

**THE MAGNITUDE OF THE REGRESSION TO THE MEAN EFFECT IN TRAFFIC CRASHES**

Ellen De Pauw<sup>1</sup>, Stijn Daniels<sup>1</sup>, Tom Brijs<sup>1</sup>, Elke Hermans<sup>1</sup>, Geert Wets<sup>1\*</sup>

<sup>1</sup>Transportation Research Institute  
Hasselt University  
Wetenschapspark 5  
BE-3590 Diepenbeek  
Belgium  
Tel: +32112691{57, 56, 55, 41, 58}  
Fax.:+32(0)11 26 91 99

Email: {ellen.depauw, stijn.daniels, elke.hermans, tom.brijs, geert.wets}@uhasselt.be

\* Corresponding author

First submitted: August 1, 2012

Revised paper submitted: November 15, 2012

Number of words: 5635

Number of tables: 3

Number of figures:

→ Total number of words: 6385

**1 ABSTRACT**

2 Regression to the mean has been recognized as a phenomenon that influences road safety evaluations and should be  
3 accounted for. However, some doubts have risen about the necessity to implement rather sophisticated techniques  
4 such as the empirical Bayes method to correct for regression to the mean whereas the use of a sufficient long before-  
5 period could reach the same objective. Present study examines the existence and the magnitude of the regression to  
6 the mean effect in crash data from 169 intersections for the injury crashes and 143 for the more severe crashes in  
7 Flanders-Belgium for whom regression to the mean was likely to occur as they were selected based on their crash  
8 history. The presence of a RTM-effect was investigated by comparing the crash numbers of this period with the  
9 crash numbers in the next three years, during which no traffic safety measure was applied. Two comparison groups  
10 were used. The results demonstrate the existence of a substantial regression to the mean effect in the investigated  
11 sample of intersections. The magnitude of the regression to the mean effect is estimated to be almost 9% for injury  
12 crashes and 37% for the most severe crashes. From this can be concluded the correction for regression to the mean  
13 in evaluation studies is highly recommended in cases when locations are selected based on their crash history. This  
14 can be applied through the Empirical Bayes method.

15  
16 Key words: regression to the mean, empirical Bayes, before-after evaluation

## 1 INTRODUCTION

2 To implement an effective traffic safety policy, there is an undisputed need to evaluate the effectiveness of traffic  
3 safety measures. Well-controlled before-and-after (B&A) studies are widely considered to be the most appropriate  
4 method to execute such evaluation (1, 2). Different B&A methods are available, which mainly differ in the extent  
5 they control for possible confounding variables. A confounding variable is defined as any exogenous variable  
6 affecting the number of crashes or injuries whose effects, if not estimated, can be mixed up with the measure being  
7 evaluated (1). Variables that are regarded as potentially confounding in observational B&A studies of road safety  
8 measures are regression to the mean, long term trends affecting the number of crashes or injured road users, general  
9 changes in the number of crashes, changes in traffic volumes and any other specific events introduced at the same  
10 time as the road safety measure (1, 3, 4). Regression to the mean (RTM) is defined as one of the most important  
11 confounding variables (3). Elvik and Vaa (6) defined it as followed: "Regression-to-the-mean denotes the tendency  
12 for an abnormally high number of accidents to return to values closer to the long term mean; conversely abnormally  
13 low numbers of accidents tend to be succeeded by higher numbers. RTM occurs as a result of random fluctuation in  
14 the recorded number of accidents around the long-term expected number of accidents". As often the decision to  
15 implement a measure is based on a high crash rate during a relatively short period (e.g. 1 year), it is plausible that  
16 the number of crashes will decrease afterwards, irrespective of the measure. In those cases RTM will lead to an  
17 overestimation of the treatment effectiveness, when not appropriately taken into account.

## 18 19 BACKGROUND

20 Several authors analyzed the RTM phenomenon when examining the effectiveness of traffic safety measures, others  
21 compared the results of studies that controlled for RTM with studies that did not controlled. Based on several real-  
22 world data sets, Hauer and Persaud (7) found at locations selected because of their poor accident history, simple  
23 B&A comparisons tend to overestimate the safety effect of a measure, due to the RTM phenomenon. Elvik (1)  
24 compared studies that did not controlled for RTM, with studies that did controlled. This analysis showed the  
25 magnitude of RTM was estimated to be 20-30% in three out of seven cases, but was of a negligible magnitude in the  
26 other four. Harwood et al. (8) applied three B&A methods to evaluate the safety-effects of the installation of added  
27 left-turn lanes, right-turn lanes and the extension of the length of existing left- or right-turn lanes. The Empirical  
28 Bayes (EB) method, that controls for RTM, resulted in lower effectiveness estimates, compared to the yoked  
29 comparison and the comparison group approach, that did not controlled for RTM. The authors ascribed this  
30 difference to the RTM phenomenon. In a study that analyzed the effectiveness of speed and red light cameras (9),  
31 the authors also examined the RTM effects on a subset of 216 cameras, through application of the EB method with  
32 an external crash model. This study showed that RTM only had a modest effect on all injury crashes, that is an  
33 average fall of 7%, but had an appreciable effect on fatal and serious injury crashes of 35%. Compared to the total  
34 decrease, the RTM represented one quarter of the observed decrease in injury crashes, and three fifths of the  
35 observed decrease in more severe crashes. A meta-analysis that examined the effectiveness of red light cameras (10)  
36 showed that studies that controlled for RTM and spillover effects, yielded less favorable results, compared to studies  
37 that did not controlled these confounding variables. The study concluded that a failure to control for RTM can lead  
38 to an overestimation of the effect of RLC (10). Sharma & Datta (11) studied the RTM effect using four B&A  
39 evaluation methods (B&A, B&A with comparison group, 2 EB methods) through which the effectiveness of low-  
40 cost safety improvements were evaluated. From this study it was found that analyses that did not controlled for  
41 RTM, produced similar results to those that did controlled, when three to five years of crash rates were included.  
42 The authors concluded that the RTM effect becomes insignificant when three or more years of crash data are  
43 applied in the evaluation of high crash locations.

44 The occurrence of RTM in B&A studies is widely accepted. However, only few studies examined the  
45 magnitude of this RTM effect. Furthermore doubt exist concerning the necessity to use sophisticated methods, from  
46 which the EB method is stated as one of the most appropriate, whereas others stated the use of many years of pre-  
47 treatment data is sufficient to control for RTM (11).

## 48 49 STUDY DESIGN

50 In order to analyze the RTM effect, a dataset of locations that knew a higher than average number of crash counts  
51 was selected. Using a B&A design, these crash rates were compared with the crash rates in the years after, during  
52 which no traffic safety measure was implemented. Three B&A methods were applied, which differ in the extent  
53 they control for trend and RTM. The magnitude of the RTM effect was examined through a comparison of the  
54 results using these three B&A methods.  
55

## 1 DATA

2 To measure the RTM effect, a subset of the data set from the black spot program in Flanders, Belgium was used. To  
 3 increase traffic safety, the government decided in 2002 to select the most dangerous traffic spots in Flanders and to  
 4 redesign these locations. A location was selected as 'dangerous' when during the years 1997-1999 at least three  
 5 injury crashes occurred, and a total severity score of 15 was reached. For this score the number of injured were  
 6 taken into account: every slightly injured (defined as every person that got injured in a traffic crash, but cannot be  
 7 defined as severely or deadly injured) got a weight of 1, every severely injured (every person that needed more than  
 8 24 hours of hospitalization as a consequence of a crash) got 3 and every deadly injured (every person that died as a  
 9 consequence of a traffic crash on the spot or within 30 days after the crash) got a score of 5. A total priority score of  
 10 minimum 15 was necessary to be selected as a dangerous spot. A total of 1014 black spots were selected, of which  
 11 the 800 most dangerous were planned to be adapted. Assuming the existence of a RTM phenomenon, it can be  
 12 expected that locations with this high number of crashes will return to a more average number in the years after,  
 13 even without the implementation of a traffic safety measure. However, generally the RTM-effect cannot effectively  
 14 be isolated from other effects since several phenomena (e.g. trend effects, physical measures, changes in traffic  
 15 volume, random fluctuations, RTM) occur more or less simultaneously. The used dataset was particularly suitable to  
 16 examine the RTM effect as the locations were based on their crash history in the period 1997-1999 and the  
 17 adaptations of the locations only started in 2003. Between the period that was used to select the black spots and the  
 18 starting point of the first adaptations, also called as the lag period (13), no traffic safety measures were implemented  
 19 on these locations. This period (2000-2002) could subsequently be used, to examine whether RTM occurred.

20 To analyze the RTM phenomenon, a subset of the black spots with the highest frequency of crashes was  
 21 selected. A first selection was made for the injury crashes, for which the 200 spots with the highest number of injury  
 22 crashes in a radius of 100meter during 1997-1999 were selected. Eventually 169 sites were included in the final  
 23 research group. Thirty-one spots could not be included, as there was not enough information about the traffic  
 24 volume data, which was necessary to apply the empirical Bayes method (see further). In order to analyze the RTM-  
 25 effect in severe crashes, spots with 3 or more crashes with serious injuries or fatalities that occurred during 1997-  
 26 1999 were selected, which led to a selection of 180 spots. The final research group for severe crashes consisted  
 27 of 143 locations, as for 37 spots not all necessary data was present.

28 As stated above, different variables are defined as confounding in B&A studies, from which RTM is one  
 29 (3, 14). Although no measures were implemented during the research period, it was important to control for all  
 30 those factors, to make sure no false conclusions were made about the magnitude of the RTM. One of these  
 31 confounding variables is the general trend in the number of crashes. To account for this variable, a comparison  
 32 group can be used (1). Since all the sites from the research group are located at intersections with a regional road, all  
 33 crashes that occurred in Flanders at intersections with a regional road were selected (=comparison group 1). As a  
 34 check for robustness, also a second comparison group is used, which encompasses all crashes in Flanders  
 35 (=comparison group 2). A second confounding variable are changes in traffic volumes. According to Elvik (1) there  
 36 is no need to control explicitly for traffic growth. He stated it is sufficient to include a large comparison group, with  
 37 several hundreds of crashes. Such a comparison group includes the effects of all factors that may produce changes  
 38 over time in the long-term expected number of crashes. In such cases there is no need to estimate the effects of  
 39 changes in traffic volume statistically, and doing so carries a great risk of double counting (1). As both comparison  
 40 groups in this article encompass national wide crashes, which counts for thousands of injury crashes and hundreds  
 41 of severe crashes, it can be concluded that the comparison groups control for general changes in traffic volumes. A  
 42 third confounding variable is chance, which is taken into account through the application of a large comparison  
 43 group, several years of B&A crash data and the application of point estimates and confidence intervals. Fourthly,  
 44 also the introduction of any other event introduced during the study period can be defined as a confounding factors.  
 45 This was not the case in this study as all the locations remained the same during the total research period, and traffic  
 46 safety measures that were introduced, were more widely implemented for the total of Flanders. These were  
 47 subsequently taken into account through the comparison groups.

48 To examine whether the comparison group is comparable to the research group, the odds ratio (OR) for the  
 49 crash rates during the before period (1997-1999) can be estimated (15).

$$50 \quad OR = \frac{R_t/R_{t-1}}{C_t/C_{t-1}} \quad [1]$$

51  $R_t$  = number of crashes in the research group in year t

52  $R_{t-1}$  = number of crashes in the research group in year t-1

53  $C_t$  = number of crashes in the comparison group in year t

54  $C_{t-1}$  = number of crashes in the comparison group in year t-1

1 An OR near to one indicates the comparison group is comparable to the research locations. Maximum standard  
 2 deviation should not be higher than 0.20. Those analyses are applied on all injury crashes as on crashes with only  
 3 severely or deadly injured. The results in table 1 show for comparison group 1 and 2, both for all injury crashes as  
 4 for the severe crashes, the odds ratios are near to 1, and the standard deviations do not exceed 0.20. We can  
 5 conclude these are appropriate comparison groups.

7 **TABLE 1: Odds Ratio (OR) and Standard Deviation per Year, both for Injury and More Severe Crashes**

	Comparison group 1		Comparison group 2	
	Injury crashes	Severe crashes	Injury crashes	Severe crashes
	OR (s)	OR (s)	OR (s)	OR (s)
97-98	0.95 (0.05)	0.95 (0.11)	0.89 (0.05)	0.87 (0.10)
98-99	1.09 (0.05)	0.97 (0.11)	1.08 (0.05)	1.02 (0.11)
Average	1.02 (0.05)	0.96 (0.11)	0.98 (0.05)	0.94 (0.11)

8  
 9 **METHOD**  
 10 To calculate the magnitude of the RTM effect, the crash rates from the after period are compared with the crash  
 11 rates from the before period, using three different methods. This comparison can be expressed by an index, similar  
 12 to the index of effectiveness (3), however slightly different here as no measure was applied, and thus it cannot be  
 13 referred to as 'effectiveness'. Therefore 'index of difference' is a better expression. The proportion of the number of  
 14 crashes before to after can be expressed by a relative count that indicates how much of the crashes remain during the  
 15 after period:

$$16 \text{ Index of difference} = \frac{\text{counted number of crashes during the after period}}{\text{counted number of crashes during the before period}} \quad [2]$$

17 When the index is equal to one, this indicates there was no difference in the number of crashes. An index lower than  
 18 one shows that the number of crashes decreased. An index higher than one designate an increase in the number of  
 19 crashes after, compared to before. In order to examine the magnitude of the RTM effect, three different methods are  
 20 used to calculate the index of difference, which differ in the extent they control for confounding variables.

21 At first the naïve B&A method was applied. This method shows how the number of crashes has evolved  
 22 over the years, without control for any confounding variable. However, even when no measure is applied, it can be  
 23 expected that the crash rates will change according to the crash trend. To examine which part of the change is due to  
 24 general crash trend, a second method is applied, which compares the crashes rates after with before, taking trend  
 25 effects into account through a comparison group. As no traffic safety measure were applied at the research  
 26 locations, it could be expected that controlling the number of crashes for trend would lead to the same results as  
 27 recorded during the before period. When however still a decrease in the crash rates will be found, even when trend  
 28 is accounted, this difference can be ascribed to RTM. To confirm these results, a third method was applied, which is  
 29 the EB method. This method calculates the evolution of crashes and controls, next to trend and chance effects, also  
 30 for RTM effects (1, 3). When the application of this method leads to a non-significant result, this is a prove of the  
 31 existence of RTM at the selected sites. Comparing the results of the application of those three methods on the  
 32 dataset could reveal what was the difference between the results of the analyses for which no confounding variables  
 33 were controlled, only trend was taken into account and both trend and RTM was controlled. Those three methods  
 34 were applied on injury and severe crashes, for which three years of before ('97-'99) and after crash data ('00-'02)  
 35 were taken into account.

### 37 Control for RTM

38 To control for RTM, Hauer et al. (16) give next equation:

$$39 L_{\text{RTM, before}} = w * (\mu_{\text{C, before}} * km_L * T_{\text{before}}) + (1-w) * (\sum_{t=1}^{\text{tbefore}} L_t) \quad [3]$$

40  
 41 With:

42  $L_{\text{RTM, before}}$  = estimated number of crashes during the before-period in location L after correction for regression to the  
 43 mean

- 1  $w$ = the importance of the comparison group  
 2  $\mu_{C,before}$ = average number of crashes per year in the comparison group during the before period  
 3  $km_L$ = length of location L in km (intersections get a length of 1km)  
 4  $T_{before}$ = length of period before the measure  
 5  $L_t$ = number of crashes at location L in year t  
 6  $1-w$ = the importance of location L

7  
 8 To calculate the average number of crashes in the comparison group ( $\mu_C$ ), two possibilities can be chosen: a  
 9 comparison group that consist of crash data from comparable locations or a model that estimates the crash rates,  
 10 based on the characteristics of comparable locations. In this study a model is used, which is also based on the  
 11 Flemish black spot dataset (17). The dependent value of the model was the number of crashes, based on crash rates  
 12 of 2000-2003. This period is located after the selection of the black spots, for which crash rates from 1997-1999  
 13 were taken into account, and through which the effect of RTM is excluded. Furthermore this period is located before  
 14 measures were implemented, as the first locations were adapted in 2004. The model was based on locations with  
 15 varying types of intersection (e.g. both signalized as priority controlled, 3leg and 4leg, urban and rural setting). In  
 16 this research a model is used which estimates the number of injury and severe crashes through traffic volumes at  
 17 major and minor roads of the intersection:

$$18 \quad E_{injury}(\lambda) = e^{-1.7131} Q_{Maj}^{0.3231} Q_{Min}^{0.2463} \quad [4]$$

$$19 \quad E_{severe}(\lambda) = e^{-3.2138} Q_{Maj}^{0.3327} Q_{Min}^{0.2009} \quad [5]$$

20

21 Where

22  $E(\lambda)$  = expected annual number of crashes (dependent variable), with  $E_{injury}$  are all crashes with injured, and

23  $E_{severe}$  are all crashes with severe injuries and fatalities

24  $Q_{maj}$  = traffic volume at major road

25  $Q_{min}$  = traffic volume at minor road

26  $e$  = natural logarithm = 2.718

27 A more detailed description of the model can be found in De Ceunynck et al. (17).

28

29

30 The weight ( $w$ ) can be calculated through next equation:

$$31 \quad w = \frac{1}{1 + (\mu_{C,before} * T) / k_{before}} \quad [6]$$

32

33 With  $k$  is an over dispersion parameter per unit of length (18, 19), which is also calculated through the model.

34

### 35 Control for Crash Trend

36 Applying control for trend through a comparison group, can be written as an odds ratio:

$$37 \quad \frac{L_{after} / L_{before}}{C_{after} / C_{before}} \quad [7]$$

38 With:

39  $L_{after}$ = number of crashes on location L during the after period

40  $L_{before}$ = number of crashes on location L during the before period (for naïve B&A and B&A with comparison group)  
 41 or estimated number of crashes during the before period after control for RTM (for EB method)

42  $C_{after}$ = number of crashes in the comparison group during the after period

43  $C_{before}$ = number of crashes in the comparison group during the before period

44

45 With a 95% confidence interval (CI):

46 EFF, below limit =  $\exp[\ln(\text{EFF}) - 1.96 * s]$

47 EFF, above limit =  $\exp[\ln(\text{EFF}) + 1.96 * s]$

[8]

48 And a standard deviation ( $s$ ) as the root of the variance ( $s^2$ ):

$$s_L^2 = \frac{1}{L_{after}} + \frac{1}{L_{before}} + \frac{1}{C_{after}} + \frac{1}{C_{before}} \quad [9]$$

2

### 3 Meta-Analysis

4 Next to the individual evaluation per location, the total effect across different locations can be calculated by means  
5 of a fixed effects meta-analysis, which results in one overall effect estimate and in more statistically reliable results  
6 (15). Every location within the meta-analysis gets a weight, which is the inverted value of the variance.

7 Subsequently locations at which many crashes occurred, are given a higher weight.

$$8 \quad w_l = \frac{1}{s_l^2} \quad [10]$$

9 Supposing that the measure is executed at n different places, the weighted mean index of difference of the measure  
10 over all places is:

$$11 \quad \text{Overall index of difference} = \exp \left[ \frac{\sum_{l=1}^n w_l \cdot \ln(EFFI)}{\sum_{l=1}^n w_l} \right] \quad [11]$$

12 The estimation of a 95% confidence interval is:

$$13 \quad 95\% \text{ CI} = \exp \left[ \frac{\sum_{l=1}^n w_l \cdot \ln(EFFI)}{\sum_{l=1}^n w_l} \pm 1.96 * \frac{1}{\sqrt{\sum_{l=1}^n w_l}} \right] \quad [12]$$

14

## 15 RESULTS

16 At first all three methods (naïve B&A, B&A with comparison group, EB method) were applied on every black spot,  
17 separately for injury crashes and more severe crashes, taking three years of B&A data into account. Afterwards an  
18 overall estimation for all locations (169 for injury crashes, 143 for severe crashes) was executed with each of the  
19 three methods. The results are shown in table 2. For the injury crashes a significant decrease from before to after of  
20 13% in the number of crashes was found, using the naïve method. This is a quite high decrease as no specific  
21 measures were applied at these locations. As a consequence those decreases are attributable to confounding  
22 variables, such as trend effects and RTM. When trend and chance are taken into account, using the B&A method  
23 with comparison group 1 (all crashes at intersections at regional roads), the results show an index that is still lower  
24 than 1, but higher compared to the results of the naïve method. This indicates that the decrease from before to after  
25 is lower when trend is controlled, and subsequently that trend accounts for a part of the recorded decrease in injury  
26 crashes. However, even with this correction a nearly significant decrease of 5% was found. Using comparison group  
27 2, that encompasses all crashes in Flanders, a decrease of 9%, significant at the 5% level was found. As the most  
28 important confounding variables are controlled and no measure was implemented, this difference can subsequently  
29 be attributed to the RTM effect. To confirm this statement, the EB method is applied. When indeed RTM occurred it  
30 can be expected that analyses using the EB method -since it corrects for RTM- will show non-significant changes in  
31 the number of crashes. As from table 2 can be seen, the application of the EB method resulted in a non-significant  
32 index of 1.03 with comparison group 1, and 0.99 with comparison group 2. The results using the EB method  
33 indicate that there is no significant difference between the recorded number of crashes after and the estimated  
34 number of crashes before, when trend and RTM are controlled. This confirms the presence of RTM and the  
35 suitability of the EB method to correct for its existence.

36 For the severe crashes, the naïve B&A method shows a high decrease in crash rates of 36%. When the  
37 trend is taken into account, this resulted in a higher index of difference. However, still a significant decrease of 21%  
38 with comparison group 1 and 24% with comparison group 2 is found. From this can be concluded that RTM is also  
39 present for severe crashes. The application of the EB method, which found a non-significant increase of 6 to 11%,  
40 confirms this statement.

41

42

43

44

45

46

47

1 **TABLE 2: Results (Index of difference, [95% Confidence Interval])**

	Naïve B&A	B&A with comparison group		EB method	
	Index [95% CI]	Index [95% CI]		Index [95%CI]	
		Comparison group 1	Comparison group 2	Comparison group 1	Comparison group 2
Injury crashes	0.87 [0.82; 0.92]	0.95 [0.90; 1.01]	0.91 [0.86; 0.97]	1.03 [0.97; 1.09]	0.99 [0.93; 1.04]
Severe crashes	0.64 [0.56; 0.74]	0.79 [0.68; 0.91]	0.76 [0.65; 0.88]	1.11 [0.95; 1.30]	1.06 [0.91; 1.24]

2  
3 Table 3 gives a more detailed analysis of the magnitude of the RTM effect. This table shows the  
4 proportional change in the crash rates from before to after (see column 3), and the proportional change in crash rates  
5 from before to after, attributable to the RTM phenomenon (column 4). For these calculations the results with  
6 comparison group 1 are used, as it can be expected that crashes at intersections at regional roads are more similar  
7 with the crashes at the research locations, compared to all crashes in Flanders, which encompass for example also  
8 crashes at road sections. As the RTM effect is a group phenomenon, it is not reasonable to compute the deviation for  
9 each location separately (11). Therefore those results were grouped into various ranges of crash frequencies from  
10 1997-1999 (see column 1), for which an average value for the group in that range was calculated. In the last row the  
11 average over all locations is calculated. The results show, both for the injury and the more severe crashes, the  
12 general change in the number of crashes is negative, which indicates a decrease in crashes. Also the changes  
13 attributable to RTM is negative, for which the magnitude ranges from 1.5 to 17.7%, with an average of 8.7% for all  
14 injury crashes. For the severe crashes a decrease of 20.6 to 54%, with an average of 36.9% is found. Based on the  
15 total change in crash rates and the change attributable to RTM, column 5 shows which part of the observed decrease  
16 is represented by the RTM effect. From this can be seen that, the part of the observed decrease attributable to RTM  
17 raises proportionally with the mean number of crashes from the before period (as shown in the first column). For the  
18 injury crashes 18 to 66%, with an average of 51% of the observed decrease from before to after can be ascribed to  
19 the RTM phenomenon. For the severe crashes this ranges from 59 to 81%, with an average of 74%. From these  
20 results can subsequently be concluded the RTM effect for severe crashes is higher compared to all injury crashes.

21 **TABLE 3: Detailed Analysis of the Magnitude of the RTM Effect, Using Comparison Group 1**

Column 1	Column 2	Column 3	Column 4	Column 5
Mean nr crashes in before period	Nr of locations in that group	Total change in crash rates (%)	Change in crash rates attributable to RTM (%)	Part of the observed decrease attributable to RTM (%)
<b>All injury crashes</b>				
<b>0-3.9</b>	<b>26</b>	-8.1	-1.5	18.0
<b>4-4.9</b>	<b>65</b>	-14.9	-6.1	41.1
<b>5-5.9</b>	<b>39</b>	-13.8	-5.5	40.2
<b>6-6.9</b>	<b>15</b>	-23.1	-14.6	63.4
<b>7-7.9</b>	<b>11</b>	-15.7	-6.8	43.1
<b>≥8</b>	<b>12</b>	-26.9	-17.7	65.9
<b>Average</b>		-17.1	-8.7	51.0
<b>Severe crashes</b>				
<b>0-1</b>	<b>74</b>	-34.9	-20.6	59.1
<b>1.01-1.5</b>	<b>30</b>	-38.8	-28.3	72.9
<b>1.51-2</b>	<b>24</b>	-59.4	-44.6	75.1
<b>≥2</b>	<b>15</b>	-66.5	-54.0	81.2
<b>Average</b>		-49.9	-36.9	73.9

22  
23 **DISCUSSION**

24 Present study shows that the number of crashes at locations that were selected based on their high crash counts  
25 returns to a more average number during the years after, even when no measures were applied and possible other



1 influencing factors were eliminated. This decrease can subsequently be attributed to the RTM effect. The magnitude  
2 of this effect is estimated to be almost 9% for all injury crashes, which represents more than half of the observed  
3 decrease from before to after. For severe crashes a decrease of 37%, attributable to RTM is found, which count for  
4 almost three fourths of the observed decrease. Even when three years of B&A data are taken into account, high  
5 decreases can be found. This is in contrast with the study of Sharma and Datta (11), who stated that the RTM effect  
6 becomes insignificant when three or more years of crash data are applied in the evaluation of high crash locations.  
7 Furthermore it can be concluded that the RTM effect is higher for the more severe crashes compared to all injury  
8 crashes. This is in line with an earlier study (9), which showed that RTM only had a modest effect on all injury  
9 crashes, but had an appreciable effect on fatal and serious injury crashes.

10 A problem that especially needs attention in current study, and more generally in all B&A studies that  
11 examine the effectiveness of traffic safety measures, is the occurrence of zero crashes in the before or after period.  
12 In current research, no zero counts appeared for all injury crashes, but for the severe crashes 33 locations had zero  
13 counts during the after period. Elvik (21) described these zero counts as a problem for three reasons. He stated it is  
14 highly implausible that the true long-term mean number of accidents at any location is zero. Secondly zero counts  
15 suggest that a safety treatment could be either a hundred percent crash reduction (if there was a positive count  
16 before and a zero count after) or an infinite increase in the number of crashes (if there was a zero count before and a  
17 positive count after), both of which are highly implausible. Thirdly, zero counts have to be adjusted when in a meta-  
18 analysis a statistical weight is to be assigned to each result. A conventional method to solve this problem, is to add  
19 0,5 to each of the four variables, as shown in equation [7]. However previous research indicated that applying such a  
20 factor can lead to deviant results in meta-analyses (20). Subsequently it is an important initiative to explore this  
21 problem more thoroughly in future.

22 Summarized we can conclude it is of high importance that RTM is taken into account in future B&A  
23 studies that examine the effectiveness of a traffic safety measure. If not done so, the effectiveness is likely to be  
24 overestimated at locations with a higher than average crash count during the before period. Controlling for RTM can  
25 be done through application of the EB method. This method controls for different confounding variables, including  
26 RTM. However, this method had some criticism, as skeptics stated that the sophistication and data that are needed  
27 to execute this method is not worth the effort. They argued that alternative, less complex methods can produce  
28 equally valid results (12). Besides these weaknesses and criticisms, the EB method is stated as a method that does  
29 produce results that are substantially different and more valid than those generated by more traditional methods, and  
30 is indicated as the best to control for RTM (3, 12, 22). Different EB estimates are applicable, for which previous  
31 research found the EB-estimates based on crash prediction models showed the lowest prediction errors, and were  
32 defined as the state-of-the-art approach for observational B&A studies (22). Next to the EB method, also the full  
33 Bayes method can be used. For this method different advantages are defined, such as the possibility to use the  
34 method for quite complex model forms that are not easily handled in conventional generalized modeling  
35 approaches. Also it allows the estimation of valid models with smaller sample sizes. The disadvantage however is  
36 that the methodology is quite complex and may require a high level of statistical training (12). Next to the EB  
37 method, some other techniques can be defined. An ideal way of avoiding the impact of RTM is selecting locations  
38 on a random manner from amongst a specified set of potential sites, with those locations that are not selected then  
39 acting as controls. However this is rarely possible, because of ethical reasons. Maher and Mountain (13) defined an  
40 alternative possibility, that is to postpone the implementation of the treatment for some years, after the locations are  
41 selected on the basis of their crash record in the before period. This is defined as the lag period. The lag period can  
42 be used as an unbiased estimate of the true crash rate before the treatment is applied, and instead of comparing the  
43 crashes after with before, the crashes could be compared from after with the lag period (13). However this will not  
44 be applicable in every situation as often the government or traffic safety organizations try to implement a measure  
45 as soon as possible and a delay can encounter opposition.

## 46 47 **CONCLUSIONS**

48 Based on these results, the following conclusions can be made:

- 49 • This study demonstrates the existence of a substantial regression to the mean effect in a sample of  
50 intersections that were particularly prone to this effect.
- 51 • The magnitude of the regression to the mean effect in the investigated sample is estimated to be almost 9%  
52 for crashes with at least a slight injured and 37% for severe crashes.
- 53 • The Empirical Bayes method effectively corrects for regression to the mean.

- 1
  - 2
- Correction for regression to the mean in evaluation studies is highly recommended in cases when locations are selected based on their crash history.

## REFERENCES

- (1) Elvik, R. The importance of confounding in observational before-and-after studies of road safety measures. *Accident Analysis & Prevention*, Vol. 34, 2002, pp. 631–635.
- (2) Shinar, D. *Traffic Safety and Human Behavior*. Elsevier, Oxford, 2007.
- (3) Hauer, E. *Observational before-after studies in road safety: estimating the effect of highway and traffic engineering measures on road safety*. Elsevier Science Inc., Tarrytown, N.Y., 1997.
- (4) Hirst, W.M., L. J. Mountain, and M. J. Maher. Sources of error in road safety scheme evaluation: a quantified comparison of current methods. *Accident Analysis & Prevention*, Vol. 36, 2004, pp. 705–715.
- (5) Tamburri, T.N., C. J. Hammer, J. C. Glennon, and A. Lew. Evaluation of minor improvements. *Highway Research Record*, Vol. 257, 1986, pp. 34–79.
- (6) Elvik, R., and T. Vaa. *The Handbook of Road Safety Measures*. Elsevier, 2004.
- (7) Hauer, E. and B. Persaud. Common bias in before-and-after accident comparisons and its elimination. In *Transportation Research Record: Journal of the Transportation Research Board*, No.905, Transportation Research Board of the National Academies, Washington, D.C, pp. 164–174, 1983.
- (8) Harwood, D.W, K. M. Bauer, I. B. Potts, D. J. Torbic, K. R. Richard, E. R. Kohlman, E. Hauer, and L. Elefteriadou. Safety effectiveness of intersection left- and right-turn lanes. In *Transportation Research Record: Journal of the Transportation Research Board*, No.1840, Transportation Research Board of the National Academies, Washington, D.C, 2003, pp. 131–139.
- (9) Gains, A., M. Nordstrom, B. G. Heydecker, and J. Shrewsbury. The national safety camera programme: Four-year evaluation report. PA Consulting group and University College, London, 2005.
- (10) Erke, A. Red light for red-light cameras?: A meta-analysis of the effects of red-light cameras on crashes. *Accident Analysis & Prevention*, Vol. 41, 2009, pp. 897–905.
- (11) Sharma, S.L., and T. K. Datta. Investigation of regression-to-mean effect in traffic safety evaluation methodologies. In *Transportation Research Record: Journal of the Transportation Research Board*, No.2019, Transportation Research Board of the National Academies, Washington, D.C, 2007, pp. 32–39.
- (12) Persaud, B. and C. Lyon. Empirical Bayes before-after safety studies: Lessons learned from two decades of experience and future directions. *Accident Analysis & Prevention*, Vol. 39, 2007, pp. 546–555.
- (13) Maher, M. and L. Mountain. The sensitivity of estimates of regression to the mean. *Accident Analysis & Prevention*, Vol. 41, 2009, pp. 861–868.
- (14) Elvik, R. Evaluations of road accident blackspot treatment: A case of the iron law of evaluation studies? *Accident Analysis & Prevention*, 1997, Vol. 29, pp. 191–199.
- (15) Fleiss, J.F. *Statistical Methods for Rates and Proportions*. John Wiley, New York, 1981.
- (16) Hauer, E., D. W. Harwood, F. M. Council, and M. S. Griffith. Estimating safety by the empirical Bayes method: a tutorial. *Transportation Research Record: Journal of the Transportation Research Board*, No.1784, Transportation Research Board of the National Academies, Washington, D.C, pp. 126–131, 2002.
- (17) De Ceunynck, T., S. Daniels, T. Brijs, E. Hermans, and G. Wets. *Explanatory models for crashes at high-risk locations in Flanders, Belgium*. Steunpunt Verkeersveiligheid, Diepenbeek, In revision.
- (18) Abbess, C., D. Jarrett, and C. C. Wright. Accidents at blackspots: estimating the effectiveness of remedial treatment, with special reference to the ‘regression-to-mean’ effect. *Traffic Engineering and Control*, Vol. 22, 1981, pp. 535–542.
- (19) Ogden, K.W. *Safer Roads: A Guide to Road Safety Engineering*. Avebury, Aldershot, 1996.
- (20) Sweeting, M.J., A. J. Sutton, and P. C. Lambert. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. *Statistics in Medicine*, Vol. 23, 2004, pp. 1351–1375.

- (21) Elvik R. Treatment of zero counts in before-and-after road safety evaluation studies: an exploratory study of continuity corrections. Presented at the 91<sup>st</sup> Annual Meeting of the Transportation Research Board, Washington, D.C., 2011.
- (22) Elvik, R. The predictive validity of empirical Bayes estimates of road safety. *Accident Analysis & Prevention*, vol. 40, 2008, pp. 1964–1969.