0 and u = 0
β_1 = the effect on log-odds of one unit increase in x for individuals in same group
u_j = is the effect of being in group j on the log-odds that y = 1
σ_u^2 = is the level 2 (region) variance

4. Method

The analytical approach is divided into two steps. First, developing individual models for both the UK (Northumbria) and Belgium. A descriptive analysis and Shapiro-Wilk test for normality, using SPSS, were carried out in order to gain a better understanding of the data before interpreting the results. This was followed by a logistic regression analysis using STATA based on relaxing the p-value criterion. A simple linear regression, calculating the predictive margins, were applied in order to interpret the results and develop suggestions for policy makers and design engineers.

Secondly, a comparative analysis was carried out including proxy data to equivalence variables collected in Belgium with those that differed in precise nature from those in the UK variables for both countries. Three-way chi square test of independence using SPSS was applied to gain a deeper understanding of the relationship between variables. Simple and multiple multilevel logistic regressions using STATA were carried out to determine the regional estimated variance.

Regarding the data availability, the first analysis covered Northumbria region in the north east part of England. Cyclist casualty records (between 2005 and 2016) including the sociodemographic characteristics of each injured cyclist, the prevailing environmental conditions, the speed limit at the collision site and behaviour-related contributory factors were obtained from the CIRTAS database (developed based on Stats19 by Gateshead Council). The second stage of the analysis was conducted to investigate the impact of the variables on cyclist casualty severity recorded in all areas in Belgium. The data for the same period (i.e. 2005-2016) was obtained from VIAS Institute located in Brussels, Belgium. Sociodemographic characteristics of cyclists and environmental conditions were recorded in the same way in both datasets. However, the data recording protocol used in the UK is different to Belgium, descriptors needed to be associated with each other in order to standardise the data across the two countries before carrying out a comparative analysis. In this research the Northumbria data is used as representative of the UK because the required variables across 2005-2016 were not available for other cities. Given that there are stringent government guidelines that are used for the design of roundabouts across the UK, as well as the a standard format STATS19 for recording details of the accidents, it is a reasonable assumption to use Northumbria as representative of the UK. For the remainder of this manuscript the Northumbrian data used in the analysis will be referred to as that of the UK.

Speed limit was given in units of kilometres per hour in the Belgium data. Therefore, speeds were converted to miles per hour. Not all variables were collected/defined in the same way. Therefore, reconciliation of the data was carried out to determine variables from the UK which could be considered as equivalent to and therefore proxies for those in the Belgium dataset and vice versa. Four contributory factors that were reconciled with each other for the UK and Belgium are as follows: *Junction overshoot*, *Poor turn or manoeuvre*, *Passing too close to cyclists*, *Loss of control*. The data descriptor and numeric code were assigned to the cyclist casualty severity, cyclist gender, lighting, weather and road surface condition. Finally, proxy data was applied in the merged multilevel analysis.

The UK data included 729 slight, 133 serious and 2 fatal casualties that occurred between years 2005 and 2016. On the other hand, 8 unharmed, 855 Slight, 60 serious and 1 fatal casualty were recorded in Belgium during the same period. Due to the low number of unharmed and fatal casualties, slight and unharmed data were merged as *Slight*; serious and fatal were combined as *Killed or Seriously Injured* (KSI). This is common practice in accident analysis because a low number of records may mislead the interpretation of the results. Throughout the analysis, unless otherwise stated, the values for 95% confidence interval are provided in the regression models.

Descriptive Statistics

This section continues with the descriptive statistics for considered predictor variables namely cyclist age and gender, lighting, weather, road surface condition and speed limit (See Table 1).

Table 1- Descriptive Statistics for Sociodemographic Characteristics, Environmental Conditions and Speed Limit

Variable (Unit)	**Northumbria**	**Belgium**
Cyclist age (years)	Mean= 37.8; Min.= 4; Max.= 83	Mean= 39.2; Min.= 1 Max.= 99
Cyclist gender (Female, Male) (number of people)	Female= 95; Male= 769	Female= 326; Male= 598
Lighting (Darkness; Daylight; Unknown)	Darkness= 135; Daylight= 666; Unknown= 0	Darkness= 107; Daylight= 813; Unknown= 4
Weather (Fine; Rain/Snow; Fog/Mist)	Fine= 726; Rain/Snow= 115; Fog/Mist= 23	Fine= 793; Rain/Snow= 98; Fog/Mist= 37
Road Surface (Dry; Wet/Damp; Snow/Frost/Ice; Unknown)	Dry= 587; Wet/Damp= 265; Snow/Frost/Ice= 12; Unknown= 0	Dry= 623; Wet/Damp= 170; Snow/Frost/Ice= 114; Unknown= 17
Speed limit (20; 30; 40; 50; 60; 70) (mph)	20= 5;30= 647; 40= 83; 50=23; 60= 67; 70= 39	
Speed limit (30; 40; 50; 60; 70; 90; 120) (kph)		30= 63; 40= 1; 50= 767; 60= 2; 70= 57; 90= 28; 120= 1

Shapiro-Wilk test for normality suggested that the null hypothesis (states that the data is normally distributed) was rejected (p-value=0.00) with 95% confidence level for cyclist age for both UK and Belgium. Given that *other* data, such as number of roundabouts for each speed band or cycling volume at mixed traffic roundabouts, was not known, further investigation was not possible.

The UK data did not register which contributory factor was associated to which person (cyclist or driver). Therefore, in this analysis, the contributory factors are allocated to everyone involved in the collision. Table 2 shows that '*failed to look properly'* was recorded in 67% of the KSI casualties. This by far was found to be the most frequently reported contributory factor for the UK. This variable was followed by '*failed to judge other person's path'*, '*passing too close to cyclist'* and '*poor turn or manoeuvre'*.

Table 2 - Descriptive Statistics for Contributory Factors Recorded in the UK (Northumbria)

Variable Name	**Slight**			**KSI**		
	Yes	**No**	**% of Presence/ Absence**	**Yes**	**No**	**% of Presence/ Absence**
Junction overshoot	33	696	4.74	8	127	6.30
Junction restart	26	703	3.70	7	128	5.47
Poor turn or manoeuvre	74	655	11.30	10	125	8.00
Failed to signal or misleading signal	15	714	2.10	2	133	1.50
Failed to look properly	517	212	243.87	91	44	206.82
Failed to judge other person's path or speed	159	570	27.89	17	118	14.41
Passing too close to cyclist	83	646	12.85	13	122	10.66
Sudden braking	12	717	1.67	6	129	4.65
Swerved	3	726	0.41	0	135	0.00
Loss of control	9	720	1.25	3	132	2.27
Aggressive driving	8	721	1.11	1	134	0.75
Careless, reckless or in a hurry	154	575	26.78	28	107	26.17
Nervous, uncertain or panic	6	723	0.83	1	134	0.75
Driving too slow for condition or slow vehicle	63	729	8.64	0	135	0.00
Learner or inexperienced driver/rider	3	726	0.41	0	135	0.00
Inexperience of driving on left	1	728	0.14	0	135	0.00
Unfamiliar with model of vehicle	63	729	8.64	0	135	0.00

In contrast to UK, contributing factors are registered separately for cyclists and drivers in the Belgian data. This separation provided an advantage to explore any differences of influence for each road user separately. Table 3 showed that driver's '*non respect of the priority'* was by far the most reported contributory factor for both *Slight* and *KSI* casualties in Belgium.

Table 3- Descriptive Statistic for Contributory Factors Recorded in Belgium

Variable	**Cyclist**				**Driver**			
	Slight	**Slight % of Presence/ Absence**	**KSI**	**KSI % of Presence/ Absence**	**Slight**	**Slight % of Presence/ Absence**	**KSI**	**KSI % of Presence/ Absence**
Non respect of the priority	0(827) 1(36)	4.35	0(55) 1(6)	10.91	0(579) 1(284)	49.05	0(42) 1(19)	45.24
Cross the white line continues	0(862) 1(1)	0.12	0(61) 1(0)	0	0(863) 1(0)	0	0(61) 1(0)	0
Performs in extremis an avoidance manoeuvre	0(854) 1(9)	1.05	0(61) 1(0)	0	0(861) 1(2)	0.23	0(60) 1(1)	1.67
No respect for the distance between users	0(856) 1(7)	0.82	0(60) 1(1)	1.67	0(841) 1(22)	2.62	0(59) 1(2)	3.38
Loss of control of the vehicle	0(857) 1(6)	0.70	0(61) 1(0)	0	0(860) 1(3)	0.35	0(61) 1(0)	0
Wrong overtaking	0(860) 1(3)	0.35	0(60) 1(1)	1.67	0(855) 1(8)	0.94	0(61) 1(0)	0
Illegal place on the roadway	0(807) 1(56)	6.94	0(55) 1(6)	10.91	0(857) 1(6)	0.70	0(60) 1(1)	1.67
Fall	0(746) 1(117)	15.68	0(54) 1(7)	12.96	0(859) 1(4)	0.47	0(60) 1(1)	1.67

[a] 1= presence; 0= absence

5. Data Analysis

The following sections present the results for the UK and Belgium separately followed by the analysis of the generalised model with proxy data to illustrate the regional influence.

Analysis of UK cyclist casualty data

Number of *killed or seriously injured* cyclist casualties was 135; however, the number of variables was 18. This does not meet with the EPV limitation for logistic regression analysis. Therefore, as suggested by (Sperandei, 2013) relaxing p-value criteria and comparing the results of simple logistic regression and a full model of multiple logistic regression, allows statistically significant variables at 90% and 95% confidence level to be determined for inclusion in the final analysis. Relaxing p-value analysis showed that *speed limit*, *cyclist age*, '*failed to judge other* person's path or speed' and '*sudden braking*' were all statistically significant at the 95% confidence level in both simple and multiple logistic regressions (See

Table 4). The statistical significance for *cyclist gender*, *weather* and *road surface condition* was not consistent in both regressions because other predictors either increased or decreased their dominance.

Table 4- Results of the analysis of Relaxing P-Value for UK

	Simple Logistic Regression					**Multiple Logistic Regression**				
Variables^a	**Coefficient**	**P- Value**	**Odds ratio (OR)**	**95% confidence interval for OR Lower**	**Upper**	**Coefficient**	**P- Value**	**Odds ratio (OR)**	**95% confidence interval for OR Lower**	**Upper**
Speed limit	0.02	**0.02****	1.02	1.00	1.03	0.02	**0.05****	1.02	1.00	1.03
Cyclist age	0.02	**0.00****	1.02	1.01	1.03	0.02	**0.00****	1.02	1.01	1.04
Cyclist gender	-0.34	0.22	0.71	0.41	1.22	-0.54	**0.06***	0.59	0.33	1.03
Lighting	0.15	0.51	1.16	0.74	1.82	-0.19	0.45	0.83	0.51	1.35
Weather	-0.41	**0.10***	0.66	0.41	1.07	-0.20	0.47	0.82	0.47	1.41
Road surface condition	-0.34	**0.09***	0.71	0.48	1.05	-0.38	0.11	0.69	0.43	1.09
Junction overshoot	0.28	0.48	1.33	0.60	2.94	0.35	0.41	1.42	0.62	3.28
Junction restart	0.39	0.37	1.47	0.63	3.48	0.48	0.29	1.62	0.67	3.95
Poor turn or manoeuvre	-0.35	0.33	0.71	0.36	1.41	-0.44	0.23	0.64	0.31	1.32
Failed to signal or misleading signal	-0.33	0.67	0.72	0.16	3.17	-0.18	0.82	0.84	0.18	3.79
Failed to look properly	-0.16	0.41	0.85	0.57	1.23	-0.15	0.47	0.86	0.57	1.30
Failed to judge other person's path or speed	-0.66	**0.02****	0.52	0.30	0.88	-0.65	**0.02****	0.52	0.30	0.91
Passing too close to cyclist	-0.19	0.55	0.83	0.45	1.54	-0.22	0.51	0.80	0.41	1.54
Sudden braking	1.02	**0.05****	2.78	1.02	7.53	1.21	**0.03****	3.36	1.17	9.67
Loss of control	0.60	0.38	1.82	0.49	6.80	0.62	0.38	1.87	0.46	7.57
Aggressive driving	-0.40	0.71	0.67	0.08	5.42	-0.80	0.48	0.45	0.05	4.10
Careless, reckless or in a hurry	-0.02	0.92	0.98	0.62	1.54	0.00	0.99	1.00	0.63	1.61
Nervous, uncertain or panic	-0.11	0.92	0.90	0.11	7.53	-0.22	0.85	0.81	0.09	7.48
Constant						-2.10	0.00	0.12	0.05	0.32

[a] cyclist gender (female vs. male); lighting (darkness vs. daylight, unknown); weather (fine vs. rain/snow; fog/mist); road surface condition (dry vs. wet/damp, snow/frost/ice, unknown); contributory factors (yes vs. no)
* Statistically significantly at 90% confidence level
**Statistically significantly at 95% confidence level

The selected variables (found to be at statistically significant at least with a 90% confidence) were used as predictors in the final multiple logistic regression (p-value=0.00 at 95% confidence level of significance of the likelihood ratio test) (See Table 5). One unit higher in *speed limit* decreases the safety resulting in a 1% higher probability in *killed or seriously injured* compared to *slight casualty* (odds ratio=1.01) whilst in the case of a cyclist of a higher compared to lower *age group* (odds ratio=1.02). The casualties were twice as likely to be *slight* for reported contributory factor namely '*failed to judge other person's path* or ‘*speed*’ (odds ratio=0.51). The chances of a cyclist casualty being *killed or seriously injured* compared to only being *slightly injured* was three times higher when ‘*sudden braking’* was noted as a possible contributory factor (odds ratio=3.02). Given that the number of collisions involving ‘*sudden braking*’ was very low (n=?), the wide range of the 95% confidence interval from 1.09 to 8.38 suggests that a *killed or seriously injured* casualty compared to a *slight* casualty could have almost equal or up to eight times as much chance of happening. In addition, the relationship between ‘sudden braking’ and geometric design suggested no statistically significant relationship; therefore, the impact on severity of accident of ‘*sudden braking’* is related to other influences rather than the *geometry* of the roundabout.

Table 5- Multiple Logistic Regression Including Selected Variables (with a p-value < 0.1 in either the simple logistic regression or the multiple logistic regression)

Variables ^{a}	**Coefficient**	**P- Value**	**Odds ratio (OR)**	**95% confidence interval for OR**	
				Lower	**Upper**
Speed limit	0.02	**0.05****	1.01	1.00	1.03
Cyclist age	0.02	**0.00****	1.02	1.01	1.03
Cyclist gender	-0.50	0.07*	0.60	0.35	1.06
Weather	-0.18	0.51	0.84	0.49	1.43
Road surface condition	-0.31	0.18	0.74	0.47	1.15
Failed to judge other person’s path or speed	-0.67	**0.02****	0.51	0.30	0.88
Sudden braking	1.11	**0.03****	3.02	1.09	8.38
Constant	-2.30	0.00	0.10	0.04	0.23

[a] cyclist gender (female vs. male); weather (fine vs. rain/snow; fog/mist); road surface condition (dry vs. wet/damp, snow/frost/ice, unknown); contributory factors (yes vs. no)
* Statistically significantly at 90% confidence level
**Statistically significantly at 95% confidence level

Considering the constant odds ratios of cyclist age in both simple and final multiple logistic regressions, the age at which cyclists are found to ride most safely can be determined. This knowledge is helpful to policy makers by identifying those groups to target education and training campaigns. Therefore, the model was reversed, and a simple linear regression was applied because cyclist age (i.e. a constant) became the dependent variable and *casualty severity* was the predictor in the model. The result of simple linear regression suggested that cyclists who below the age of 41 years old were less likely to be *killed or seriously injured* than the older population groups.

Analysis of Belgian cyclist casualty data

A similar analytical approach, which was relaxing p-value criterion for developing a logit model, was applied to the Belgian data (See Table 6). Cyclist *age* and cyclist's '*non respect of the priority'* was statistically significant at the 95% confidence level in both simple and multiple logistic regressions. On the other hand, *speed limit* in simple logistic regression and '*vehicle's performs in extremis an avoidance manoeuvre'* in the multiple logistic regression were statistically significant at 90% confidence level.

Table 6- Results of the analysis of Relaxing P-Value for Belgium

Variables[a]	**Simple Logistic Regression**					**Multiple Logistic Regression**				
	Coefficient	**P- Value**	**Odds ratio (OR)**	**95% confidence interval for OR**		**Coefficient**	**P-Value**	**Odds ratio (OR)**	**95% confidence interval for OR**	
				Lower	**Upper**				**Lower**	**Upper**
Cyclist age	0.02	**0.00****	1.02	1.01	1.04	0.02	**0.00****	1.02	1.01	1.04
Speed limit (kph)	0.02	**0.06***	1.02	1.00	1.04	0.02	0.21	1.02	0.99	1.04
Cyclist gender	0.20	0.49	1.22	0.70	2.14	0.15	0.61	1.16	0.65	2.10
Weather	0.01	0.96	1.01	0.58	1.76	0.10	0.73	1.11	0.62	1.98
Road surface condition	-0.05	0.78	0.95	0.68	1.34	-0.01	0.97	0.99	0.69	1.43
Lighting	0.54	0.26	1.71	0.67	4.35	0.64	0.22	1.90	0.69	5.29
Cyclist's non respect of the priority	0.92	**0.05****	2.51	1.01	6.20	1.13	**0.02****	3.10	1.19	8.08
Cyclist's wrong overtaking	1.56	0.18	4.78	0.49	46.63	1.70	0.15	5.47	0.54	55.14
Cyclist's illegal place on the roadway	0.45	0.32	1.57	0.65	3.81	0.72	0.13	2.06	0.81	5.22
Cyclist's no respect for the distance between users	0.71	0.51	2.04	0.25	16.84	0.89	0.43	2.42	0.27	21.99
Cyclist's fall	-0.19	0.65	0.83	0.37	1.86	-0.26	0.56	0.77	0.33	1.82
Driver's non respect of the priority	-0.08	0.78	0.92	0.53	1.61	0.09	0.75	1.10	0.61	1.98
Vehicle driver performs in extremis an avoidance manoeuvre	1.97	0.11	7.18	0.64	80.26	2.30	**0.09***	10.00	0.72	139.53
Vehicle's illegal place on the roadway	0.87	0.43	2.38	0.28	20.09	1.25	0.26	3.49	0.39	31.14
Vehicle's no respect for the distance between users	0.26	0.73	1.30	0.30	5.64	0.41	0.59	1.51	0.33	6.85
Driver's fall	1.28	0.26	3.58	0.39	32.53	1.44	0.21	4.24	0.43	41.38
Constant						-5.42	0.00	0.00	0.00	0.02

[a] cyclist gender (female vs. male); lighting (darkness vs. daylight, unknown); weather (fine vs. rain/snow; fog/mist); road surface condition (dry vs. wet/damp, snow/frost/ice, unknown); contributory factors (yes vs. no)

* Statistically significantly at 90% confidence level

**Statistically significantly at 95% confidence level

The logistic regression (p-value=0.00 at 95% confidence level of significance of the likelihood ratio test) including selected variables suggested that a unit increase of cyclist age and cyclist's non respect of the priority increased the probability of *killed or seriously injured* casualty occurrence compared to *slight* with 2% and 171%, respectively (See Table 7). Regarding the constant values for cyclist *age* in whether for the simple or final multiple regressions cyclists involved in an accident who are over age 48 were more likely to be *killed or seriously injured* compared to *slight* casualty. Interestingly, the influence of speed limit on casualty severity do no longer exist in the multiple logistic regression (see Table 7) Instead the regression is dominated by cyclist age and cyclist's failure to respect priority in mixed traffic.

Table 7- Multiple Logistic Regression Including Selected Variables (with a p-value < 0.1 in either the simple logistic regression or the multiple logistic regression)

Variables[a]	**Coefficient**	**P- Value**	**Odds ratio (OR)**	**95% confidence interval for OR**	
				Lower	**Upper**
Cyclist age	**0.02**	**0.00****	**1.02**	**1.01**	**1.04**
Speed limit	0.02	0.14	1.02	0.99	1.04
Cyclist's non respect of the priority	**1.00**	**0.04****	**2.71**	**1.07**	**6.87**
Vehicle's performs in extremis an avoidance manoeuvre	1.62	0.20	5.06	0.42	60.70
Constant	-4.64	0.00	0.01	0.00	0.04

[a] contributory factors (yes vs. no)
* Statistically significantly at 90% confidence level
**Statistically significantly at 95% confidence level

Including the statistically significant variables only the final logistic regression Equation is:

$$Logit\ (p) = ln\frac{p}{1-p} = -3.75 + 0.02 * Cyclist\ age + 1.04 * Cyclist's\ non\ respect\ of\ \ priority$$

Where the predicted probability (predictive margin):

$$p = \frac{exp^{logit(p)}}{1 + exp^{logit(p)}}$$

The predictive margins for the logarithmic relationship between cyclist *age* and a cyclist's *'non respect of priority'* is shown in Figure 3.

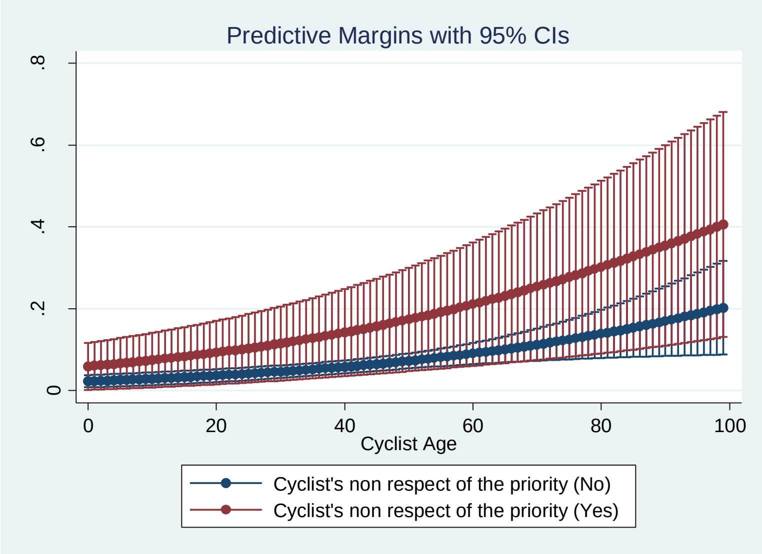

Figure 3- Predictive Margins for Cyclist Age and Cyclist's Non Respect of the Priority

Exploring Differences between UK and Belgium with Proxy Data

A further comparative analysis with ten proxy variables, namely '*cyclist age*', '*cyclist gender*', '*speed limit*', '*lighting*', '*weather*', '*road surface condition*', '*junction overshoot*', '*poor turn or manoeuvre*', '*passing too close to cyclist' and 'loss of control'* was conducted in order to study the relative contribution of the variables to the casualty severity and the differences between the UK and Belgium.

The analysis starts with understanding the association between cyclist casualty severity, the ten considered variables and countries to explore differences between UK and Belgium which are known to have different roundabout design approaches and environments. The three-way chi square test of independence (χ^2), which explores the statistical dependency between these variables, was deemed appropriate (See Figure 4). The variables were ordinal or nominal and consisted of two or more categorical values, therefore the data met both χ^2 assumptions. The rejected null hypothesis in the χ^2 test (with p-value$\leq$0.05 at 95% confidence level) suggested that there is statistically significant evidence of dependence between variables.

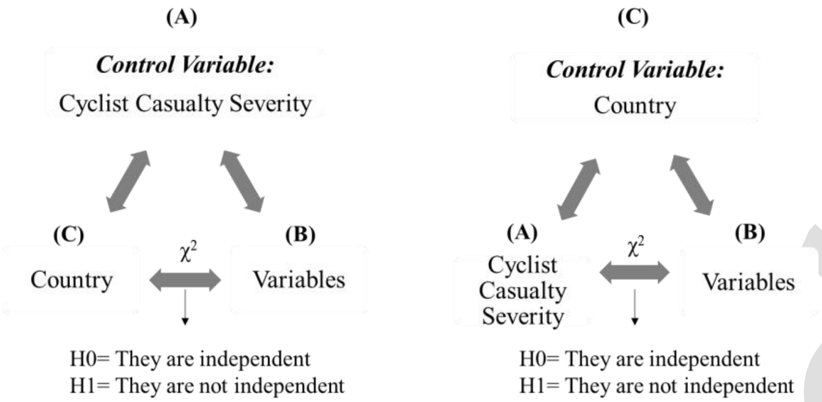

Figure 4 - Principle of Thee-Way Chi Square Test of Independence

The three-way chi square test of independence was applied, and the results at 95% confidence level suggested that a statistically significant relationship was found between country (i.e. UK and Belgium) and certain variables (i.e. cyclist *age group*, cyclist *gender, speed limit, lighting*, *road surface condition*, '*junction overshoot'*, '*poor turn or manoeuvre*' and '*passing too close to cyclists*') among *slight* casualties (See Table 8). Regarding *killed or seriously injured* casualties, the national influence was statistically significant for six variables, namely cyclist *age group*, cyclist *gender, speed limit*, *lighting*, *road surface condition* and '*junction overshoot'* with χ^2 values 29.26, 7.84, 11.59, 6.19, 16.08 and 34.35 (p-values≤0.05), respectively. Regarding the total casualty severity, the null hypothesis was not rejected for weather (χ^2= 4.08 with p-value=0.13) and loss of control (χ^2= 0.66 with p-value=0.42).

Table 8- Three-way Chi-Square Test of Independence between Country and Variables by Casualty Severity

		Cyclist age group	**Cyclist gender**	**Speed limit**	**Lighting**	**Weather**	**Road surface condition**	**Junction overshoot**	**Poor turn or manoeuvre**	**Passing too close to cyclist**	**Loss of control**
Slight Casualty	Pearson χ^2	101.03	136.80	122.83	38.96	4.09	104.57	232.86	61.60	50.09	0.13
	Df[a]	10	1	5	2	2	3	1	1	1	1
	P-value	**0.00****	**0.00****	**0.00****	**0.00****	0.13	**0.00****	**0.00****	**0.00****	**0.00****	0.72
KSI Casualty	Pearson χ^2	29.26	7.84	11.59	6.19	2.06	16.08	34.35	2.64	2.40	1.38
	Df[a]	9	1	5	1	2	2	1	1	1	1
	P-value	**0.00****	**0.01****	**0.04****	**0.01****	0.36	**0.00****	**0.00****	0.10	0.12	0.24
Total Casualty	Pearson χ^2	115.11	146.29	138.24	43.80	4.08	119.51	268.26	62.36	51.69	0.66
	Df[a]	10	1	5	2	2	3	1	1	1	1
	P-value	**0.00****	**0.00****	**0.00****	**0.00****	0.13	**0.00****	**0.00****	**0.00****	**0.00****	0.42

[a] Df (Degrees of freedom)
**Statistically significant at 95% confidence level

The following step explored the three-way chi square test of independence between variables and casualty severity by country (See Table 9). The results suggest that there is a statistically significant dependence between *casualty severity* and *speed limit* for only UK (χ^2= 13.74 with p-value=0.02). On the other hand, it is seen that the dependency between casualty *severity* and cyclist *age group* was statistically significant for Belgium only (χ^2= 24.37 with p-value=0.01). Regarding the combined data, both *speed limit* and cyclist *age group* were statistically significantly dependent on casualty *severity*.

The outcome of the chi square test of independence suggests the statistical associations between three variables; however, it does not provide information on the strength of the dependency. Therefore, a further analysis recommended by (Li *et al*, 2011) was required to estimate the effect of covariates at country level using multilevel logistic regression including proxy data.

Table 9- Three-way Chi-Square Test of Independence between Severity and Variables by Country

		Cyclist age group	**Cyclist gender**	**Speed limit**	**Lighting**	**Weather**	**Road surface condition**	**Junction overshoot**	**Poor turn or manoeuvre**	**Passing too close to cyclist**	**Loss of control**
Northumbria	Pearson χ^2	9.88	1.55	13.74	0.43	2.92	3.74	0.49	0.98	0.36	0.81
	Df[a]	10	1	5	1	2	2	1	1	1	1
	P-value	0.45	0.21	0.02**	0.51	0.23	0.15	0.48	0.32	0.55	0.37
Belgium	Pearson χ^2	24.37	0.49	7.82	1.92	0.70	2.01	0.25	0.06	0.12	0.64
	Df[a]	10	1	4	2	2	3	1	1	1	1
	P-value	0.01**	0.48	0.10	0.38	0.71	0.57	0.62	0.80	0.73	0.42
Total Casualty	Pearson χ^2	20.85	2.11	27.58	0.59	1.90	6.34	3.00	0.03	0.31	0.24
	Df[a]	10	1	5	2	2	3	1	1	1	1
	P-value	0.02**	0.15	0.00**	0.75	0.39	0.10	0.08	0.87	0.58	0.62

[a] Df (Degrees of freedom)

**Statistically significant at 95% confidence level

The null model of multilevel logistic regression suggested that there was a statistically significant (p-value = 0.00 at the 95% confidence level) difference considering the casualty *severity* between the two countries with effects parameter of 0.22. Simple multilevel logistic regression on each variable separately and a full model of multiple multilevel logistic regression were conducted (See Table 10).

Table 10- Comparative Analysis with Multilevel Logistic Regression Model

Variables	**Simple Multilevel Logistic Regression**				**Multiple Multilevel Logistic Regression**		
	Coefficient	**P-Value**	**Estimated Variance of Region**	**P-Value of Residual**	**Coefficient**	**P-Value**	**Estimated Variance of Region**
Speed limit	0.02	0.00	0.17	0.00	0.02	0.01	
Cyclist age group	0.18	0.00	0.24	0.00	0.18	0.00	
Cyclist gender	-0.05	0.82	0.23	0.00	-0.13	0.52	
Lighting	0.22	0.28	0.23	0.00	0.01	0.96	
Weather	-0.25	0.19	0.22	0.00	-0.15	0.47	
Road surface condition	-0.19	0.17	0.21	0.00	-0.15	0.32	
Junction overshoot	0.15	0.52	0.24	0.00	0.17	0.47	
Poor turn or manoeuvre	-0.28	0.40	0.23	0.00	-0.33	0.33	
Passing too close to other vehicle	-0.12	0.69	0.23	0.00	-0.18	0.56	
Loss of control	0.24	0.71	0.22	0.00	0.50	0.44	
Constant					-3.86	0.00	0.25
					P-value of regional residual = 0.00		

Statistically significance at 95% confidence level

The result suggested that there was a statistically significant difference between the two countries with regard to multiple multilevel logistic regression (estimated variance of country=0.25 with p-value=0.00) and more specifically for variables speed limit and cyclist age. In addition, a consistent result emerged from the simple logistic multilevel regression which suggested that the statistically significant variance (with p-value=0.00) for countries was observed for all individual variables taken separately.

6. Discussion

A review of a wide range of previous studies available in the literature revealed that cyclist safety at roundabouts was an under researched area. This was unexpected given that government authorities are investing in schemes to promote sustainable transport and cycling is increasing year on year. An in-depth critique of the literature identified that whilst several studies had been made in different countries no effort has been made to compare across country differences in analysis of accident statistics and associated data concerning the circumstances of the accident and demographics of the victims. More specifically with an aim to give a richer understanding of cycle safety at a roundabout revealed by any differences that might emerge from the analysis.

Regarding the results of the paper, speed limit was statistically significant in UK analysis while it did not show any statistical influence on cyclist casualty severity for Belgium. Speed limits at UK roundabouts do matter because these roundabouts in some cases are designed in a way that vehicles more or less can drive through the roundabout at the speed of the posted limit. For radial roundabouts as in Belgium this is much less likely to be the case because – regardless of the posted speed limit – each vehicle has to slow down as the geometry is much tighter. An earlier study (Akgün *et al.*, 2018) concluded that speed limit can be considered as a proxy of geometric design parameters and it was stated that a higher speed limit (i.e. wider entry geometry) increases the probability of a serious cyclist casualty occurring at roundabouts with mixed traffic. This might help explain the difference between the influence of speed limit in the UK and Belgian case studies.

Furtado (2004) pointed out that determining one size for the safest balanced design is difficult to achieve as there are several design solutions applied in different regions. A main design feature of the UK priority roundabouts is to optimise for vehicle capacity. Allowing higher approach capacity results in a wider approach and entry geometry. Whilst delivering an increased level of service and less delay to traffic encourages higher entry speeds which in turn increases the number of cyclist collisions, (Davies *et al.*, 1997; Lawton *et al.*, 2003). The research presented in this manuscript clearly shows that the radial design, the basic principle for Belgium roundabouts, is safer. This is at the expense of loss of vehicle capacity. However, because tangential roundabouts are designed for increasing the capacity for motor vehicles, this suggests that a tighter geometry does not meet with the basic design principal of roundabout in the UK. On the other hand, in Belgium non-respect of priority is a significant contributor to increasing accidents, whilst in the UK the priority to the right is mandatory. Therefore, an important policy outcome of this research is when cycle *flows are high*, the wider geometry at tangential roundabouts to increase vehicle capacity should be carefully balanced with more focus towards a radial design with wider traffic management measures to reduce traffic flow at the junction. Notwithstanding, mandatory priority to the right, with associated driver education to ensure compliance should be given serious consideration. In addition, research should be conducted to identify the optimum balance of approach capacity and safety for cyclists as well as to understand the role of environmental issues. Novel roundabout designs which introduce a degree of segregation of cyclist flows should be explored, also.

Daniels *et al.* (2010) reported that for all road user types (i.e. including cyclists), higher age groups were more likely to be involved in severe casualties using Belgian data. A finding consistent with the results from our research also suggest that older cyclists are more likely to be involved in a *killed or seriously injured* collision compared to *slight* casualties not only in Belgium but also in the UK as well. This might be an expected result given that older population tend to have slower reaction times and lower physical ability. In addition, there was no evidence in either the Daniels (2010) research or this research here that the gender of the cyclist is associated with casualty severity.

The uniqueness of the research presented here is due consideration of the influence of behaviour-related contributory factors along with environmental and sociodemographic variables on cyclist casualty *severity* occurring at roundabouts therefore discussion is limited to comparing the outcomes of the behaviour related contributory factors. Rasanen and Summala (2000) suggested that driver's speed decreases yielding behaviour towards cyclists (Rasanen and Summala, 2000) and driver's behaviour has a strong impact on cyclist position at the roundabout (Silvano *et al.*, 2015). The result for UK demonstrated that there is a strong relationship between cyclist *casualty* and *sudden braking*, which increased the probability of *serious severity* consistent with the speed influence for tangential design. Interestingly, whilst *vehicle driver*-related contributory factors were not statistically significant the one cyclist behaviour-related predictor, '*cyclist's non respect of the priority*', did show an influence on casualty *severity* in a collision increased the likelihood to be *killed or serious injury* versus *non-killed or serious injury* with 171%. The *cyclist's 'non-respect of the priority'* could be intentional in which case steps for enforcement of the priority to the right (in the UK) rule at roundabouts should be stepped-up perhaps introducing media campaigns with on the spot fines for non-compliance of cyclists. On the other hand, cyclist's '*non-respect of the priority'* could be unintentional resulting from cyclist error and that might be due to lack of experience or unfamiliarity with the environment and the Highway Code (the rules of the road in the UK) etc. in which case education and training is more appropriate. Whilst the introduction of a formal roadworthy test for cyclists with a subsequent issue of a license, similar to that required of vehicle drivers, would be beneficial for cyclists the introduction of such regulation may deter the public from choosing cycling as a travel mode. However formal training could be developed along the lines of the current voluntary UK National Cycling Proficiency Test conducted in schools to ensure safety for children, but more advanced. Notwithstanding, further research needs to be conducted to investigate the level of support for mandatory cycling training, the content and format of a handbook, the level of acceptance of, and barriers presented by, the introduction of a licence to cycle on public roads.

The later results which used a three-way chi-square test of independence followed by multilevel logistic regression gave a deeper understanding of the differences between the UK and Belgium in the influences of behavioural and non—behavioural variables on the severity of accidents. Both simple and multiple multilevel logistic regressions showed that there were statistically significant differences between countries. Also, despite the need to use proxy data the different characteristics influencing thee predictors of casualty severity in the UK and Belgium were revealed.

From the chi square test, all non-behavioural variables such as cyclist *age group* and *gender*, *lighting*, *road surface condition* and more importantly *speed limit* were all found to be statistically different between the UK and Belgium for both slight and *killed and seriously injured* casualty severity. On the other hand of the three behavioural variables, only *'junction over-shoot'* emerged as having statistically significantly different influence on casualty severity in the two countries. This suggests that driver/cyclist interaction and behaviour in cities in the two countries is generally similar whilst speed limit, sociodemographic characteristics of cyclists and environmental conditions are specific for each country. Therefore, further research specifically into driver-rider behaviour influence on cyclist casualty severity in different countries would be useful. However, the possibility of including other variables besides particular geometric design, in the same model is limited due to the different protocols applied in recording details of accidents in the different countries.

7. Limitations

The statistically significant residual in the multilevel model for associated data of the UK and Belgium is expected to be due to the inconsistency in design protocol and the differences in the method of recording casualty data in each country suggesting the need for each country would require its own independent model. It is difficult to carry out a comparative study between regions with different design approaches, environment and data collection methodologies and this has been overcome in the research reported here by using proxies for some variables which in turn has given rise to a limitation of the current study. Ideally, comparative analysis should be carried out to include several countries to increase the statistical outcome of the regional residual. This may provide a more reliable generic model which is likely to include additional variables. But again, limitations in data availability, and non-standardisation of recording data remains an issue.

Studies that considered casualty analysis rely on police records as the data source. Relying on cyclist casualty records has limitations such as high degree of under-reporting (Laureshyn *et al.*, 2017). Hels and

Orozova-Bekkevold (2007) suggested that nearly 75% of cyclist casualty records in hospital are not registered in police collision databases. This suggests the possibility that there is a higher risk of cyclist collision at roundabouts and certainly researchers should be aware of potential bias caused by difficulties in some data collection methods. Therefore, achieving complete records for cyclists is the main challenge for safety studies. Considering this limitation, it was suggested (Laureshyn *et al.*, 2017) that field observations and measurements should be applied. However, these observations may not be possible always due to the lack of time and sources.

The study in this paper used very reliable police records in the UK (STATS19) and Belgium (VIAS), however, the authors are well aware of the limitation of unreported cyclist casualties. There is a tremendous attention on accuracy in STATS19 data collection because the accumulated data is available for public consumption by permission for the UK DfT, Department for Transport and widely used in scientific research. However, contributory factors started to be recorded based on a police officer's subjective assessment of whether they believe the specific factor has contributed to the collision or not (Rolison *et al.,* 2018). Efforts have been made to eliminate subjective recording by enhanced description needed in STATS21 with responsibility given to Local Authorities to ensure accuracy of records by with the need to create and maintain a database and applying automatic checks using validity protocols to expose errors in the database (TS, 2013). A specific issue in the current STATS19 data revealed in this research was that the contributory factor data was not assigned individually to the cyclist or driver. This created a limitation in the research reported here.

Traffic flow is an important parameter because at higher cyclist and driver flow the number of potential encounters increases. This study cyclist and driver flow at roundabouts could not be considered as they were not available at the location where the casualty was recorded. In addition, the study could not include actual speed of vehicles. Traffic related data would have improved the prediction models.

8. Conclusions

The key outcomes of this research are presented as follow:

1. For UK roundabouts, the higher the speed limit the greater the probability of a cyclist being killed or seriously injured in a collision compared to being slightly injured. The speed limit was not significant in explaining cyclist injury severity at Belgium roundabouts.
2. The older a cyclist is, the more likely they are to be seriously or fatally injured in a collision. The likelihood increases by 2% for each year of a cyclist's age.
3. Cyclists over the age of 41 in the UK and 48 in Belgium we are more likely to be a KS I casual tea than slight.
4. Gender and environmental conditions were not found to influence the severity of cyclist casualty at either UK or Belgium roundabouts.
5. Where sudden braking was recorded as a potential contributory factor, the likelihood of a cyclist being killed or seriously injured increases by about three times at a UK roundabout compared to a Belgium roundabout.
6. A cyclist casualty is three times more likely to be killed or seriously injured when 'non respect of priority' is a potential contributory factor at radial roundabouts. In the UK there is a mandatory give-way to the right at roundabouts.
7. In the comparative analysis of the UK with Belgium the regional residual was highly statistically significant. This can be attributed to the difference of design in two different countries, with tangential for the UK and radial for Belgium.

8. Whilst respecting the limitations of the different data recording protocols this research suggests, for cyclists, radial design can be considered safer than tangential roundabouts although at reduced capacity and improved safety for vehicles.
9. Harmonisation of the ways accident records are currently generated across countries in Europe with the inclusion of geometric design metrics would be valuable and worthwhile.

Acknowledgements

Accident data were obtained from CIRTAS-Gateshead Council and VIAS Institute Belgium. This study was supported by the Ministry of National Education, The Republic of Turkey.

References

Agresti, A., 2007. An introduction to categorical data analysis. 2nd Edition. Haboken, New Jersey: John Wiley and Sons, Inc.

Austroads, 2009. Guide to Road Design Part 4B Roundabouts. Australian Design Standards.

Akgün, N., Dissanayake, D., Thorpe, N., Bell, M.C., 2018. Cyclist casualty severity at roundabouts – To what extent do the geometric characteristics of roundabouts play a part? Journal of Safety Research 67, 83-91.

Akgün, N., 2019. Investigating the Impacts on Cyclist Casualty Severity at Give Way Roundabouts with Mixed Traffic, Newcastle University

Arnold, L.S., Flannery, A., Ledbetter, L., Bills, T., Jones, M.G., Ragland, D.R., Spautz, L., 2010. Identifying Factors that Determine Bicyclist and Pedestrian: Involved Collision Rates and Bicyclist and Pedestrian Demand at Multi-Lane Roundabouts. UC Berkeley Safe Transportation Research and Education Center, I.O.T.S. University of California, Berkeley.

Campisi, T., Akgün, N., Ticali, D., Tesoriere, G. 2020. Exploring Public Opinion on Personal Mobility Vehicle Use: A Case Study in Palermo, Italy. Sustainability, 12, 5460.

Daniels, S., Nuyts, E., Wets, G., 2008. The Effects of Roundabouts on Traffic Safety for Bicyclists: An Observational Study. Accident Analysis and Prevention 40 (2), 518-526.

Daniels, S., Brijs, T., Nuyts, E., Wets, G., 2009. Injury Crashes with Bicyclists at Roundabouts: Influence of Some Location Characteristics and the Design of Cycle Facilities. Journal of Safety Research 40 (2), 141-148.

Daniels, S., Brijs, T., Nuyts, E., Wets, G., 2010. Externality of Risk and Crash Severity at Roundabouts. Accident Analysis and Prevention 42 (6), 1966-73.

Daniels, S., Brijs, T., Nuyts, E., Wets, G., 2011. Extended Prediction Models for Crashes at Roundabouts. Safety Science 49 (2), 198-207.

Davies, D.G., Taylor, M.C., Ryley, T.J., Halliday, M.E., 1997. Cyclists at Roundabouts — the Effects of 'Continental' Design on Predicted Safety and Capacity. Transport Research Laboratory.

De Brabander, B., Vereeck, L., 2007. Safety Effects of Roundabouts in Flanders: Signal Type, speed Limits and Vulnerable Road Users. Accident Analysis and Prevention 39 (3), 591-599.

Efron, B., 2013. Empirical Bayes Modeling, Computation and Accuracy. Stanford University, USA.

Field, A., 2009. Discovering Statistics Using SPSS. 3rd Edition. SAGE Publications Ltd.

Furtado, G., 2004. Accommodating Vulnerable Road Users in Roundabout Design. Annual Conference of the Transportation. Canada, Quebec City.

Ghasemi, A., Zahediasl, A., 2012. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. International Journal of Endocrinal Metabolism 10 (2), 486-489.

Gross, F., Lyon, C., Persaud, B., Srinivasan, R., 2013. Safety Effectiveness of Converting Signalized Intersections to Roundabouts. Accident Analysis and Prevention 50, 234–241.

Harrell, F., 2001. Regression Modelling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. First Edition. Springer-Verlag New York.

Hels, T., Orozova-Bekkevold, I., 2007. The Effect of Roundabout Design Features on Cyclist Accident Rate. Accident Analysis and Prevention 39 (2), 300-307.

Jensen, S.U., 2017. Safe Roundabouts for Cyclists. Accident Analysis and Prevention 105, 30-37.

Laureshyn A, Goede MD, Saunier N, Fyhri A. 2017. Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in Norway. Accident Analysis & Prevention. 105:11-20. https://doi.org/10.1016/j.aap.2016.04.035

Lawton, B.J., Webb, P.J., Wall, G.T., Davies, D.G., 2003. Cyclists at Continental Style Roundabouts: Report on Four Trial Sites. Transport Research Laboratory, TRL report TRL584.

Li, B., Lingsma, H., Steyerberg, E., Lesaffre, E., 2011. Logistic Random Effects Regression Models: A Comparison of Statistical Packages for Binary and Ordinal Outcomes. BMC Medical Research Methodology, 11.

Møller, M., Hels, T., 2008. Cyclists' Perception of Risk in Roundabouts. Accident Analysis and Prevention. 40 (3), 1055-1062.

Montella, A., 2011. Identifying Crash Contributory Factors at Urban Roundabouts and Using Association Rules to Explore Their Relationships to Different Crash Types. Accident Analysis and Prevention 43 (4), 1451-1463.

Ogundimu, E.O., Altman, D.G., Collins, G.S., 2016. Adequate Sample Size for Developing Prediction Models is not Simply Related to Events Per Variable. Journal of Clinical Epidemiology 76, 175-182.

Parkin, J., 2018. Designing for Cycle Traffic: International Principles and Practice. ICE Publishing. ISBN-10: 9780727763495

Patterson, F., 2010. Cycling and Roundabouts: An Australian Perspective. Road and Transport Research: A Journal of Australian and New Zealand Research and Practice 19 (2), 4-19.

Peduzzi, P., Concato, J., Kemper, E., Holford, T.R., Feinstein, A.R., 1996. A Simulation Study of the Number of Events Per Variable in Logistic Regression Analysis. Journal of Clinical Epidemiology 49 (12), 1373-1379.

Polders, E., Daniels, S., Casters, W., Brijs, T., 2014. Identifying Crash Patterns on Roundabouts. Traffic Injury Prevention 16 (2), 202-207.

Rasanen, M., Summala, H., 2000. Car Drivers' Adjustments to Cyclists at Roundabouts. Transportation Human Factors 2 (1), 1-17.

Razali, N.M., Wah, Y.B., 2011. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. Journal of Statistical Modelling and Analytics 2 (1), 21-33.

Rolison, J.J., Regeva, S., Moutari, S., Feeney, A., 2018. What are the factors that contribute to road accidents? An assessment of law enforcement views, ordinary drivers' opinions, and road accident records. Accident Analysis and Prevention 115, 11-24.

Sakshaug, L., Laureshyn, A., Svensson, A., Hyden, C., 2010. Cyclists in Roundabouts: Different Design Solutions', Accident Analysis and Prevention, 42(4), pp. 1338-51.

Silva, A.B., Santos, S., Vasconcelos, L., Seco, Á., Silva, J.P., 2014. Driver Behavior Characterization in Roundabout Crossings. Transportation Research Procedia 3, 80-89.

Silvano, A.P., Ma, X., Koutsopoulos, H.N., 2015. When Do Drivers Yield to Cyclists at Unsignalized Roundabouts. Transportation Research Record: Journal of the Transportation Research Board 2520 DOI: 10.3141/2520-04.

Silvano, A.P., Linder, A., 2017. Traffic Safety for Cyclists in Roundabouts: Geometry, Traffic and Priority Rules. VTI Notat, no. 31A: 35.

Sommet, N., Morselli, D., 2017. Keep Calm and Learn Multilevel Logistic Modeling: A Simplified Three-Step Procedure Using Stata, R, Mplus, and SPSS. International Review of Social Psychology, 30 (1).

Sperandei, S., 2013. Understanding Logistic Regression. Croatian Society of Medical Biochemistry and Laboratory Medicine.

Steele, F., 2009. Multilevel Models for Binary Responses. Centre for Multilevel Modelling, Bristol.

Tollazzi, T., 2015. Alternative Types of Roundabouts: An Informational Guide. Springer International Publishing. DOI: 10.1007/978-3-319-09084-9

TS, 2013. Procedures for Submitting Road Accident Data to Transport Scotland from non-CRASH Sources Transport Scotland.