SIXTH FRAMEWORK PROGRAMME
Priority 1.6 Sustainable Development, Global Change and Ecosystem
1.6.2: Sustainable Surface Transport

Police Enforcement Policy and Programmes on European Roads

Contract No: 019744

Deliverable 9

Good practice in the selected key areas: Speeding, drink driving and seat belt wearing: Results from meta-analysis

Due date of deliverable: **30/04/2008**
Actual submission date: **19/09/2008**

Start date of the project: 1/3/2006 Duration: 30 months

Organisation name of lead contractor for this deliverable: **TØI**

Authors: Alena Erke, TØI,
Charles Goldenbeld, SWOV,
Truls Vaa, TØI

Final

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination level:		
PU	Public	**PU**
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

PEPPER Project Consortium

1	VTT Technical Research Centre of Finland	FI
2	4Sight, Ergonomics & Safety	ISR
3	IBSR-BIVV Institut Belge Pour La Sécurité Routière	BE
4	KLPD/TISPOL Korps Landelijke Politiediensten /	NL
5	bfu Schweizerische Beratungsstelle für Unfallverhütung	CH
6	CERTH/HIT Hellenic Institute of Transport	GR
7	BASt Bundesanstalt für Strassenwesen	D
8	CDV Transport Research Centre	CZ
9	DTU Department of Transport, Technical University of Denmark	DK
10	IBDIM Road and Bridge Research Institute	PL
11	INRETS Institut National de Recherche sur les Transports et leur Sécurité	FR
12	KfV Kuratorium für Verkehrssicherheit (Co-ordinator)	AT
13	SWOV Institute for Road Safety Research	NL
14	TØI Institute of Transport Economics	NO
15	TRL Limited	UK
16	UPM Universidad Politécnica de Madrid	ES
17	VTI Swedish National Road and Transport Research Institute	SE
18	ETSC European Transport Safety Council	INT

Abstract

The main objective of task 4.2 in the EU-project PEPPER on traffic law enforcement measures has been to give a systematic review of evaluation studies on speed, drink driving and seat-belt enforcement by applying meta-analyses to assess the best estimates of the effects of enforcement measures on accidents and behaviour. The report separates between stationary speed enforcement using laser/radar, mobile patrolling, composite police controls with stationary/visible elements: and speed cameras. The overall accident-reducing effect is 18 % (-23: -13). Of these, mobile patrolling, mobile/hidden speed cameras and stationary speed enforcement "American type" (same unit measures, pursue and sanction the violator) do not have statistically significant effects on reducing the number of accidents. Visible/fixed speed cameras reduce the number of accidents with -34 % (-25; -42) while stationary and visible speed enforcement show a tendency in reducing the number of accidents of 11%, however insignificant (-22: +1). Concerning drink driving enforcement a distinction between patrolling measures and DUI checkpoints is justified. The former exhibits a significant effect on accidents by -8 (-12; -3), the latter somewhat stronger by a reduction of -15% (-18; -11). Finally, a meta-analysis of seat belt enforcement shows a significant increase in wearing rates of + 21 (during-period) and + 15 % (after-period)

Contents

EXECUTIVE SUMMARY

The main objective of task 4.2 of the EU-project PEPPER has been to give a systematic review of evaluation studies on speeding, drink driving and seat belt wearing by applying meta-analysis to assess the best estimates of the effects of enforcement measures on accidents and behaviour. From previous experiences within this field of research it is known that the number of studies is very large, which calls for methods that can systematize the high number of results from evaluation studies. By using meta-analysis one is able to calculate best estimates of effects of any given enforcement measure by a weighted average across a large number of studies. Task 4.2 is done in cooperation with SWOV and has been led by TØI.

Categories of speed enforcement measures

The preparation for meta-analysis of speed enforcement measures draws upon the distinction between measures adopted in the review of literature in Elvik and Vaa (2004). At that time the following subgroups of speed enforcement measures were separated:

- Stationary speed enforcement
- Patrolling
- Automatic speed enforcement ("speed cameras")

A distinction should be made between speed enforcement using stationary methods and speed enforcement using mobile methods (patrolling) because time and distance halo effects have been found to be at work for stationary enforcement, but not for mobile patrols (Shinar and McKnight 1985; Vaa 1993). *Time halo effects* means that an effect can be found in a given period of time after enforcement have ended, while *distance halo effects* means that effects on speed have been found at a certain distance from the spot where the speed enforcement is carried out.

In task 4.2 a total of 45 evaluation studies on speed enforcement were identified and found to be of an acceptable quality to be included in the database. The 45 studies have been published in 14 countries and comprise a total of 129 results. USA, Australia, UK and Sweden are the ones which have had the largest numbers of results from speed enforcement evaluations.

The following methods of speed enforcement are distinguished:

- **Stationary speed enforcement using laser or radar** that measures speed from one, usually unobtrusive or hidden observation site, or instruments that measure mean speed between two fixed observation sites, and clearly visible apprehension sites staffed by uniformed police officers and marked cars.
- **Stationary radar enforcement "American type".** The police observer (sometimes one officer alone in a car) measures speed by a radar mounted on the window and then pursues offending vehicles straightaway in order to apprehend and sanction the speeding driver. This technique is also used in Australia and the observation unit may be overt (marked car) as well as covert (unmarked car).
- **Patrolling**: Mobile police patrols with uniformed cars or motorcycles.

- **Composite police controls with stationary and visible elements**: This term is used to illustrate that speed enforcement may utilize a whole range of different techniques and methods, but also that it comprises some element that is stationary and that some of the activity is visible to the drivers passing by.
- **Speed cameras:** Several deployment patterns can be distinguished: 1) Fixed speed cameras, most often visible, on fixed locations/poles or portals with a mobile camera moving around, or 2) Mobile cameras, less obtrusive or even hidden cameras used on different locations, and 3) section control of speed where the average speed between two fixed sites is calculated and enforced if the speed limit is violated.

The distribution according to year of publication shows that two thirds of the studies are quite recent as they have been published in the 1990s or after year 2000.

Results of meta-analysis of speed enforcement studies

Results from tests of heterogeneity and summary effects of the estimated effects of speed enforcement on accidents are shown in Table E-1. The overall result is a significant reduction of the number of accidents of 18 % (-23; -13). There are large differences between the estimated effects of the different types of speed enforcement. The trim-and-fill analyses indicate that the results for mobile speed cameras are affected by publication bias, but this is the only subgroup affected.

Table E-1. Summary effects of speed enforcement measures

	Test of heterogeneity			Change of number of accidents (%)	
	Cochran's Q	**df**	**p**	**Summary effect**	**95% confidence interval**
All measures	5307.82	128	0.000	**-18**	(-23; -13)
Stationary manual	1854.17	22	0.000	**-11**	(-22; +1)
Patrolling	62.7573	10	0.000	**-6**	(-16; +4)
Radar/laser US/AUS	22.3372	30	0.841	**-0**	(-3; 4)
Speed cameras (all types)	1693.9	42	0.000	**-30**	(-38; -23)
- Subgroup: Mobile speed cameras	168.476	12	0.000	**-17**	(-34; 4)
- Subgroup: Fixed speed cameras	1513.02	27	0.000	**-34**	(-42; -25)
Composite Other	454.306	20	0.000	**-18**	(-33; +1)

The following variables were chosen as potential predictor variables:

Visibility: The enforcement measures can roughly be classified as visible enforcement when police cars, measurement equipment or cameras are either made visible or at least not hidden from being viewed by the drivers.

Randomization*:* In most of the studies randomization procedures are not clearly specified. In a few studies it is stated that there has been some kind of randomization of time of observation and/or site of apprehension.

Publicity: Publicity campaign: The enforcement measure is part of a wider campaign. Local publicity: The enforcement measure is not part of a wider campaign, but accompanied by local publicity.

Intensity: The studies are classified as a) A new type of enforcement or b) Increased enforcement. The studies are classified according to the increase of the amount of enforcement: a) Less than doubling the amount of enforcement, b) Doubling or larger increase of the amount of enforcement.

Country: According to the availability of studies, the following countries are compared: Australia, Sweden, UK, USA and a rest group of “other countries”.

Accident severity and types of accidents: Effects are estimated for the following groups of accident severity: Fatal accidents and/or serious injury vs lower injury levels (including unspecified/PDO.

Study methodology: Based on the study methodology all studies are roughly classified as “good” or “weak” study designs. Weak study designs are studies without a comparison group and studies that have controlled for time trends only. Good study designs are studies that have used a control or comparison group.

There are significant amounts of heterogeneity in all results, except for radar / laser of the type used in USA and Australia. Comparing the summary effects between the subgroups shows that larger accident reductions have been found when:

- enforcement and signposted; effects are not different dependent on whether or not enforcement is visible;
- there is no randomization (all results with randomization refer to radar / laser US / AUS);
- there is local publicity, compared to a publicity campaign or no publicity;
- a new form of enforcement is introduced and when the intensity is increased by a large amount;
- in the UK compared to other countries;
- accidents are severe;
- in studies with a weak study design.

In meta-regression analysis it is investigated how the potential predictor variables affect the effectiveness of speed enforcement when controlling for all other variables in the multivariate model at the same time. In the meta-regression analysis all predictor variables are coded as dummy variables. The meta-regression analysis is conducted in four steps:

Step (1) Stepwise forward analysis: The predictor variables are included in the regression model successively, one at a time.

Step (2) Complete regression model with all predictor variables, based on all studies: The regression coefficients are estimated for the complete model which contains all predictor variables.

Step (3) Complete regression model with all predictor variables, based on the studies with a good design only.

Step (4) In a fourth step, coefficients are estimated for a partial model which is based on all results. In the partial model only the predictor variables are included that have been identified as relevant in the preceding steps.

Step (3) and (4) are conducted in order to provide a test of the consistency and robustness of the results. The results for step (1), (2) and (3) are shown in Table E-2. The results from step (4) are not shown in the table. Table E-2 shows how each of the predictor variables is dichotomized as a dummy variable with the values 0 and 1. The categories for which the larger accident reductions have been found are underlined. Table E-3 shows the model summaries of the regression models that have been estimated in step (2) and (3).

Table E-2: Effects of speed enforcement on accidents, results from meta regression.

		Step (1)	Step (2)		Step (3)	
		R^2 when predictor included into model	All results		Results from good study designs	
Predictor	Dummy variables (1; 0)		Coeff.	p(Coeff.)	Coeff.	p(Coeff.)
Type of enforcement	(Radar US/AUS; <u>other</u>)	0.096	0.224	0.050	0.142	0.026
	(<u>Speed cameras</u>; other)		-0.030	0.831	-0.036	0.674
	(<u>Composite</u>; other)		-0.090	0.442	-0.089	0.245
Signposted	(<u>Signposted</u>; other)	0.169	-0.231	0.130	-0.126	0.261
Accident severity.	(<u>Fatal/serious</u>; other)	0.207	-0.166	0.020	-0.166	0.001
Increase	(Incr.; <u>change of enf.</u>)	0.230	0.124	0.150	0.084	0.203
Method	(Good st.; <u>weak study</u>)	0.237	0.054	0.523		
Visibility	(<u>Visible</u>; other)	0.238	-0.133	0.117	-0.162	0.003
Country	(Australia; other)	0.240	-0.055	0.602	0.054	0.405
	(UK; other)		-0.051	0.679	0.162	0.050
	(US A; other)		-0.191	0.076	0.093	0.195
Publicity	(<u>Local publicity</u>; other)	0.230	-0.085	0.328	-0.004	0.949
	(<u>Pub. campaign</u>; other)		-0.028	0.807	-0.034	0.616

Table E-3: Model summaries of the results from meta-regression.

Model	df	F	p(F)	R^2	Tau^2
Step (2) complete model based on all results	115	4.01	0.000	0.234	0.075
Step (3) complete model based on good study designs only	77	4.55	0.000	0.324	0.013

The results from meta-regression are only partly in accordance with the results from the subgroup analyses. Differences can be explained with the control for other variables in the regression model. When the effect of a variable changes when other variables are controlled for, it is likely that this effect has been, at least partly, due to confounding effects of the other variables.

- **Type of speed enforcement**: The type of enforcement is related to the effectiveness of the enforcement. The American / Australian type of radar / laser enforcement is less effective than the other methods. This result is consistent between the subgroup and the meta-regression analysis. Speed cameras and composite enforcement are, according to the results from meta-regression, not different from other types of enforcement.

- **Signposting**: Speed enforcement has been found to be more effective when signposted than when not signposted. The regression coefficient is not significant in the complete model with all predictor variables. In the partial models (step 4) however, which are based only on those predictor variables that have been identified in the stepwise procedure (step 1), the coefficient is somewhat larger (-0.337) and significant (p = 0.002).

- **Accident severity**: The effects of speed enforcement have been found to be greater for more serious accidents. The coefficient is significant both when all studies are included in the study and when the regression analysis is based on good studies only. In the partial models (step 4) the result is similar as in the complete model (coefficient = -0.184; p = 0.007). The result is also consistent with the subgroup analysis.

- **Increase vs. change of enforcement**: The effectiveness of enforcement has been found to be greater when a changed form of enforcement is introduced, compared to an increase of the amount of enforcement. When the type of enforcement is changed the intensity (e.g. in terms of person hours) if often increased as well. This result is consistent between the subgroup analysis and the meta-regression analysis. The regression coefficients are however not significant. The increase in R^2 when including this predictor variable in the model in the stepwise procedure is smaller than for the previously introduced variables.

- **Method**: Greater effects of speed enforcement have been found in studies with a weak study design, compared to studies with a good study design. The regression coefficient is however not significant and the increase in R^2 when including this predictor variable in the model in the stepwise procedure is only small.

- **Visibility**: The visibility of speed enforcement has not been found to be relevant for the effectiveness of speed enforcement in the subgroup analysis. In the meta-regression analysis no significant effect has been found and R^2 is not improved when this variable is introduced in the model in the stepwise procedure. All the same, the visibility of enforcement may not be irrelevant since signposting has been fount to be a significant predictor for the effectiveness. Additionally, in the subgroup analysis, the effect of non-visible enforcement is far smaller than the effect of visible enforcement when it is controlled for publication bias.

- **Country**: The results for the effects of speed enforcement in different countries are inconsistent. In the meta-regression analysis the coefficients have the opposite sign in the model based on all studies and in the model based on studies with a good study design. It is also curious that all three dummy variables have the same sign, although the subgroup analysis would indicate that the effectiveness is smaller in Australia than in other countries and greater in the UK than in other countries. When adding the country dummy variables in the stepwise regression analysis, the predictive power of the model does not improve markedly. It is therefore concluded that the effectiveness of speed enforcement is not systematically different between different countries.

- **Publicity**: Publicity is, according to the results from meta-regression, not a relevant predictor variable. In the subgroup analysis local publicity has been found to increase the effectiveness of enforcement, while publicity campaigns have been found to decrease its effectiveness. All except one studies of enforcement that is accompanied by a publicity campaign have used a "good" study design, which has been found to reduce the effectiveness that is found in a study. This may explain the findings that publicity campaigns seem to reduce the effectiveness of enforcement, but that this effect disappears when study design is controlled for.

Drink driving enforcement measures

Most studies of the effects of DUI enforcement on accidents have investigated the effects of *DUI checkpoints*, and some studies have investigated the effects of *patrolling*.

Patrolling: The types of measures investigated range from a mere increase of the amount of patrolling to larger programmes where officers are trained in DUI apprehension and where other anti-DUI measures are implemented at the same time. Enforcement programmes including patrolling are STEP (Selective Traffic Enforcement Programmes) and ASAP (Alcohol Safety Action Projects). The measures have in common that most of them have been accompanied by publicity..

DUI-checkpoints: DUI-checkpoint refers to all police operations where one or more police cars are standing at the roadside (not driving) and where police officers pull out drivers in order to check whether or not a driver has an illegal BAC level. According to this definition, checkpoints may vary with respect to how large and how visible DUI-checkpoints are, to what degree DUI-checkpoints are conducted at random times or places or only on roads and at times with high frequencies of DUI accidents, if all drivers are stopped at the DUI-checkpoint as far as the capacity of the checkpoints allows, or if only some drivers are stopped. When not all drivers are stopped, drivers may be stopped randomly or on suspicion only. In many countries, there are DUI-checkpoint programmes where the checkpoints are combined with media campaigns. In many of these programmes, the checkpoint operations are conducted in a very specific way, e.g. with a specific type of cars or buses (e.g. "booze buses"), and known under specific names, e.g. Random Breath Testing (RBT, Australia) or Compulsory Breath Testing (CBT, New Zealand). Only the following variables have reported to an extent that allows further sub-group analyses:

- Testing of drivers who are stopped at the checkpoint
- Accompanying publicity
- New type of or intensified enforcement
- Country
- Injury severity and types of accidents

- Time period over which DUI-checkpoints are conducted and evaluated
- Study methodology

Describing categories of drink driving measures

Testing of drivers who are stopped at the checkpoint: Based on available information, all evaluation studies are classified according to whether or not all drivers who are stopped are tested.

- **All drivers tested**: At DUI-checkpoints where all drivers are tested either a passive sensor is used or a breath test is taken.
- **Not all drivers tested**: In some checkpoint programmes it is either not required that all drivers are tested, or tests are not taken from all drivers even if it may be required.

Accompanying publicity: DUI-Checkpoint programmes are classified according to the type of accompanying publicity:

- **Publicity campaign:** Publicity with paid media, i.e. advertising in newspaper or TV-spots (publicity campaign with paid media may also include non-paid publicity).
- **Unpaid publicity**: Publicity without paid media, e.g. information in regular newspaper or TV coverage, and posters at the road side.
- **No publicity**: No activities are conducted in order to achieve publicity for the DUI-checkpoints.

Change of type or amount of enforcement: The studies have been classified as either A) A new type of enforcement or B) Increased enforcement.

Country: According to the availability of studies, the following countries are compared: Australia, New Zealand, USA and the rest group of “others countries”.

Accident severity and types of accidents: The meta-analysis effects are estimated for the following groups of accident severity: *Unspecified severity*: Results where accident severity is not specified (these include most likely injury and property-damage-only accidents), and where it is specified that accident and property-damage-only accidents have been investigated. *Injury accidents*: Mostly including fatal accidents, and finally *Fatal accidents.*

Accidents involving alcohol: Many studies have aimed at investigating effects on accidents involving alcohol. Since precise information on the BAC of drivers involved in accidents is not always available, many studies have used some substitute measure for alcohol accidents. Mostly, weekend night accidents have been used as a substitute measure for accidents involving alcohol. In short, the studies are grouped as follows: A) Accidents involving alcohol, B) All accidents, and C) Daytime accidents.

Time period over which DUI-checkpoints are conducted and evaluated: Studies vary with respect to the study period after DUI-checkpoints have been introduced or after their intensity has been increased. Studies are classified as follows: A) 3 months, B) between 3 and 6 months, C) between 6 months and 1 year, D) between 1 and 2 years, E) between 2 and 4 years, F) between 4 and 8 years.

Study design: Several types of study design have been applied, but the studies of the present sample are roughly classified as “good” or “weak” study designs: *Weak study designs* are studies without a comparison group and studies that have controlled for time trends only. *Good study designs* are the remaining types of study design i.e. studies that have applied a comparison or control group.

DUI: Patrolling studies: Results of meta-analysis

A comparison between the results from good and weak study designs shows that a significant accident reduction only was found in studies with a weak design, not in studies with a good design. For injury/unspecified accidents, effect estimates are available for all types of accidents (day and night) and for night-time accidents. The rest of table E-4 shows the results that are based on studies with a good design only. For both groups of effect estimates there are significant amounts of heterogeneity and the summary effects are almost identical. For fatal accidents, no significant effect on accidents has been found.

Table E-4. Test of heterogeneity and summary effects (Random Effects models) of patrolling.

	Test of heterogeneity			Change of number of accidents (%)	
Type of accidents affected	**Cochran’s Q**	**df**	**p**	**Summary effect**	**95% confidence interval**
		All results			
All accidents	156.41	26	0.000	-8	(-12; -3)
		Good vs. weak study design			
All accidents, good study design	54.38	20	0.000	-4	(-8; +0)
All accidents, weak study design	2.680	5	0.749	-24	(-28; -20)
		Injury, fatal and night time accidents (good study designs only)			
Injury accidents / unspecified severity; all types of accidents	39.62	6	.000	-6	(-11; 0)
Injury accidents / unspecified severity; all types of accidents – outlier omitted[1]	25.98	5	.000	-7	(-14; -1)
Injury accidents / unspecified severity; night time accidents	22.97	4	.000	-9	(-17; 0)
Fatal accidents; all types of accidents	14.71	13	.325	-1	(-7; +5)

In summary, the results do not indicate that patrolling has had significant effects on accidents.

DUI-checkpoints studies: Results of meta-analysis

The results from the trim-and-fill analysis indicate that the results for DUI checkpoints are affected by publication bias. The results presented below are therefore also adjusted for publication bias. These show to which groups of injury severity, type of checkpoint, time period, and accompanying publicity the effect estimates refer. The overall result is a significant reduction of the number of accidents (table E-5). When the results are adjusted for publication bias in trim-and-fill analyses the summary effects become smaller, but all of them remain significant.

Table E-5. Test of heterogeneity and summary effects (Random Effects models) of DUI checkpoints.

	Test of heterogeneity			Change of number of accidents (%)		Trim-and-fill analysis: Change of number of accidents (%)		
	Cochran's Q	df	p	Summary effect	95% confidence interval	Summary effect	95% confidence interval	Number of new effect estimates
All results	1007.221	96	0.000	-17	(-20; -14)	-15	(-18; -11)	7

There are significant amounts of heterogeneity in all results. Moderator variables are therefore likely to be present.

DUI: Subgroup analyses

The effects of a number of potential moderator variables are investigated by comparing summary effects between subgroups of results. It has been found that the results are likely to be affected by publication bias. Comparing the summary effects between the subgroups of each of the potential moderator variables shows that *larger accident reductions* have been found when:

- shorter time periods are studied, i.e. the largest accident reductions are found during the first half year;
- in Australia compared to New Zealand and USA, and New Zealand and USA compared to other countries;
- in studies with a weak study design;
- accidents involving alcohol; this effect becomes even stronger when the results are controlled for publication bias;
- when not all drivers are tested at the checkpoints, compared to checkpoints where all drivers are tested, the difference is however not large;
- for injury accidents compared to fatal accidents, the difference becomes larger when the results are controlled for publication bias;
- when publicity involves paid media, compared to publicity with unpaid media only.

All results are significant, with the exception of the results for checkpoints in countries other than Australia, New Zealand and USA.

DUI: Results from meta-regression

In meta-regression analysis it is investigated how the potential moderator variables affect the effectiveness of DUI checkpoints when controlling for all other potential moderator variables at the same time. The meta-regression analysis is conducted in four steps.

Table E-6. Effects of DUI checkpoints on accidents, results from meta-regression.

Predictor	(1; 0)	Step (1) R^2 when predictor included into model	Step (2) All results		Step (3) Results from good study designs	
			Coeff.	p (Coeff)	Coeff.	p (Coeff.)
1. Time period	> 6 months (1) vs. < 6 months (0)	0.136	0.176	0.001	0.102	0.104
2. Country	Australia (1) vs. other countries (0)	0.225	-0.224	0.000	-0.180	0.003
3. Study design	Good design (1) vs. weak design (0)	0.277	0.120	0.009	-	-
4. Accident types	Involving alc. (1) vs. all accidents (0)	0.293	-0.080	0.052	-0.058	0.136
5. Testing of drivers	All tested (1) vs. not all tested (0)	0.311	-0.111	0.053	-0.085	0.171
6. Accident severity	Fatal accidents (1) vs. injury accidents (0)	0.314	-0.053	0.233	-0.061	0.159
7. Publicity	Paid media (1) vs. no paid media (0)	0.311	0.037	0.417	0.051	0.280

The first three variables that are included in the regression model are, in this order, time period studied, country and study design. All three variables have highly significant regression coefficients in the model that has been developed in step (2). The results are consistent with the findings from the subgroup analyses. The results for the *time period* that has been studied show that the largest accident reductions are found during the first half year. When the time period after introducing or changing a DUI programme is longer, the effectiveness seems to decrease. This does not mean that longer time periods reduce the effectiveness of checkpoints during the first half year after introduction, only that accident reductions are becoming smaller over time

The *country* in which DUI checkpoints have been studied also has a highly significant effect on the results. The largest accident reductions have been found in Australia. Two obvious differences between Australian and other DUI-checkpoint programmes are the use of highly visible “booze buses”, which were introduced in Australia in 1989 in the state of Victoria, and the large amount of publicity accompanying the DUI-checkpoint programmes in Australia. New Zealand is the only other country where booze buses are being used.

Study design is also a highly significant predictor for the effects of DUI-checkpoints, but study design is not related to any of the characteristics of DUI enforcement.

Accidents involving alcohol are according to the results in Table E-6 more strongly reduced than other accidents. Mostly, some substitute measure of alcohol accidents is used, such as weekend night accidents.

Testing of drivers seems to affect the effectiveness of DUI checkpoints. The negative regression coefficient indicates that checkpoints where all drivers are tested are more effective in reducing accidents. The last two variables that were included in the regression model are accident severity and accompanying publicity. None of these variables has a significant regression coefficient, and the amount of variance explained by the regression model does not increase noticeably when these variables are included in addition to the other predictor variables.

Accident severity and use of media: The results from the regression analysis do, however, not indicate that effects are significantly different between injury and fatal accidents. The results for accompanying publicity do not say anything about the effect of the amount or intensity of publicity, which it was not possible to classify in a consistent way for all studies.

Seat belt enforcement measures

The following types of seat belt enforcement measures have been investigated:

- Stationary control at the roadside, checkpoints, mostly combined with speed or DUI enforcement
- Canadian and USA STEP program
- Combinations of checkpoints and mobile controls
- Educational enforcement of use of child restrains with leaflets and warnings instead of fines

Most studies have investigated the effects of primary seat belt law enforcement on seat belt wearing. Although all seat belt enforcement measures differ in several ways, there are no clearly distinguishable groups of different types of enforcement measures. The Canadian and USA STEP programs have been investigated in only one study each. Meta-analysis is conducted based on the studies of the effects of seat belt enforcement on seat belt wearing.

Drivers, front and back seat passengers: Effects of seat belt enforcement have been investigated on car occupants in different seating positions: drivers, front seat occupants, front seat passengers, and back seat passengers.

Daytime vs. night time: Most studies have not specified whether effects have been investigated at day or at night. Three studies have investigated effects on seat belt wearing at day and at night separately.

Visibility of enforcement: Most studies have investigated seat belt enforcement that was not signposted or there where no information whether or not signposting was used. Only two studies have specified that seat belt enforcement was signposted.

Randomization: In most of the studies randomization procedures are not clearly specified. Subgroup analyses have therefore not been conducted.

Change of type or amount of enforcement: All studies refer to either increased seat belt enforcement or to a combination of increased and changed seat belt enforcement.

Country: Most studies have been conducted in the Netherlands or USA. Some studies have also been conducted in Australia, Belgium and Canada. Subgroup-analyses are conducted for each of these countries.

Publicity: The studies have been classified according to the amount of accompanying publicity resulting in four subgroups: A) No publicity B) Local publicity C) Publicity campaign, i.e. enforcement is accompanied by a campaign incl. a mass media component, and D) Seat belt enforcement is part of a wider enforcement programme, most likely including publicity.

Study methodology: All studies have compared seat belt wearing rates before and after the amount of seat belt enforcement has been increased/changed. Six studies have used a control group and in the remaining studies no control group has been used.

Meta analysis of seat-belt enforcement studies

Publication bias: Trim and fill analysis is conducted with all effect estimates. No new effect estimates have been generated in any of these analyses, which indicate that the results are not affected by publication bias.

Results from tests of heterogeneity and summary effects of the estimated effects of seat belt enforcement on seat belt wearing rates are shown in Table E-7. The overall result is a significant increase in the seat belt usage rates. Larger increases of the use of seat belts have consistently been found in comparisons of seat belt use before and during the implementation of enforcement measures.

Table E- 7: Overall summary effects (Random Effects models) of all seat belt enforcement studies

Before – During					Before - After				
Test of heterogeneity			Change of seat belt usage(%)		Test of heterogeneity			Change of seat belt usage(%)	
Cochran's Q	df	P	Summary effect	95% confid-ence interval	Cochran's Q	df	p	Summary effect	95% confide-ence interval
924.28	29	0.000	+21	(+16; +27)	391.37	29	0.000	+15	(+10; +20)

The subgroup analyses show that:

- seat belt enforcement is more effective in the USA than in other countries when regarding before-during comparisons, and least effective in Belgium;
- in before-during comparisons larger increase of seat belt use have been found in studies which have not applied a control group; however, in before-after comparisons the difference between studies with and without control group is only small,
- there are no systematic differences in the effects on drivers and front seat passengers,
- seat belt enforcement is more effective in increasing seat belt use at night; this result refers only to before-during comparisons and is based on only one study,
- an increase of seat belt enforcement is more effective than a simultaneous increase and change of the type of enforcement
- seat belt enforcement that is conducted without signposting is more effective than signposted enforcement,

- local publicity and a publicity campaign increase the effectiveness of seat belt enforcement compared to no publicity or enforcement programmes (it is likely that enforcement programmes include at least some components of a publicity campaign).

The amount of heterogeneity remains significant in almost all subgroups. This indicates that the results within each of the subgroups are affected by further moderator variables.

Enforcement of the use of child restraints

Only one study has been found that has investigated the effect of enforcement on the use of child restraints. In this study a non-significant increase of the use of child restraints by 15% has been found (-13; +25). This study refers not strictly speaking to enforcement, since no fines have been issued. Only information was provided and drivers not properly using child restraints got warnings but no fines.

1. INTRODUCTION

The acronym PEPPER means *"Police Enforcement Policy and Programmes on European Roads"*. The objective of the PEPPER project is to enhance the effectiveness and efficiency of the police enforcement of road traffic. The project looks critically at all relevant aspects of enforcement, such as target behaviours, the detection of infringements, administrative and legal handling after infringement, decisions concerning the volume, location and timing of enforcement, effects of enforcement on road user behaviour and accidents, enforcement methods and tools, collection of enforcement data, and enforcement in a social context. Speeding, drink driving and use of seat belts are the key areas which are targeted in the PEPPER project.

1.1 Workpackages of the PEPPER project

The PEPPER project has organized its activities in five workpackages (WPs):

WP1, entitled "*Strategic, legal, administrative and social context of TLE*" – studies the role of enforcement in traffic safety policies, and analyses the roles of different stakeholders. The results indicate how the enforcement chain could be strengthened.

WP2 – "*Model for enforcement data collection systems and associated pilots*" – develops models for strategic enforcement monitoring databases. The results serve the development of enforcement methods and monitoring and planning of enforcement.

WP3 – "*Innovative technologies and approaches for improving compliance with traffic laws"* – studies the possibilities and cost benefit ratio of modern machine vision and communication technologies in enforcement.

WP4 – "*Good practices in traffic enforcement*" – defines good practices in traffic enforcement by studying current practices, producing scientific estimates of the effectiveness and efficiency of different enforcement methods, assessing monitoring and evaluation methods and surveying current realities in TLE.

WP5 – "*Dissemination*" – concentrates on spreading the results across relevant stakeholders in Europe. To ensure maximum penetration and easy access, the results are disseminated also in targeted seminars, on CD-rom and in the internet, in addition to more conventional media. New member states are especially targeted.

The consortium includes leading European road safety research institutes and the European traffic police.

1.2 Objectives and tasks of Work Package 4

Work package 4 assessed good practices in Traffic Law Enforcement (TLE) with special emphasis on speeding, drink driving and seat belt wearing. It identifies current good practices of TLE in EU Member States and presents the state of the art of good practice according to scientific knowledge including estimates of the effects of good TLE practice on road user behaviour and road safety. The WP4 is divided in four tasks:

- Task 4.1: Good practices in strategic planning and tactical deployment of TLE
- Task 4.2: Good practice in the selected key areas: Speeding, drink driving and seat belt wearing
- Task 4.3: Good practice in monitoring, evaluating and predicting the impacts of TLE on behaviour and accidents.
- Task 4.4: An in-depth survey at police forces in selected countries to identify the current realities of TLE in the European Union

Task 4.3 is further divided in to parts – a) and b):

- 4.3a: Good practice in data, data collection, and data use for monitoring and evaluating Traffic Law Enforcement. Task 4.3a is reported in Deliverable 4a (van Schagen et al, 2008)
- 4.3b: Method for the prediction of the effects on safety of traffic enforcement measures Task 4b is reported in Deliverable 4b (Kallberg and Uwe, 2008).

1.3 Good practice in the selected key areas: Speeding, drink driving and seat belt wearing

The main objective of task 4.2 is to give a systematic review of evaluation studies on speeding, drink driving and seat belt wearing by applying meta-analyses to assess the best estimates of the effects on accidents and behaviour. Further, it includes cost-benefit analyses in order to identify the most efficient TLE measures to the extent that key figures are reported. In brief, police enforcement and sanction measures comprise the following: Stationary (visible) speed enforcement; enforcement of driver behaviour by mobile patrols; laws regulating drink driving; enforcement of drink driving and sanctions; measures against recidivism; enforcement of seat belt wearing among drivers and passengers (child restraint systems included); automatic speed enforcement (speed cameras); ticketing, fining and imprisonment, and penalty point systems.

From previous experiences within this field of research it is known that the number of studies is very large, which calls for methods that can systematize the high number of results from evaluation studies. Meta-analysis is considered as the best instrument for giving systematic, valid and reliable overviews of the kind needed and expected from the EC-project PEPPER. By using meta-analysis one is able to calculate best estimates of effects of any given enforcement and sanction measure by a weighted average across a large number of studies.

The cost-benefit analyses (CBA) require information about the costs and the effects of a given measure. Effects on accidents are expected from the meta-analysis. WP4 underlines the importance of public information and this issue will be focused especially in task 4.2 in order to establish specific good practice recommendations for combining enforcement with public information.

As yet there is not much information on drink driving enforcement issues in the countries with 0.0 BAC limit (Czech Republic, Slovakia, Hungary and Croatia). Therefore, in a subtask of

4.2, studies will be carried out in these countries to evaluate their approach to enforcement of drink-driving regulations, including measures against recidivism and other, related issues in the countries. This subtask is led by the Czech institute CDV and reported in working paper 41 (Rocakova, 2008).

1.4 Effects of enforcement measures: State-of-the-art per 2004

TØI's Handbook of Road Safety Measures addresses a number of traffic law enforcement measures (Elvik and Vaa, 2004). The motivation for the presenting the below list is to show that the key areas addressed in PEPPER can be seen from different angles, i.e. from applying different measures which all, more or less, could be directed towards speeding and drink driving. It should be underlined, however, that this extended list comprises more measures than the ones that will be considered in PEPPER. It follows, that a consideration and meta-analysis will be requested only for a subset of the measures. The list below is more or less a complete list of enforcement measures, where we have made a separation between enforcement measures addressed by PEPPER task 4.2 (list A) and an additional list of enforcement measures (list B), which are not specifically addressed by task 4.2 (an (M) indicates that an estimate of effect has been assessed by meta-analysis).

List A: Enforcement measures addressed by PEPPER:

- Stationary, visible enforcement of speed (M)
- Stationary, visible enforcement of drink driving (M)
- Mobile speed enforcement (marked cars, civilian cars) (M)
- Mobile enforcement of drink driving
- Stationary speed cameras (M)
- Mobile speed cameras
- ”Publicized enforcement of drink driving” (”Australian Random Breath Testing”-campaigns) (M)
- Enforcement of seat belt wearing, including child restraint systems (M)

List B: Other enforcement measures:

- Law regulation of drink driving (per se laws) (M)
- Lowering the BAC-level – all drivers (M)
- Lowering the BAC-level – new/young drivers
- Increasing the age limit for serving alcohol (M)
- Reducing the age limit for serving alcohol (M)
- Treatment/rehabilitation as alternative to licence revocation (M)
- Fining, licence revocation, imprisonment (all together) (M)
- Change of sanction: From imprisonment to fining (M)
- Automatic enforcement of read light crossing (M)
- Ticketing and fining, including increasing the amount (of relevance for seat belt wearing rates)
- Warning letters (alone) (M)
- Penalty point systems
- Licence revocation (alone) (M)

- Penalty point systems, warning letters and licence revocations (all together) (M)

The meta-analysis which has been done previously for most of the measures on list A and B, were done in 1997. The estimates are then more than 10 years old. Meta-analysis of the measures on list A is considered a mandatory activity of task 4.2.

When considering the literature addressing measures on list A, it is expected that task partners will come across information on issues regarding strategic planning and tactical deployment of police enforcement, i.e. information on issues like

- Effect mechanisms of enforcement: Time and distance halo effects – ”spreading in time and space”
- Randomisation of sites and times
- Level of enforcement/enforcement intensities (hours per day or week)
- Effects of using unmarked/civilian cars
- Criteria for deploying speed cameras (as accident risk levels, accident densities and average speeds, etc)

The present update of previous estimates of enforcement measures is justified because more studies have accumulated over the years, new enforcement measures have been added, and more sophisticated evaluation methods have contributed to improved effects estimates.

2. METHOD: THE USE OF META-ANALYSIS TO SUMMARISE STUDY RESULTS

The basic entity of a meta-analysis is a *result*. By a *result* is meant an estimate of the change in the number of accidents, odd ratio, accident risk, relative risk, etc. Meta-analysis may be described as a procedure for summing up all the individual results from different studies on a given variable, by a weighted average. The weights of each of the results are calculated in such a way that the statistical uncertainty in the weighted average is minimised. This is done by assigning a statistical weight, which is inversely proportional to the variance of each of the individual results (Fleiss 1981). The weights in turn depend on the accident counts, which mean that the more accidents an individual result is based on, the higher is the statistical weight of that result. A single study can contain more than one result. In such cases, all results, or the most important results from studies with a very large number of results, have been included in the meta analyses. Multiple results from the same study have been treated as statistically independent, although this assumption may not always be correct.

There are two methods of combining estimates of effects in meta-analysis. These are referred to as the *fixed-effect* model and the *random-effect* model. A fixed-effect model is assumed when a given measure is supposed to have the same effect or consequence across contexts that may vary – for example across countries, cultures, sub-groups, times, etc. The effect of a given condition would naturally not be exactly the same from context to context, some variation would be expected, but under a fixed-effect model the variation is regarded as random, not systematic. A random-effect model would be more appropriate where the effect of a given condition is considered to vary systematically. Adequate steps in meta-analysis can be described as follows:

2.1 Meta-analysis step-by-step

Recent developments in the field of meta-analysis have led to recommendations that the procedure be supplemented with certain statistical methods to improve validity. These methods attempt to account for potential biases and any non-random variation in the dataset summarised. Since they did not employ such methods, those meta-analyses carried out under the GADGET project are now considered to be relatively basic.1 At the time of GADGET we knew too little about the conditions to be met for valid application of meta-analyses, about weaknesses in the method, about tests that could identify biases in the data, or about the extent to which overall best estimates could be generalised. New knowledge in each of these areas means that the meta-analysis procedure can be supplemented with formal steps that improve robustness of the results. The supplemented process is set out below.

1. Carry out meta-analysis on the assembled set of effects using the log-odds method, in which odds ratios of effects are transformed to their natural logarithms (log-odds).

[1] The present description of steps in meta-analysis is developed under three projects on road safety campaigns: GADGET (EU-project), INFOEFFEKT (a joint Norwegian/Swedish project) and CAST (EU-project to be completed in January 2009). The description is considered to be equally relevant for the PEPPER project as it is for the CAST project.

2. Test the set of effects (the set of natural logarithms of the effects) for publication bias.
3. Compensate for this bias as necessary using the trim-and-fill method to simulate any 'unpublished' effects missing from the dataset.
4. Test the set of effects for heterogeneity, both before and after adjusting for publication bias.
5. Use the result of this test to decide whether to use a fixed- or random-effect model to describe the overall effect for a set of effects.
6. Assess the output of steps 2 to 5 to make a conservative best estimate of overall effect.
7. Divide the whole set of effects for a single outcome measure into sub-groups, according to variables defining aspects of enforcement evaluation, or enforcement content. Obtain overall effects for each sub-group.
8. Use differences between overall effects for different sub-groups to develop hypotheses about the causes of systematic variation in enforcement effect.
9. Test hypotheses by using them to inform a model for meta-regression analysis, and thereby accept or reject ideas about the partial effects of different enforcement elements.

Apart from the first step, each of these steps represents a development of the method used to estimate overall campaign effect in GADGET (Delhomme et al, 1999). The new steps were followed and refined during the INFOEFFEKT project, which also estimated the effect of campaigns on accident counts (Vaa et al, 2004). However, the criteria employed for study selection under INFOEFFEKT were narrow, such that the number of campaign effects selected for the final set summarised was low. CAST represents the first use of the refined method on a substantially broader dataset.

To explain the technical developments responsible for improvements in meta-analysis, a full step-by-step description is now given.

Step 1*. *The log-odds method: There are several reasons to transform odds ratios to their natural logarithms, but in simple terms such transformation addresses statistical demands, the most fundamental of which is that the set of effects should be normally distributed.

Step 2. Publication bias. Publication bias describes a tendency for authors and editors to publish only those studies demonstrating desirable, statistically significant effects. To put it another way, it is a tendency to avoid publishing those studies that fail to demonstrate desirable effect. There is documented evidence for publication bias, and it means that any set of effects gathered from accessing a selection of available studies will be incomplete in that it will lack those undesirable or non-significant effects that were never published. Such an incomplete set of effects will not be representative of the true overall effect. Fortunately, statistical tests are now available that can be used to detect publication bias in a set of effects.

Step 3. Correction for publication bias. Several methods have been developed to correct the overall effect of a set of effects for which publication bias has been indicated. The 'trim-and-fill' method proposed by Duval and Tweedie (2000a & b) has been developed to test and, where such bias is indicated, adjust the overall effect estimate by generating the missing, 'undesired' effects to complete the original set of effects. Publication bias can be shown using a so-called funnel-plot. This plot is the basis for the trim-and-fill method. According to the method the level of bias is equivalent to the level of asymmetry in this plot, which results from the absence of effects on the ’undesired’ side of the funnel. The method generates artificial effects and uses them to supplement and restore symmetry to the set of effects retrieved. A new, corrected overall effect estimate is then calculated, based on the new completed set of effects.

Step 4 & 5: Fixed- and random-effects**: **homogeneity vs heterogeneity. One must decide whether to use a so-called fixed-effects or random-effects model when carrying out meta-analysis. A fixed-effects model assumes that the same intervention (in our case an enforcement measure) will have the same effect regardless of whether it is carried out on different people, at different times or in different countries. It is assumed that the characteristics of the intervention are such that it will have the same effect in different contexts. This does not assume that the effect will be exactly the same every time the intervention is used, but that any variation that is observed will just be due to chance. A random-effects model, on the other hand, assumes that characteristics of the situation, as well as the intervention itself, can cause variations in effect, as indeed can characteristics of the method used to measure the effect. Fortunately again, a statistical test can be used to inform about the degree of heterogeneity (Everitt 2002). When a significant amount of heterogeneity is detected, the application of a fixed-effects model would be inappropriate. Amongst other things, the size of the confidence interval would be underestimated. In these cases, a random-effects model is more appropriate. When no significant amount of heterogeneity is detected, the fixed-effects model may be applied. When there are smaller amounts of heterogeneity, the results from a fixed- and a random-effects model are more similar than when there are larger amounts of heterogeneity. Hence, a random -effects model is applied in all cases, where the number of available effect estimates is sufficient. The advantage of always applying a random-effects model is that heterogeneity is taken into account also when the amount is not statistically significant.

Step 6. Conservative best estimate of overall effect. Using the preceding steps it is possible to generate estimates of overall effect in four different ways:

1. using a fixed-effects model without adjusting for publication bias;
2. using a fixed-effects model and adjusting for publication bias;
3. using a random-effects model without adjusting for publication bias; or
4. using a random-effects model and adjusting for publication bias.

For the sake of simplicity however, only the results that are based on a Random Effects model will be presented in this report. The results from a Fixed Effects model are only presented when no Random Effects model could be applied (because of too few effect estimates).

The statistical output on the level of publication bias and heterogeneity in the dataset is assessed together to give the *best estimate* for an intervention’s overall effect. Together these statistical assessments help answer questions of generalisation. They address to what extent, based on the set of effects retrieved, one can expect to achieve overall effect in different situations, and therefore what level of confidence we should employ to describe the overall effect. They therefore help to make more useful predictions about future situations. By informing on the choice of best estimate, the tests also allow a better approximation of the true overall effect, and thus provide an accurate basis for subsequent cost-benefit analyses.

We have chosen to follow a path of caution when choosing the overall best estimate in order to avoid overestimating the overall effects of enforcements. This path can be summarised in three points:

1. always use the random-effects model where heterogeneity is indicated;
2. where there is heterogeneity include an estimator to account for any systematic variation in each set of effects; and
3. calculate overall effects for corrected sets of effects where publication bias is evidenced.

These points summarise the strategy followed in the analyses and in the presentation of the results. Because it is cautious it will tend to give conservative effect sizes associated with higher levels of uncertainty (confidence intervals). It will also, however, minimise the risk of overestimating the size of enforcement effects.

Step 7 & 8 –Model development and choice of explanatory variables: In the PEPPER project we want to identify those factors underlying any statistically significant overall effects that a enforcement intervention is shown to have. This involves taking the meta-analysis a step further by employing subgroups analyses and meta-regression.

Meta-regression is an expansion of meta-analysis in which one models the relationships between an intervention's effect and known explanatory variables using regression (Everitt 2002). It is preferable that meta-regression is carried out on a model, effectively a set of potential explanatory factors. To identify these factors, a set of enforcement effects is divided into sub-groups of effects according to simple variables that define aspects of enforcement evaluation and enforcement content. Meta-analysis is carried out to obtain overall effects for each of these sub-groups. The differences between overall effects of different sub-groups are then used to develop hypotheses about the causes of systematic variation in enforcement effect.

Without subsequent meta-analysis, sub-group analysis itself is an unreliable way to identify factors influencing the effect of enforcements. This is because it is simply a set of bivariate tests: each sub-group analysis assesses only the effect of one variable on another variable. Sub-group analysis cannot be used to test variables in the presence competing effects from other variables, and the results can therefore be misrepresentative of the systems of overlapping influences that together alter effect sizes. To be able to assess systems of variables together we need to use meta-regression, and we can use sub-analyses to identify candidate variables for meta-regression. A series of sub-analyses were in fact carried out under the GADGET project with an ultimate aim to assess whether these factors were potential candidates for future exploration by meta-regression.

Step 9 – Meta-regression and testing hypotheses. The final step consists of multivariate regression analysis based on the model developed in steps 7 & 8. As stated, the advantage with multivariate analysis, in which several variables can be tested together, is that it can inform about effects of a variable of interest in the presence of other controlling variables in the model.

2.2 Basics of the log odds method of meta-analysis

The log odds method of meta-analysis has been applied throughout (Fleiss 1981, Shadish and Haddock 1994). According to this method, a weighted mean estimate of effect is calculated on the basis of the estimates of effect found in the studies which have been retrieved. This method of meta-analysis was chosen because the odds ratio is the most commonly found estimate of effect in road safety evaluation studies. For example, the estimate of effect (odds ratio) in a before-and-after study with a comparison group is:

Estimate of effect = odds ratio = (A/B)/(C/D)

A = Number of accidents (or injured road users) for treated group after treatment
B = Number of accidents (or injured road users) for treated group before treatment
C = Number of accidents (or injured road users) in comparison group after treatment
D = Number of accidents (or injured road users) in comparison group before treatment

In studies that employ multivariate techniques of analysis, effects are normally stated in terms of an odds ratio that has been adjusted for confounding.

A weighted summary effect is calculated as follows:

1. The natural logarithm is calculated of each estimate of effect (log-odds)
2. A statistical weight is calculated for each log-odds
3. The weighted mean is calculated of all log-odds
4. The weighted mean of the log-odds is transformed to an odds ratio by taking the inverse of the natural logarithm.

The log odds method of meta-analysis takes the logarithm of the odds ratio as the estimate of effect because the logarithm of the odds ratios follows a normal distribution. Combining logarithms of odds ratios yields an unbiased estimate of the weighted mean effect of a set of studies.

Each estimate of effect is assigned a statistical weight which is inversely proportional to its variance. The variance of the logarithm of the odds ratio is:

$$v_i = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}$$

A, B, C, and D are the four numbers that enter the calculation of the estimate of effect. In studies that do not use comparison groups, the terms 1/C and 1/D drop out. The same applies to studies that state the effects of a road safety measures in terms of an accident rate ratio. Statistical weights are estimated on the basis of the recorded number of accidents. In case of zero accidents, 0.5 is added to all four (or two) numbers used in estimating the statistical weight of a result.

2.3 Literature search and coding of studies

In the search of literature, the studies used in meta-analysis done earlier (Elvik and Vaa, 2004), formed the basis of the search. In addition, several literature searches have been done to update the initial basis and identify the accumulation of new, relevant studies. Lead by SWOV and supplied by TØI, reference lists and code-books of each of the three key areas were elaborated. Subsequently, SWOV and TØI made agreements regarding the division of studies to be coded and fed into the databases that served as the bases of subsequent meta-analyses.

3. SPEED ENFORCEMENT

3.1 Speed enforcement measures

The preparation for meta-analysis of speed enforcement measures draws upon the distinction between measures adopted in the review of literature in Elvik and Vaa (2004). At that time the following subgroups of speed enforcement measures were separated:

- Stationary speed enforcement
- Patrolling
- Speed cameras (automatic speed enforcement)

Halo effects: A distinction should be made between speed enforcement using stationary methods and speed enforcement using "mobile" methods, or police patrols (Elvik and Vaa, 2004). The distinction is relevant because halo effects have been found in time and space for stationary enforcement, but not for mobile patrols (Shinar and McKnight 1985; Vaa 1993). "Halo effects" in time and space means that an effect can be found during a given period of time and/or at a certain distance from the spot where the speed enforcement is carried out.

The following methods of speed enforcement are distinguished:

- **Stationary speed enforcement using laser or radar** that measures speed from one, usually unobtrusive or hidden observation site, or instruments that measure mean speed between two fixed observation sites and clearly visible apprehension sites staffed by uniformed police officers and marked cars.
- **Stationary radar enforcement "American type".** The police observer (sometimes one officer alone in a car) measures speed by a radar mounted on the window and then pursues offending vehicles straightaway in order to apprehend and sanction the speeding driver. This technique is also used in Australia and the observation unit may be overt (marked car) as well as covert (unmarked car).
- **Patrolling**: Mobile police patrols with uniformed cars or motorcycles.
- **Composite police controls with stationary and visible elements**: This term is used to illustrate that speed enforcement may utilize a whole range of different techniques and methods, but also that it comprises some element that is stationary and that some of the activity is visible to the drivers passing by.
- **Speed cameras:** Several deployment patterns can be distinguished: 1) Fixed speed cameras, most often visible, on fixed locations/poles with a mobile camera moving around, or 2) Mobile cameras, less obtrusive or even hidden cameras used on different locations, and 3) section control of speed (where the average speed between two fixed sites is calculated and enforced if the speed limit is violated).

In the PEPPER project task 4.2 the scope has been to collect more studies and results and also to differentiate more specifically between speed enforcement measures as new measures might have entered the speed enforcement scene since 1997 which was the year when the previous meta-analysis of traffic law enforcement measures was performed (Elvik and Vaa, 2004).

3.2 Description of speed enforcement evaluation studies

In task 4.2 a total of 45 evaluation studies were identified and found to be of an acceptable quality to be included in the data. Table 1 presents the expected subgroups of speed enforcement measures that might be found in evaluation studies, but, as can be seen, some subgroups were empty.

Table 1. Amount of research evaluating safety effects of speed enforcement measures. Number of results and sum of statistical weights of the studies.

Measure	Number of results
Stationary control (observation hidden, apprehension visible/uniformed)	17
Speed control with radar (" American type": same car detects and apprehends)	15
Visible speed cameras on fixed locations (mobile camera)	28
Hidden speed cameras used on different locations	11
Section speed control (with speed cameras)	2
Patrol/Mobile enforcement - marked car	11
Patrolling - civilian car	0
Patrol/mobile with motorcycle	0
Composite police controls with stationary and visible elements	40
Unmanned radar (no fines)	2
Radar control "American type" also used in Australia same car detects and apprehends)	3
Total	129

The 45 studies have been published within 14 countries and comprise a total of 129 results. USA, Australia, UK and Sweden are the ones which have had the largest numbers of results from speed enforcement evaluations. Table 2 presents the number of results according to country.

Table 2. Amount of research evaluating safety effects of speed enforcement measures. Number of results according to country

Australia	26
Belgium	1
Denmark	1
Finland	2
Germany	9
Israel	1
Korea	4
The Netherlands	8
NewZealand	2
Norway	4
Sweden	12
:UK	23
USA	33
Irland	3
Total	129

The distribution according to year of publication (table 3) shows that two thirds of the studies are quite recent as they have been published in the 1990s or after 2000, but there are also some studies that date back to the 1960 and there is even one from 1958 (Shumate, 1958).

Table 3. Distribution of studies evaluating safety effects of speed enforcement measures. Number of studies according to year of publication

1958	1
1960 – 1969	3
1970 – 1979	5
1980 – 1989	6
1990 – 1999	16
2000-2008	14
Total	45

3.3 Factors affecting the effectiveness of speed enforcement

A number of potential moderator variables is investigated in meta-analysis, i.e. variables that are assumed to affect the effectiveness of speed enforcement measures. Moderator analyses are conducted in two ways: In subgroup analyses and in meta-regression.

In subgroup analyses studies are divided into subgroups according to a potential moderator variable. The summary effects of the subgroups of studies are then compared. In most cases there are only two subgroups. An important point in subgroup analysis is to keep the subgroups comparable with respect to other comparable moderator variables. Subgroups analyses are mostly only made when there is significant heterogeneity in the results for a group of effect estimates (e.g. all injury accidents). Heterogeneity indicates that the results are affected by some factor or factors that are different between the studies. If there is no heterogeneity, all effect estimates can be regarded as representatives of the same outcome and no systematic differences between studies can be expected.

A grouping variable is assumed to be a relevant moderator when two criteria are fulfilled:

- the summary effects for the subgroups are different,
- heterogeneity within the subgroups is reduced as compared to the overall summary effect.

The second criterion is however not a very strict criterion. When it is not fulfilled (one finds for example frequently reduced heterogeneity only in one of the subgroups), this may indicate that the moderator variable is not relevant. It is however also possible that the moderator variable all the same is relevant, but that the results within one or more subgroups are affected by further moderator variables. This can be investigated by further subgroup analyses or by meta-regression. In a meta-regression analysis, the effect of several potential moderator variables is investigated at the same time.

Visibility: The enforcement measures can roughly be classified as

- **Visible** enforcement when police cars, measurement equipment or cameras are either made visible or at least not hidden from view.
- **Non-visible** enforcement when civil police cars are used or when measurement equipment or cameras are hidden.

Signposting: Additionally, enforcement may be signposted. Studies are classified as

- Signposted enforcement.
- Not signposted.

Randomization: In most of the studies randomization procedures are not clearly specified. In a few studies it is stated that there has been some kind of randomization and some studies state that enforcement has been conducted at times and places that have been selected based on e.g. high accident numbers or high speeds. These studies are classified as

- Randomization, either according to time or place,
- No randomization.

Accompanying publicity: Speed enforcement measures are classified according to the type of accompanying publicity:

- Publicity campaign: The enforcement measure is part of a wider campaign (independent of whether or not paid media is used).
- Local publicity: The enforcement measure is not part of a wider campaign, but accompanied by local publicity.
- No publicity: No activities are conducted in order to achieve publicity for the speed enforcement measure.

Change of type or amount of enforcement: The speed enforcement measures have been evaluated either after the introduction or the intensity of enforcement has been changed. Accordingly the studies are classified as

- New type of enforcement or
- Increased enforcement.

In almost half of all studies however it is not clear whether a new for of enforcement has been introduced or the intensity has changed.

Increase of the amount of enforcement: The studies where the amount of enforcement has been increased are classified according to the increase of the amount of enforcement:

- Less than doubling the amount of enforcement.
- Doubling or larger increase of the amount of enforcement.

The amount of the increase of enforcement has not been reported in all studies. In some of the studies where the type of enforcement has been changed, the total amount (e.g. in terms of the number of person hours) has increased as well.

Country: Speed enforcement measures often follow certain procedures and use specific equipment in different countries. Effects of speed enforcement are therefore compared between different countries. According to the availability of studies, the following (groups of) countries are compared:

- Australia
- Sweden
- UK
- USA

Others; other countries are Belgium, Denmark, Finland, Germany, Ireland, Israel, Korea, New Zealand, Norway, and The Netherlands

Accident severity and types of accidents: In most studies it is specified if accident numbers refer to injury accidents or to all accidents including property damage only accidents. Some studies have reported results for different degrees of accident severity, e.g. fatal, serious and slight injuries. Accordingly, in the meta-analysis effects are estimated for the following groups of accident severity:

- All injuries: Results for all injuries include results for all injury accidents. These include for the most part fatal accidents.
- Fatal accidents.
- Serious injury.
- Slight injury.
- Unspes/Property-Damage-Only (PDO) accidents includes accidents where the severity is not specified (for the most part injury and property damage only accidents) and property damage only accidents.

Study methodology: Different study methodologies have been used to evaluate the effects of speed enforcement. Most studies have compared accident numbers before and after the introduction of a measure. These studies differ with respect to the type of control group that has been used as a comparison. Control groups that have been used are:

- **No control group:** The number of accidents before and after the introduction of the measure are compared and no other control than the time period before and after is used. In these studies changes in accident numbers may be affected by all other factors that have changed at the same time or in the same time period as the measure has been introduced and conducted.

- **Trend only controlled**: Some studies have applied time series models in order to investigate to what degree accident numbers after a measure has been introduced deviate from the accident numbers that would have been expected if the trend in the accident numbers before the introduction of the measure had continued. These studies are better controlled than studies without comparison group, the results may however be affected by other changes that occurred at the same time as the introduction of the measure.

- **With control group**: Studies with a control group have compared accident numbers before and after the introduction of an enforcement measure and compared the change with the change that occurred during the same time period on roads or in areas where no enforcement measures has been introduced.

- **Multivariate studies**: A few studies have applied multivariate models in order to examine the contribution of speed enforcement and numerous other factors on accident numbers. Factors controlled for vary between studies.

Study design: Based on the study methodology all studies are roughly classified as “good” and “weak” study designs.

- **Weak study designs** are studies without a comparison group and studies that have controlled for time trends only.

- **Good study designs** are the studies that have used a control group and multivariate studies.

Both within the “good” and the “weak” study designs, study quality varies, dependent for example on the choice of comparison group, the statistical methods applied, and the quality of the available accident data. However, it can be assumed that most studies that are classified as “good study design” actually have controlled for confounding variables to a larger extent than studies classified as “weak study design”.

3.4 Meta analysis of speed enforcement studies

Studies included in the meta-analysis: An overview of the available studies that are included in the meta-analysis is shown in Table 4. The classification of injury severity and study design is made as described in the section above.

Table 4: Studies of the effects of speed enforcement on accidents.

Authors	Year	Country	Type of measure	Study design	Sum of statistical weights1
Amann	1990	Germany	Stationary manual	Weak study	16,474.7
Andersson	1991	Sweden	Stationary manual	Good study	237.7
Bourne & Cooke	1993	Australia	Mobile speed cameras / radar	Good study	344.6
Brackett	1980	USA	Radar/laser US/AUS	Good study	1,216.1
Cameron, Newstead, Diamantopoulou & Oxley	2003	Australia	Mobile speed cameras / radar	Good study	573.8
Carlsen	1990	Norway	Stationary manual	Good study	63.4
Diamantopoulou	2002	Australia	Radar/laser US/AUS	Good study	158.0
Diamantopoulou	1998	Australia	Radar/laser US/AUS	Good study	236.3
Dobrotka	1998	USA	Composite	Weak study	1,723.3
Ekstrøm	1967	Sweden	Stationary manual	Good study	17.4
Ekström, Kritz og Strömgren	1966	Sweden	Composite	Good study	8.7
Elvik	1997	Norway	Fixed speed cameras	Good study	852.0
Fuller	2002	Ireland	Composite	Good study	200.4
Gains	2005	UK	Fixed speed cameras	Weak study	9,822.2
Goldenbeld & Van Schagen	2005	The Netherlands	Mobile speed cameras / radar	Good study	198.5
Ha, Kang & Park	2003	Korea	Fixed speed cameras	Weak study	616.4
Hakkert, Gitelmann, Cohen, Doveh & Umansky	2001	Israel	Composite	Good study	173.4
Hooke	1995	UK	Fixed speed cameras	Weak study	3.3
Jones, Sauerzapf & Haynes	2008	UK	Fixed speed cameras	Good study	142.3
Jorgenson	1973	Denmark	Patrolling	Good study	82.1
Kang	2002	Korea	Fixed speed cameras	Weak study	1,204.2
Lamm & Kloeckner	1984	Germany	Fixed speed cameras	Weak study	161.6
Leggett	1988	Australia	Radar/laser US/AUS	Good study	130.7
London Accident Analysis Unit	1997	UK	Fixed speed cameras	Good study	1,204.3
Machemer	1995	Germany	Stationary manual	Good study	13.4
Mason	1971	USA	Patrolling	Good study	320.0
Mason	1970	USA	Patrolling	Good study	1,022.3
McCartt & Rood	1989	USA	Radar/laser US/AUS	Weak study	749.2
Munden	1966	UK	Composite	Good study	80.5
Newstead	2003	Australia	Stationary manual	Good study	15,950.7
Novak & Shumate	1961	USA	Radar/laser US/AUS	Good study	21.1
Nuyts	2006	Belgium	Fixed speed cameras	Good study	37.0
Oei & Polak	1992	The Netherlands	Fixed speed cameras	Good study	50.1
Pez	2002	Germany	Composite	Good study	578.2
Pigman, Agent, Deacon & Kryscio	1987	USA	Mobile speed cameras / radar	Weak study	137.2

Authors	Year	Country	Type of measure	Study design	Sum of statistical weights1
Roszbach & Blokpoel	1991	The Netherlands	Composite	Good study	747.5
Salusjärvi & Mäkinen	1988	Finland	Stationary manual	Good study	1,421.9
Saunders	1977	Australia	Composite	Good study	60.2
Saunders	1977	Australia	Patrolling	Good study	71.1
Shumate	1958	USA	Patrolling	Good study	1,726.9
Statens vegvesen Buskerud/UP	1996	Norway	Stationary manual	Good study	30.8
Stefan	2006	Australia	Section control	Good study	18.5
Swali	1993	UK	Fixed speed cameras	Weak study	1,067.8
Tay	2000	NewZealand	Radar/laser US/AUS	Good study	268.8
Veling	1992	The Netherlands	Composite	Good study	9.4
Wildervanck	1993	The Netherlands	Composite	Weak study	24.4
Winnett	1994	UK	Fixed speed cameras	Good study	250.3

[1] Statistical weights in fixed effects model

Data inspection, publication bias: A scatterplot of all effect estimates for speed enforcement is shown in Figure 1. It is indicated in the figure for all effect estimates whether they are based on a good or a weak study design. The natural logarithm of the overall summary effect that is based on all studies is -0.181, which corresponds to an odds ratio of 0.835 (Fixed Effects). In the fixed effects model a somewhat larger accident reduction has been found in studies with a good design, compared to studies with a weak design. This difference is further explored below, it is shown that the difference is reversed in the Random Effects model.

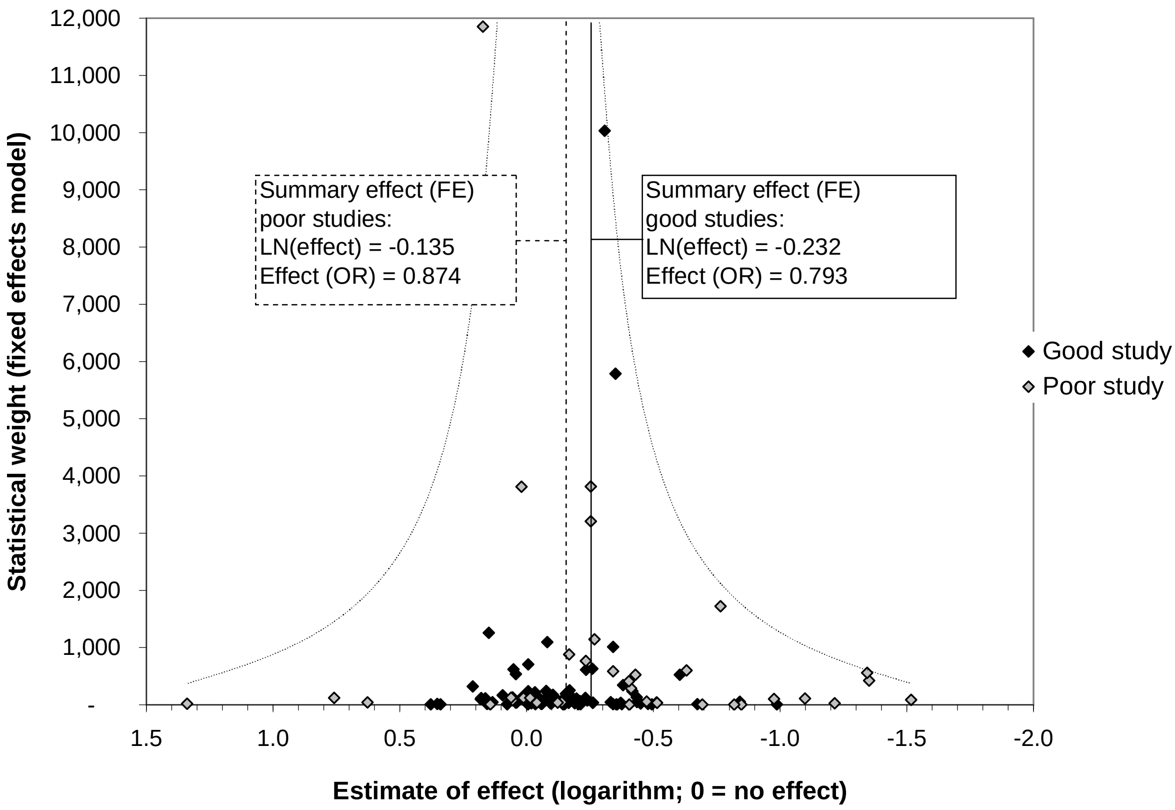

Figure 1: Scatterplot of effect estimates for speed enforcement.

The effect estimates are distributed in a funnel-like pattern. A trim and fill analysis is conducted with all effect estimates. No new effect estimates are generated in the analysis. There is no indication that the results are affected by publication bias. A trim and fill analysis that is based on the weak study designs only does not generate new effect estimates either. In a trim and fill analysis that is based on the good study designs two new effect estimates are generated and the summary effect changes only slightly. These results are in more detail described and commented in the sections below. On the whole, the trim and fill analysis do not indicate the presence of publication bias.

Summary effects and tests for heterogeneity: Results from tests of heterogeneity and summary effects of the estimated effects of speed enforcement on accidents are shown in Table 5. The overall result i a significant reduction of the number of accidents and there is no indication of publication bias according to the result from the trim and fill analysis. There are large differences between the estimated effects of the different types of speed enforcement. The trim and fill analyses indicate that the results for mobile speed cameras are affected by publication bias. When controlling for publication bias, the accident reduction of mobile speed cameras is smaller than without control for publication bias.

Table 5: Test of heterogeneity and summary effects of speed enforcement (Random-effects models).

	Test of heterogeneity			Change of number of accidents (%)		Trim and fill analysis: Change of number of accidents (%)		
	Cochran's Q	df	p	Summary effect	95% confidence interval	Summary effect	95% confidence interval	Number of new effect estimates
All speed enforcement measure								
All measures	5307.82	128	0.000	-18	(-23; -13)			0
All injury accidents	1777.28	**79**	0.000	**-18**	(-24; -13)	-14	(-20; -7)	4
Fatal accidents	13.70	13	0.395	-29	(-35; -23)	-29	(-34; -23)	1
Serious injury	313.3	31	0.000	-30	(-38; -21)	-27	(-36; -16)	1
Slight injury	123.8	10	0.000	-3	(-17; +13)	-2	(-16; 14)	2
Unspes/PDO	3294.1	33	0.000	-15	(-25; -3)	-15	(-21; -9)	4
Results for each speed enforcement measure: All accident severities								
Stationary manual	1854.17	22	0.000	-11	(-22; +1)			0
Patrolling	62.7573	10	0.000	-6	(-16; +4)			0
Radar/laser US/AUS	22.3372	30	0.841	-4	(-7; 0)	0	(-3; 4)	6
Speed cameras	1693.9	42	0.000	-30	(-38; -23)			0
- Mobile speed cameras	168.476	12	0.000	-23	(-38; -4)	-17	(-34; 4)	2
- Fixed speed cameras	1513.02	27	0.000	-34	(-42; -25)			0
Composite Other	454.306	20	0.000	-18	(-33; +1)			0
Results for each speed enforcement measure: Injury accidents								
Stationary manual	323.19	**9**	0.000	**-12**	(-26; +5)			
Patrolling	62.07	**6**	0.000	**-6**	(-20; +10)			
Radarl/aser US/AUS (FE)	12.08	**16**		**+1**	(-3; +5)	+2	(-3; +8)	3
Speed cameras	1040.17	**30**	0.000	**-28**	(-36; -20)	-22	(-33; -10)	2
- Mobile speed cameras	58.15	**9**	0.000	**-14**	(-29; +5)			
- Fixed speed cameras	929.94	**18**	0.000	**-35**	(-43; -25)			0
Composite Other	57.30	**14**	0.000	**-21**	(-33; -7)	-18	(-30; -4)	1
Results for each speed enforcement measure: Fatal accidents								
Stationary manual	2.98	1	0.084	-27	(-45; -2)			
Patrolling	0.15	1	0.695	-12	(-35; +18)			
Radar/laser US/AUS	3.94	4	0.414	-21	(-42; +7)			
Speed cameras	0.86	2	0.649	-33	(-39; -27)			
- Mobile speed cameras	0.00	0		-32	(-39; -24)			
- Fixed speed cameras	0.03	1	0.853	-39	(-50; -25)			
Composite Other	0.32	1	0.573	-16	(-36; +11)			

There are significant amounts of heterogeneity in all results that refer to all accident severities, except for Radar / lased US/AUS. Moderator variables are therefore likely to be present. Speed cameras have been found to yield the largest accident reductions, both of all severities, fatal, and injury accidents. For all measures larger accident reductions have been found for fatal accidents compared to injury accidents.

Effects of potential moderator variables: Subgroup analyses: The effects of a number of potential moderator variables are investigated by comparing summary effects between subgroups of results. For all summary effects that are based on at least 10 effect estimates trim and fill analyses are therefore conducted. The results are summarized in Table 6a, based on all accident severities. The results for injury accidents and for fatal accidents are shown in table 6b and 6c, respectively.

Table 6a: Test of heterogeneity and summary effects (RE models) of speed enforcement – moderator analyses: results for all accident severities.

	Test of heterogeneity			Change of nr. of accidents (%)		Trim and fill analysis: Change of number of accidents (%)		
	Cochran's Q	Df	p	Summary effect	95% conf. interval	Summary effect	95% conf. interval	New effect estimates
Visibility								
Visible	4088.3	77	0.000	-23	(-28; -17)			0
Not visible	173.5	15	0.000	-22	(-34; -7)	-11	(-28; 10)	2
Signposting								
Signposted	1322.5	28	0.000	-38	(-45; -30)	-28	(-38; -17)	5
Not signposted	524.8	43	0.000	-9	(-16; -1)	-3	(-11; 6)	2
Randomization								
Randomization (fixed effects)	3.7	10	0.961	-1	(-6; +5)	+3	(-2; 8)	8
No randomiz.	2508.3	37	0.000	-25	(-31; -18)	-22	(-28; -15)	2
Accompanying publicity								
No publicity	249.6	53	0.000	-14	(-19; -9)	-13	(-18; -8)	2
Local publicity	2384.5	44	0.000	-24	(-31; -17)	-20	(-28; -11)	2
Publicity campaign	193.3	17	0.000	-7	(-22; +11)			0
Change of type or amount of enforcement								
Increased	2650.2	67	0.000	-14	(-20; -6)			0
Changed form	1763.3	50	0.000	-25	(-32; -18)	-21	(-29; -13)	2
Increase of the amount of enforcement								
< 200%	269.2	29	0.000	-3	(-9; +4)	-3	(-9; 4)	1
> 200%	47.6	22	0.001	-13	(-22; -3)	-11	(-21; -1)	1
Country								
Australia	165.7	24	0.000	-10	(-16; -4)			0
Sweden	13.8	11	0.245	-14	(-26; -2)			0
UK	1305.1	23	0.000	-38	(-47; -28)	-32	(-42; -20)	2
USA	861.6	32	0.000	-15	(-27; -3)	-4	(-20; 16)	7
Other	1034.4	34	0.000	-13	(-20; -4)			
Study methodology								
Good study	862.9	89	0.000	-14	(-18; -9)	-13	(-17; -8)	2
Weak study	4300.5	38	0.000	-27	(-37; -17)			0

Table 6b: Test of heterogeneity and summary effects (RE models) of speed enforcement – moderator analyses: results for injury accidents.

	Test of heterogeneity			Change of nr. of accidents (%)		Trim and fill analysis: Change of number of accidents (%)		
	Cochran's Q	Df	p	Summary effect	95% conf. interval	Summary effect	95% conf. interval	New effect estimates
				Visibility				
Visible	1426.18	**45**	0.000	**-25**	(-31; -18)	-17	(-25; -8)	4
Not visible	92.79	13	0.000	-15	(-27; -2)	-10	(-24; 7)	1
				Signposting				
Signposted	890.69	**19**	0.000	**-38**	(-45; -28)	-27	(-37; -14)	4
Not signposted	129.38	**24**	0.000	**-8**	(-18; +2)	-4	(-15; +7)	1
				Randomization				
Randomization (fixed effects)	1.07	**3**	0.784	**-16**	(-25; -5)			
No randomiz.	685.42	**25**	0.000	**-23**	(-29; -18)	-22	(-27; -16)	1
				Accompanying publicity				
No publicity	150.92	**31**	0.000	**-11**	(-18; -4)	-10	(-17; -3)	2
Local publicity	1113.27	**27**	0.000	**-25**	(-33; -16)	-23	(-30; -12)	1
Publicity campaign	55.20	**11**	0.000	**-11**	(-24; +5)			0
				Change of type or amount of enforcement				
Increased	616.20	**45**	0.000	**-16**	(-22; -10)	-14	(-20; -8)	1
Changed form	982.93	**29**	0.000	**-23**	(-34; -12)	**-16**	(-30; +1)	2
				Increase of the amount of enforcement				
< 200%	95.26	**19**	0.000	**-5**	(-12; +2)	-4	(-11; +3)	1
> 200%	37.66	15	0.001	-10	(-22; +4)	-7	(-20; 9)	1
				Country				
Australia	146.89	**18**	0.000	**-5**	(-17; +10)	+8	(-13; +34)	7
Sweden	10.59	**5**		**-15**	(-38; +15)			
UK	895.79	**21**	0.000	**-36**	(-44; -26)	-32	(-41; -21)	1
USA	102.98	**15**	0.000	**-8**	(-18; +4)	-2	(-13; +12)	3
Other	134.90	**16**	0.000	**-15**	(-23; -6)	-9	(-17; 0)	4
				Study methodology				
Good study	410.73	**57**	0.000	**-12**	(-17; -6)	**-10**	(-16; -4)	2
Weak study	1362.22	**21**	0.000	**-31**	(-41; -20)			0

Table 6c: Test of heterogeneity and summary effects (RE models) of speed enforcement – moderator analyses: results for fatal accidents.

	Test of heterogeneity			Change of nr. of accidents (%)		Trim and fill analysis: Change of number of accidents (%)		
	Cochran's Q	Df	p	Summary effect	95% conf. interval	Summary effect	95% conf. interval	New effect estimates
Visibility								
Visible	7.09	6	0.312	-30	(-39; -20)			
Not visible								
Signposting								
Signposted (fixed effects)	0.03	1	0.853	-39	(-50; -25)			
Not signposted (fixed effects)	0.32	3	0.956	-11	(-29; +11)			
Randomization								
Randomization (fixed effects)	0.27	2	0.874	-17	(-35; +6)			
No randomiz.	5.62	4	0.230	-28	(-36; -19)			
Accompanying publicity								
No publicity	6.33	6	0.388	-27	(-38; -14)			
Local publicity	3.28	4	0.512	-32	(-33; -30)			
Publicity campaign (fixed effects)	0.00	0		-1	(-48; +89)			
Change of type or amount of enforcement								
Increased	5.47	4	0.242	-23	(-34; -10)			
Changed form	3.74	5	0.588	-35	(-38; -32)			
Increase of the amount of enforcement								
< 200% (fixed effects)	0.30	3	0.961	-12	(-27; +8)			
> 200%								
Country								
Australia (fixed effects)	0.27	1	0.604	-33	(-38; -26)			
USA	4.37	6	0.627	-15	(-24; -4)			
Other	5.55	4	0.235	-27	(-38; -13)			
Study methodology								
Good study	7.86	8	0.447	-29	(-34; -24)			
Weak study	5.48	4	0.242	-34	(-46; -18)			

Comparing the summary effects between the subgroups of each of the potential moderator variables shows that larger accident reductions have been found when

- enforcement is signposted;
- enforcement is visible;

- there is no randomization (all results with randomization refer to Radar / laser US / AUS, which has been found to be less effective than other types of enforcement);
- when there is local publicity, compared to a publicity campaign or no publicity;
- when a new for of enforcement is introduced and when the intensity is increased by a large amount;
- in the UK compared to other countries;
- accidents are severe;
- in studies with a weak study design.

The amount of heterogeneity remains significant in almost all subgroups. This indicates that the results within each of the subgroups are affected by further moderator variables. When the subgroup analysis are conducted based on only injury accidents or only on fatal accidents, the results are not substantially different from the results that are based on all accident severities.

3.5 Factors predicting effects of speed enforcement: Results from meta-regression

In the subgroup analyses in the preceding section, the effect of only one variable at a time is investigated. None of the investigated moderator variables has been found to be sufficient to reduce the heterogeneity in the results. In meta-regression analysis it is investigated how the potential moderator variables affect the effectiveness of speed enforcement when controlling for all other potential moderator variables at the same time. The same potential moderator variables are included as predictors that have been used as grouping variables in the subgroup analyses in the preceding section.

In the meta-regression analysis all predictor variables are coded as dummy variables as shown in Table 7. Those dummy variables that refer to the same original variable are treated as sets of dummy variables, i.e. in the stepwise procedure all dummy variables that are based on the same original variable are included simultaneously in the model.
The meta-regression analysis is conducted in four steps:

Step (1) Stepwise forward analysis: The predictor variables are included in the regression model successively, one at a time. The first predictor is the one which produces the largest value of R^2 (variance explained by the model). The second predictor is the one which produces the largest value of R^2 when included in the model in addition to the first predictor. The third through seventh predictors are chosen accordingly. Table 7 shows the predictors in the order in which they were included in the model, and the values of R^2 that are obtained when the respective predictor and the preceding predictors are included in the model.

Step (2) Complete regression model with all predictor variables, based on all studies: The regression coefficients are estimated for the complete model which contains all predictor variables.

Step (3) Complete regression model with all predictor variables, based on the studies with a good design only.

Step (4) In a fourth step, coefficients are estimated for a partial model which is based on all results. In the partial model only the predictor variables are included that have been identified as relevant in the preceding steps. The regression coefficients and p-values of the coefficients are then compared to the coefficients in the complete model with all predictor variables.

Step (3) and (4) are conducted in order to provide a test of the consistency and robustness of the results. The results for step (1), (2) and (3) are shown in Table 7. The results from step (4) are not shown in table 7. Table 7 shows how each of the predictor variables is dichotomized as a dummy variable with the values 0 and 1. The categories for which the larger accident reductions have been found are underlined. Table 8 shows the model summaries of the regression models that have been estimated in step (2) and (3).

Table 7: Effects of speed enforcement on accidents, results from meta regression.

		Step (1)	Step (2)		Step (3)	
		R^2 when predictor	All results		Results from good study designs	
Predictor	Dummy variables (1; 0)	included into model	Coeff.	p(Coeff.)	Coeff.	p(Coeff.)
Type of enforcement	(Radar US/AUS; <u>other</u>)	0.096	0.224	0.050	0.142	0.026
	(<u>Speed cameras</u>; other)		-0.030	0.831	-0.036	0.674
	(<u>Composite</u>; other)		-0.090	0.442	-0.089	0.245
Signposted	(<u>Signposted</u>; other)	0.169	-0.231	0.130	-0.126	0.261
Accident severity.	(<u>Fatal/serious</u>; other)	0.207	-0.166	0.020	-0.166	0.001
Increase	(Incr.; <u>change of enf.</u>)	0.230	0.124	0.150	0.084	0.203
Method	(Good st.; <u>weak study</u>)	0.237	0.054	0.523		
Visibility	(<u>Visible</u>; other)	0.238	-0.133	0.117	-0.162	0.003
Country	(Australia; other)	0.240	-0.055	0.602	0.054	0.405
	(UK; other)		-0.051	0.679	0.162	0.050
	(US A; other)		-0.191	0.076	0.093	0.195
Publicity	(<u>Local publicity</u>; other)	0.230	-0.085	0.328	-0.004	0.949
	(<u>Pub. campaign</u>; other)		-0.028	0.807	-0.034	0.616

Table 8: Model summaries of the results from meta-regression.

Model	df	F	p(F)	R^2	Tau^2
Step (2) complete model based on all results	115	4.01	0.000	0.234	0.075
Step (3) complete model based on good study designs only	77	4.55	0.000	0.324	0.013

The results from meta-regression are only partly in accordance with the results from the subgroup analyses. Differences can be explained with the control for other variables in the regression model. When the effect of a variable changes when other variables are controlled for, it is likely that this effect has been, at least partly, due to confounding effects of the other variables.

- **Type of speed enforcement**: The type of enforcement is related to the effectiveness of the enforcement. The American / Australian type of radar / laser enforcement is less effective than the other methods. This result is consistent between the subgroup and the meta-regression analysis. Speed cameras and composite enforcement are, according to the results from meta-regression, not different from other types of enforcement. In the stepwise regression analysis The type of enforcement is the first variable that has been included in the regression model in the stepwise analysis.

- **Signposting**: Speed enforcement has been found to be more effective when signposted than when not signposted. The regression coefficient is not significant in the complete model with all predictor variables. In the partial models (step 4) however, which are based only on those predictor variables that have been identified in the stepwise procedure (step 1), the coefficient is somewhat larger (-0.337) and significant ($p = 0.002$).

- **Accident severity**: The effects of speed enforcement have been found to be greater for more serious accidents. The coefficient is significant both when all studies are included in the study and when the regression analysis is based on good studies only. In the partial models (step 4) the result is similar as in the complete model (coefficient = -0.184; $p = 0.007$). The result is also consistent with the subgroup analysis.

- **Increase vs. change of enforcement**: The effectiveness of enforcement has been found to be greater when a changed form of enforcement is introduced, compared to an increase of the amount of enforcement. When the type of enforcement is changed the intensity (e.g. in terms of person hours) if often increased as well. This result is consistent between the subgroup analysis and the meta regression analysis. The regression coefficients are however not significant. The increase in R^2 when including this predictor variable in the model in the stepwise procedure is smaller than for the previously introduced variables.

- **Method**: Greater effects of speed enforcement have been found in studies with a weak study design, compared to studies with a good study design. The regression coefficient is however not significant and the increase in R^2 when including this predictor variable in the model in the stepwise procedure is only small.

- **Visibility**: The visibility of speed enforcement has not been found to be relevant for the effectiveness of speed enforcement in the subgroup analysis. In the meta-regression analysis no significant effect has been found and R^2 is not improved when this variable is introduced in the model in the stepwise procedure. All the same, the visibility of enforcement may not be irrelevant since signposting has been fount to be a significant predictor for the effectiveness. Additionally, in the subgroup analysis, the effect of non-visible enforcement is far smaller than the effect of visible enforcement when it is controlled for publication bias.

- **Country**: The results for the effects of speed enforcement in different countries are inconsistent. In the meta-regression analysis the coefficients have the opposite sign in the model based on all studies and in the model based on studies with a good study design. It is also curious that all three dummy variables have the same sign, although the subgroup analysis would indicate that the effectiveness is smaller in Australia than in other countries and greater in the UK than in other countries. When adding the country dummy variables in the stepwise regression analysis, the predictive power of the model does not improve markedly. It is therefore concluded that the effectiveness of speed enforcement is not systematically different between different countries.

- **Publicity**: Publicity is, according to the results from meta-regression not a relevant predictor variable. In the subgroup analysis local publicity has been found to increase the effectiveness of enforcement, while publicity campaigns have been found to decrease its effectiveness. All except one studies of enforcement that is accompanied by a publicity campaign have used a “good” study design, which has been found to reduce the effectiveness that is found in a study. This may explain the findings that publicity campaigns seem to reduce the effectiveness of enforcement, but that this effect disappears when study design is controlled for.

3.5.1 Effects of potential moderator variables: Robustness of the results

On the whole, the results from the moderator analysis are not very robust against the selection of moderator variables included in the analysis and the studies included in the analysis.

The only moderator variables that have consistently been found to be relevant are accident severity and signposting. More severe accidents have been found to be more strongly affected by speed enforcement than less severe accidents, and signposting speed enforcement has been found to increase its effectiveness (even if the regression coefficients are not large or significant in all models).

For the potential moderator variable increase vs. change of enforcement, the effects are not inconsistent, but no large difference has been found.

For the other potential moderator variables, a number of inconsistencies have been found, amongst others:

- The type of speed enforcement that has been found to be more effective than others differs between all analyses. Only the American / Australian type of radar/laser enforcement has consistently been found to be less effective than other types of speed enforcement.

- Publicity may or may not be a relevant moderator. The results are most likely dependent on the control for other potential moderator variables.

- The results that refer to the effectiveness of speed enforcement in different countries are inconsistent between most analyses. A possible explanation is that there actually are no systematic differences between countries.

- The results indicate that visibility of enforcement increases its effectiveness. The results are however only consistent when publication bias in the results for non-visible enforcement are controlled for. An additionally analysis has shown that this finding is completely dependent on one outlier effect (an effect estimate with a statistical weight that exceeds all other statistical weights in this group of results by far). When this effect estimate is removed from the analysis, there is no longer any remarkable difference between visible and non-visible speed enforcement.
- Study methodology appears to be a relevant moderator in the subgroup analysis. However, the results from meta-regression do not provide strong support for this conclusion.

When comparing the regression models that have been developed in step 1 to 4, there is no strong consistency between the size and significance levels of the coefficients between the models. Some coefficients even change the sign between different models.

3.5.2 Effects of potential moderator variables: Summary

Moderator variables that have been found to be relevant are accident severity and signposting. More severe accidents have been found to be more strongly affected by speed enforcement than less severe accidents, and signposting speed enforcement has been found to increase its effectiveness.

For the other potential moderator variables that have been investigated the results are less clear. Visibility may increase the effectiveness of speed enforcement, and weak study designs may produce somewhat more favourable effects.

Whether or not the type of enforcement is changed and the country in which speed enforcement has been evaluated do not seem to be relevant moderators.

The results that refer to the type of enforcement and publicity are inconsistent. The only conclusion that can be drawn is that the American / Australian type of radar/laser enforcement seems to be less effective than other types of speed enforcement.

3.6 Limitations of speed enforcement studies

The results of the present analyses have shown that the results of speed enforcement may be affected by study quality. The results are however not quite consistent. In the meta-regression analysis, study quality is not always a significant predictor. This may be due to confounding effects between study quality and other predictor variables included in the meta-regression models. It is therefore not possible to draw any conclusions about the degree to which a weak study design leads to more favourable results. Additionally, even if over half of all studies have been classified as having applied a “good” study design, most of these studies have not controlled for regression to the mean, and only few studies have applied matched control groups or multivariate analyses in order to control for confounding variables. More strictly controlled studies may therefore find less favourable results than indicated by the present meta-analysis.

The results from the trim-and-fill analysis do not indicate that the results are affected by publication bias. When all results are regarded together, no publication bias seems to be present. In most of the subgroups a few new effect estimates have been generated. However, the number of generated effect estimates is for the most part only small (one or two) and the changes in the summary effects are not large or significant either. Only in the studies from USA and in the studies of non-visible enforcement (visible or unspecified visibility), the trim-and-fill analysis has led to marked reductions in the effectiveness of speed enforcement.

Almost all results from meta-analysis show significant amounts of heterogeneity. All the same, the effects of further differences between studies have not been investigated. Sooner or later one would find some subgroups of results without significant amounts of heterogeneity. It would however be questionable if this would lead to the detection of relevant moderator variables. Most likely, the results would be mere products of chance.

4. DRINK DRIVING ENFORCEMENT

4.1 Drink driving enforcement measures

Most studies of the effects of DUI enforcement on accidents have investigated the effects of checkpoints, and some studies have investigated the effects of patrolling.

4.1.1 Patrolling

Studies that have investigated the effects of patrolling are quite heterogeneous. The types of measures investigated range from a mere increase of the amount of patrolling to larger programmes where officers are trained in DUI apprehension and where other anti-DUI measures are implemented at the same time. Enforcement programmes including patrolling are STEP (Selective Traffic Enforcement Programme; Amick & Marshall, 1983; Sali, 1983) and ASAP (Alcohol Safety Action Projects; Zador, 1976). The measures have in common that most of them have been accompanied by publicity. Only in two studies it is not specified whether or not the measure was accompanied by publicity. All studies of patrolling have been conducted in 1981 or earlier. In later years, checkpoints became more common.

Although the patrolling measures are quite heterogeneous, there are only a few aspects the effects of which can be investigated in meta-analysis. These are:

- accident severity and type of accidents,
- study methodology.

How the effect of these factors is studied is described below. Other characteristics of patrolling studies can not be investigated in meta-analysis, partly because there is no enough detailed information available, and partly because there are too few studies with information available.

4.1.2 DUI-Checkpoints

DUI-checkpoint refers to all police operations where one or more police cars are standing beside the road (not driving) and where police officers pull out drivers in order to check whether or not a driver is has an illegal BAC level. According to this definition, checkpoints may vary with respect to

- how large and how visible DUI-checkpoints are,
- to what degree DUI-checkpoints are conducted at random times or places or only e.g. on roads and at times with high frequencies of DUI accidents,
- whether all drivers are stopped at the DUI-checkpoint as far as the capacity of the checkpoints allows, or if only some drivers are stopped,
- when not all drivers are stopped, whether drivers are stopped randomly or on suspicion only,

- whether all drivers who are stopped are being tested e.g. with a passive alcohol sensor or a breath test, or if only those who are suspected of drink-driving are tested, and how this suspicion is established.

In many countries, there are DUI-checkpoint programmes where the checkpoints are combined with media campaigns. In many of these programmes, the checkpoint operations are conducted in a very specific way, e.g. with a specific type of cars or buses (e.g. “booze buses”), and known under specific names, e.g. Random Breath Testing (RBT, Australia) or Compulsory Breath Testing (CBT, New Zealand). Checkpoints are however also conducted in other countries and terms like RBT are not consistently used for the exactly same type of checkpoints. In order to avoid confusion, terms like RBT and CBT are not used in this paper, except when they refer to the specific programmes.

Ideally, a meta-analysis of the effects of DUI-checkpoints on accidents would investigate the effect of a number of properties of the DUI-checkpoints, such as the procedures for stopping and testing drivers or the increase and total amount of checkpoints and DUI-testing. However, information about properties of DUI-checkpoints is in most studies only sparse. There are only few characteristics on which information is available from a sufficient number of studies in order to investigate them by means of meta-analysis. These characteristics of DUI-checkpoints are:

- Testing of drivers who are stopped at the checkpoint
- Accompanying publicity
- New type of or intensified enforcement
- Country

Additionally, available information allows the study of the following factors in meta-analysis:

- Injury severity and types of accidents
- Time period over which DUI-checkpoints are conducted and evaluated
- Study methodology

These factors are described in section 3.3. Other characteristics of DUI-checkpoint operations can also be assumed to affect the effectiveness, such as the randomness (unpredictability) of when and where checkpoints are conducted, the total amount of enforcement (e.g. the numbers of drivers tested per licence holder in a region or country), whether or not drivers are stopped randomly or non-randomly, the type of measurement equipment that is used, the type of punishment for DUI offences, amongst other things. Such factors can however not be investigated in the present meta-analysis because the information that is provided from the studies is not sufficient. All the same, it has to be taken into account that such factors may affect the effectiveness of checkpoints and that they differ between different DUI-checkpoint operations. In other words, these factors must be regarded as uncontrolled confounding factors which may limit the strength of the conclusions that can be drawn from the results of the meta-analysis and the degree to which the results can be generalized.

4.2 Method

4.2.1 Description of drink driving enforcement evaluation studies

In task 4.2 a total of 49 evaluation studies were identified and found to be of an acceptable quality to be included in the data. Of these, 9 studies have evaluated patrolling and 40 studies have evaluated DUI checkpoints.

The 49 studies have been published in 8 countries and comprise a total of 143 results. Most studies are from USA or Australia.. Table 9 presents the number of results according to country.

Table 9. Amount of research evaluating safety effects of DUI enforcement measures. Number of results according to country.

Country	**Number of studies of DUI enforcement**
Australia	19
USA	16
New Zealand	6
Canada	2
Sweden	2
UK	2
France	1
Netherlands	1
Total	49

The distribution according to year of publication (table 10) shows that about half of the studies are quite recent as they have been published in the 1990s or after 2000.

Table 10. Distribution of studies evaluating safety effects of DUI enforcement measures. Number of results according to year of publication.

Year of publication	**Number of studies of DUI enforcement**
1970 – 1979	3
1980 – 1989	22
1990 – 1999	18
2000-2008	6
Total	49

4.2.2 Factors affecting the effectiveness of DUI enforcement

A number of potential moderator variables is investigated in meta-analysis, i.e. variables that are assumed to affect the effectiveness of DUI enforcement measures. Moderator analyses are conducted in two ways: In subgroup analyses and in meta-regression.

In subgroup analyses studies are divided into subgroups according to a potential moderator variable. The summary effects of the subgroups of studies are then compared. In most cases there are only two subgroups. An important point in subgroup analysis is to keep the subgroups comparable with respect to other comparable moderator variables. Subgroups analyses are mostly only made when there is significant heterogeneity in the results for a group of effect estimates (e.g. all injury accidents). Heterogeneity indicates that the results are affected by some factor or factors that are different between the studies. If there is no heterogeneity, all effect estimates can be regarded as representatives of the same outcome and no systematic differences between studies can be expected.

A grouping variable is assumed to be a relevant moderator when two criteria are fulfilled (Christensen, 2004):

- the summary effects for the subgroups are different,
- heterogeneity within the subgroups is reduced as compared to the overall summary effect.

The second criterion is however not a very strict criterion. When it is not fulfilled (one finds for example frequently reduced heterogeneity only in one of the subgroups), this may indicate that the moderator variable is not relevant. It is however also possible that the moderator variable all the same is relevant, but that the results within one or more subgroups are affected by further moderator variables. This can be investigated by further subgroup analyses or by meta-regression.

In a meta regression analysis, the effect of several potential moderator variables is investigated at the same time (see chapter 4.5).

Testing of drivers who are stopped at the checkpoint: Based on available information, all evaluation studies are classified according to whether or not all drivers who are stopped are tested.

- **All drivers tested**: At DUI-checkpoints where all drivers are tested either a passive sensor is used or a breath test is taken. Only studies where it is explicitly stated that all drivers are tested are included in this group.

- **Not all drivers tested**: In some checkpoint programmes it is either not required that all drivers are tested, or tests are not taken from all drivers even if it may be required. Many study reports do not provide quite clear information about whether or not all drivers are tested. These studies are also classified in this group.

The distinction of whether or not all drivers are tested is not always quite clear. Firstly, available information is not always very precise. Secondly, it is not always easy to classify studies, e.g. testing all drivers may be officially required but not carried out in practice. There may also be local variations within one checkpoint program, i.e. testing procedures may differ not only between but also within studies.

This variable is not relevant to the meta-analysis of patrolling studies.

Accompanying publicity: DUI-Checkpoint programmes are classified according to the type of accompanying publicity:

- **Publicity campaign:** Publicity with paid media, i.e. advertising in newspaper or TV-spots; publicity campaign with paid media may also include non-paid publicity.
- **Unpaid publicity**: Publicity without paid media, e.g. information in regular newspaper or TV coverage, posters at the road side.
- **No publicity**: No activities are conducted in order to achieve publicity for the DUI-checkpoints.

DUI-checkpoint programmes classified as "No publicity" can be assumed to have a very low level of publicity, e.g. some newspaper notices. Publicity campaigns with paid media are often associated with a high intensity of publicity. This is however not always true, even without paid media, DUI-checkpoints may be highly publicized. The intensity or amount of publicity can not be investigated in this meta-analysis because information is either not available, or, when available, not comparable between studies.

Publicity may be relevant to the effects of patrolling. No moderator analysis can however be made because most patrolling measures have been accompanied by publicity and there is not enough information available in order to compare the amount of publicity between studies.

Change of type or amount of enforcement: Many evaluations of DUI-checkpoints have been carried out in order to evaluate the effect of introducing DUI-checkpoints as a new type of measure. In some cases the effects of an increase of the amount of enforcement, e.g. increased number of checkpoints, have been studied. The studies have therefore been classified as

- New type of enforcement or
- Increased enforcement.

Country: DUI-checkpoint operations often follow certain procedures and use specific equipment in different countries. Effects of DUI-checkpoints are therefore compared between different countries. According to the availability of studies, the following (groups of) countries are compared:

- Australia
- New Zealand
- USA
- Other

Accident severity and types of accidents: In most studies it is specified if accident numbers refer to injury accidents or to all accidents including property damage only accidents. Some studies have reported results for different degrees of accident severity, e.g. fatal and injury accidents. Accordingly, in the meta-analysis effects are estimated for the following groups of accident severity:

- **Unspecified severity**: Results where accident severity is not specified (these include most likely injury and property damage only accidents), and where it is specified that accident and property damage only accidents have been investigated. There are only few studies in this group.
- **Injury accidents**, mostly including fatal accidents.
- **Fatal accidents**.

The accidents that have been investigated vary between different studies. Many studies have aimed at investigating effects on accidents involving alcohol. Since precise information on the BAC of drivers involved in accidents is not always available, many studies have used some substititute measure for alcohol accidents. Mostly, weekend night accidents have been used as a substitute measure for accidents involving alcohol. Among weekend night accidents, alcohol involvement is known to be far higher than in accidents at other times. It must however be taken into account that not all accidents at weekend nights involve alcohol. Many studies have investigated effects on all accidents, with and without alcohol involvement, and a few studies have investigated effects on daytime accidents. In short, the studies are grouped as follows:

- **Accidents involving alcohol**
- **All accidents**
- **Daytime accidents**

A few studies have investigated effects on specific types of accidents, e.g. single vehicle accidents, accidents in rural vs. urban areas, or accidents on minor vs. major roads. There are however too few effect estimates available for each of these specific accident types, and effects on different accident types are therefore not investigated in meta-analysis.

Time period over which DUI-checkpoints are conducted and evaluated: Studies vary with respect to the study period after DUI-checkpoints have been introduced or after their intensity has been increased. Studies are classified as follows:

- **3 months** (1 to 92 days from introduction)
- **between 3 and 6 months** (93 to 182 days from introduction)
- **between 6 months and 1 year** (183 to 365 days from introduction)
- **between 1 and 2 years** (366 to 730 days from introduction)
- **between 2 and 4 years** (731 to 1460 days from introduction)
- **between 4 and 8 years** (1461 to 2920 days from introduction)

Study methodology: Numerous different study methodologies have been used to evaluate the effects of DUI-checkpoints. Most studies have compared accident numbers before and after the introduction of a measure. These studies differ with respect to the type of control group that has been used as a comparison. Control groups that have been used are:

- **No control group**: The number of accidents before and after the introduction of the measure are compared and no other control than the time period before and after is used. In these studies changes in accident numbers may be affected by all other factors that have changed at the same time or in the same time period as the measure has been introduced and conducted.

- **Trend only controlled**: Many studies have applied time series models in order to investigate to what degree accident numbers after a measure has been introduced deviate from the accident numbers that would have been expected if the trend in the accident numbers before the introduction of the measure had continued. These studies are better controlled than studies without comparison group, the results may however be affected by other changes that occurred at the same time as the introduction of the measure.

- **Other types of accidents as comparison group**: A number of studies aimed at investigating effects on accidents involving DUI. Most of these studies have used some kind of substitute for DUI accidents since no or not sufficient information about alcohol involvement in accidents was available from accident statistics. Substitutes are types of accidents for which it is known that alcohol is often involved, for example weekend night accidents or single vehicle night accidents involving young male drivers. The comparison groups in these studies consist of accidents that did not involve alcohol or that are not in the group of alcohol substitute accidents. In these studies other factors that affect general accident numbers are to a certain degree controlled for, this is however not true of other factors affecting DUI accidents. Additionally, the validity of the definition of alcohol substitute accidents is only seldom proven and these accidents will always include a proportion of accidents not involving alcohol. Conversely, a proportion of accidents that are treated as non-alcohol accidents is likely to involve alcohol.

- **Comparison area**: Some studies where DUI-checkpoints were only used in a certain area, e.g. a region, city or state, have used accidents in other areas (regions, cities or states) as a comparison. In these studies only factors affecting general accident trends are to a certain degree controlled for. Factors that are specific to the study area where the measure has been introduced however are not controlled. Changes of factors that affect accidents that occur only in the control area, but not in the study area, may also affect the results.

- **Multivariate studies**: A few studies have applied multivariate models in order to examine the contribution of DUI-checkpoints and numerous other factors on accident numbers. Factors controlled for vary between studies. Among them are time trends, seasonal changes, alcohol sales, unemployment, and speed enforcement.

Since the classification of studies into 5 groups would result in quite few effect estimates within each group, the 5 types of study design are roughly classified as “good” and “weak” study designs.

- **Weak study designs** are studies without a comparison group and studies that have controlled for time trends only.

- **Good study designs** are the remaining types of study design. All of these studies have used a control group or controlled for potential confounding variables by means of multivariate methods.

Both within the "good" and the "weak" study designs, study quality varies, dependent for example on the choice of comparison group, the statistical methods applied, and the quality of the available accident data. However, it can be assumed that most studies that are classified as "good study design" actually have controlled for confounding variables to a larger extent than studies classified as "weak study design".

4.3 Results of meta-analysis of drink driving measures: Patrolling

4.3.1 Studies included in the meta-analysis

An overview of the available studies that are included in the meta-analysis is shown in Table 11. The classification of injury severity and study design is made as described in the section above. The sum of the statistical weights indicates the size of the study in term of the number of accidents that is included in the effect estimates that are included in the meta-analysis.

Table 11: Studies of the effects of patrolling on accidents.

Authors	Year	Country	Type of measure	Study design	Sum of statistical weights[1]
Stuster & Blowers	1995	USA	Patrolling by special DUI patrols	BA only trend controlled	115
Voas & Hause	1987	USA	Patrolling by officers trained in DUI apprehension	BA other accidents as control	848
Amick & Marshall	1983	USA	STEP programme	BA other accidents as control	149
Sali	1983	USA	STEP programme	BA only trend controlled	2,597
Wolfe	1983	USA	Patrolling by officers trained in DUI apprehension	BA with comparison area	15,750
Hurst & Wright	1981	New Zealand	Increased patrolling	BA no comparison	803
Ross	1977	UK	Patrolling and increased DUI testing	BA no comparison	27
Zador	1976	USA	ASAP programme	BA with comparison area	1,170
Toomath	1974	New Zealand	Increased patrolling	BA only trend controlled	15

[1] Statistical weights in fixed effects model

4.3.2 Plot of studies and effects

A plot of studies and effect estimates is shown in Figure 2.

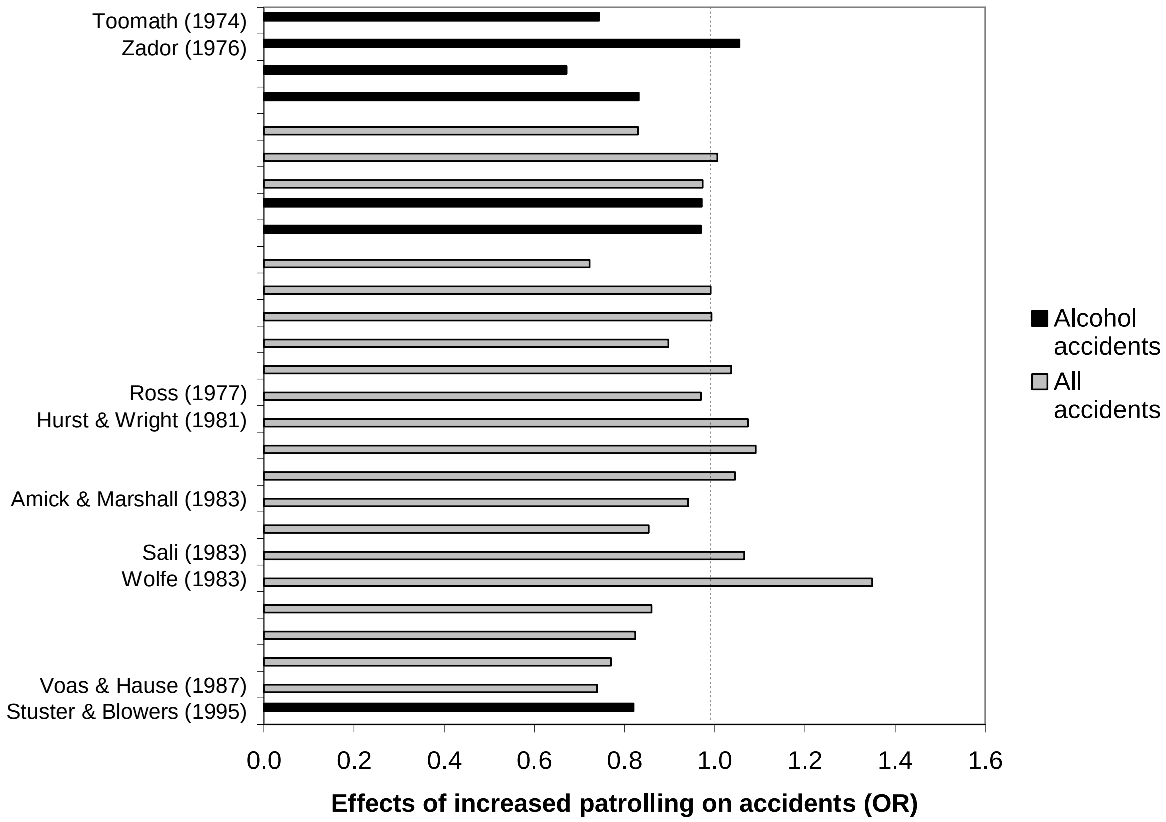

Figure 2: Plot of studies and effect estimates (odd ratios) of patrolling studies.

4.3.3 Data inspection, publication bias

A scatterplot of all effect estimates for patrolling is shown in Figure 3. It is indicated in the figure if the effect estimates refer to a good or a weak study design. The summary effects (Fixed Effects) are indicated by a dotted line for the weak study designs and by a straight line for the good study designs. The natural logarithm of the overall summary effect is -0.045, which corresponds to an odds ratio of 0.956. One of the effect estimates that is based on a good study design has a far larger statistical weight than all other effect estimates. This effect estimate is lying far outside the figure (as indicated by the arrow).

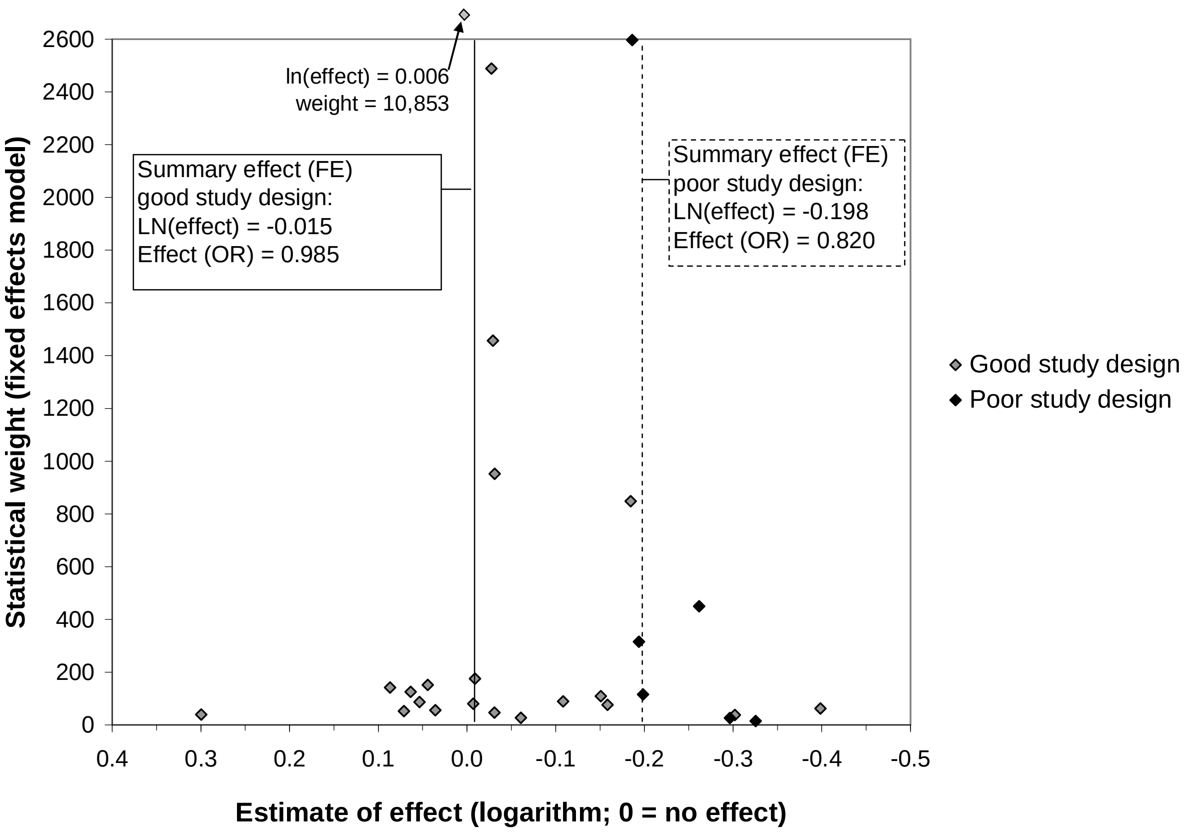

Figure 3: Scatterplot of effect estimates for patrolling.

A visual inspection of figure 3 does not reveal any clear funnel pattern. The effect estimates from studies with a weak design are however lying further to the right, i.e. show larger accident reductions, than the effect estimates that are based on good study designs. A trim and fill analysis is conducted with all effect estimates and with the effect estimates from good study designs only (from weak study designs there are not sufficient effect estimates for a trim and fill analysis). None of the trim and fill analyses generates new effect estimates. There is no indication that the results are affected by publication bias. This result may however be due to the heterogeneity in the result and the lack of a funnel-like distribution. Even if there were publication bias it is likely that it would remain undetected.

4.3.4 Summary effects and tests for heterogeneity

Results from tests of heterogeneity and summary effects of the estimated effects of patrolling on accidents are shown in Table 12.

A comparison between the results from good and weak study designs shows that a significant accident reduction only was found in studies with a weak design, not in studies with a good design. There remains significant heterogeneity in the results that are based on a good study design, which indicates the presence of further moderator variables.

For injury / unspecified accidents, effect estimates are available for all types of accidents (day and night) and for night-time accidents. Table 12 shows the results that are based on studies with a good design only. For both groups of effect estimates the are significant amounts of heterogeneity and the summary effects are almost identical. When the effect estimate that has a statistical weight of over 10,000 that is based on the study by Wolfe (1983; see section above) is omitted from the analysis, the results do not change significantly. Due to the large amount of heterogeneity in the results, the statistical weights only have a small effect on the summary effects in the Random Effects model.

For fatal accidents, no significant effect on accidents has been found. There is no significant amount of heterogeneity in these results. However, 13 of the 14 available effect estimates are from the study by Zador (1976). The one other effect estimate is based on the study by Hurst & Wright (1981). Both studies have used accidents in control areas as comparison group and are thereby classified as "good study design".

Table 12: Test of heterogeneity and summary effects (Random Effects models) of patrolling.

	Test of heterogeneity			Change of number of accidents (%)	
Type of accidents affected	Cochran's Q	df	p	Summary effect	95% confidence interval
All results					
All accidents	156.41	26	0.000	-8	(-12; -3)
Good vs. weak study design					
All accidents, good study design	54.38	20	0.000	-4	(-8; +0)
All accidents, weak study design	2.680	5	0.749	-24	(-28; -20)
Injury, fatal and night time accidents (good study designs only)					
Injury accidents / unspecified severity; all types of accidents	39.62	6	.000	-6	(-11; 0)
Injury accidents / unspecified severity; all types of accidents – outlier omitted[1]	25.98	5	.000	-7	(-14; -1)
Injury accidents / unspecified severity; night time accidents	22.97	4	.000	-9	(-17; 0)
Fatal accidents; all types of accidents	14.71	13	.325	-1	(-7; +5)

[1] The effect estimate that has a statistical weight of over 10,000 that is based on the study by Wolfe (1983) has been omitted.

In the present meta-analysis the overall effect of patrolling is estimated to a reduction in the number of accidents (all severities) of 8% (-12;-3). However, when excluding studies with a weak study design the effect is down to a marginal 4% (-8;+0), while an estimate based on weak designs only is as high as -24% (-28;-20). However, this result is found when there is no control for confounding variables. The overall impression from patrolling is that there is some reduction in the number of accidents, but the effect is rather small. Excluding studies with "Unspecified accident types", i.e. studies that may comprise property-damage-only accidents, the cultivated effect on injury accidents only is a reduction of 11% (-16;-4) while there is no effect on fatal accidents.

4.4 Results of meta-analysis: DUI-checkpoints

4.4.1 Studies included in the meta-analysis

An overview of the available studies that are included in the meta-analysis is shown in Table 13. The classification of injury severity and study design is made as described in the section above. Several studies are form the same period and same country. The results that are included in the meta-analysis have however been selected so as to avoid double coding. When a measure has been evaluated by two or more studies in the same region at the same time, the results from the study with the better study design have been selected. This has not led to the exclusion of whole studies, only some individual effect estimates have been excluded from the meta-analysis because of double coding (the excluded results are not included in the overview in table 13). The sum of the statistical weights indicates the size of the study in term of the number of accidents that is included in the effect estimates that are included in the meta-analysis.

Table 13: Studies of the effects of DUI checkpoints on accidents.

Authors	Year	Country	Type of measure	Study design	Sum of statistical weights[1]
Fell, Langston & Tippetts	2005	USA	Checkpoints	BA other acc as control	814
Tay	2005	Australia	Checkpoints	multivariate	2,939
Mathijssen & de Craen	2004	Netherlands	Checkpoints	BA with comparison area / sites	174
Miller, Blewden & Zhang	2004	New Zealand	Checkpoints, all drivers tested	BA only trend controlled	534
Agent, Green & Langley	2002	USA	Checkpoints	BA no comparison	433
Lacey & Jones	2000	USA	Checkpoints	BA only trend controlled	73
Diamantopoulou & Cameron	1998	Australia	Checkpoints	BA other acc as control	336
Newstead, Cameron & Narayan	1998	Australia	Checkpoints	multivariate	3,942
Cameron, Diamantopolou et al.	1997	Australia	Checkpoints	multivariate	423
Henstridge, Homel & Mackay	1997	Australia	Checkpoints	multivariate	6,288
Holder, Voas & Gruenwald	1997	USA	Checkpoints	BA with comparison area / sites	371
Lacey, Jones & Fell	1997	USA	Checkpoints, all drivers tested	BA no comparison	670
Ryan, Hendrie & Allotey	1997	Australia	Checkpoints	BA no comparison	86
Mara, Davies & Frith	1996	New Zealand	Checkpoints, all drivers tested	BA only trend controlled	3,312
Mercer, Cooper & Kristiansen	1996	Canada	Checkpoints	BA with comparison area / sites	102
Bailey	1995	New Zealand	Checkpoints, all drivers tested	BA other acc as control	5,594
Jones, Joksch, Lacey et al.	1995	USA	Checkpoints, all drivers tested	BA with comparison area / sites	134
Stuster & Blowers	1995	USA	Checkpoints	BA only trend controlled	796
Törnros	1995	Sweden	Checkpoints	BA with comparison area / sites	1,458
Cameron, Cavallo & Sullivan	1992	Australia	Checkpoints	multivariate	1,141

Authors	Year	Country	Type of measure	Study design	Sum of statistical weights[1]
Wells, Preusser & Williams	1992	USA	Checkpoints, all drivers tested	BA other time period as control	839
Evans	1990	USA	Checkpoints	multivariate	2,818
Smith, Maisey & McLaughlin	1990	Australia	Checkpoints	BA other acc as control	993
Barnes	1988	Australia	Checkpoints	BA with comparison area / sites	4,498
Homel	1988	Australia	Checkpoints	BA with comparison area / sites	1,772
King	1988	Australia	Checkpoints	BA no comparison	206
Derby & Hurst	1987	New Zealand	Checkpoints	BA only trend controlled	2,672
Frank	1986	Australia	Checkpoints	BA other time period as control	884
Åberg, Engdahl & Nilsson	1986	Sweden	Checkpoints, all drivers tested	BA other acc as control	33
Armour	1985	Australia	Checkpoints	BA other acc as control	182
Hardes et al	1985	Australia	Checkpoints	BA only trend controlled	759
L'Hoste, Duval & Lassarre	1985	France	Checkpoints	BA with comparison area / sites	490
Mercer	1985	Canada	Checkpoints, all drivers tested	BA no comparison	261
Voas, Rhodenizer & Lynn	1985	USA	Checkpoints, all drivers tested	BA other acc as control	306
Kearns	1984	Australia	Checkpoints	BA other time period as control	1,776
McLean, Clark, Dorsch, Holubowycz, & McCaul	1984	Australia	Checkpoints	BA other time period as control	496
Thomson & Mavrolefterou	1984	Australia	Checkpoints	Multivariate	367
Cameron & Strang	1982	Australia	Checkpoints	BA other acc as control	63
Ross	1982	UK	Checkpoints	BA with comparison area / sites	99
Cameron, Strang & Vulcan	1981	Australia	Checkpoints	BA other time period as control	150

The studies have been classified as “good design” vs. “weak design”. The following Table 14 summarizes the characteristics of the DUI checkpoints and the studies according to good vs. weak study design. Good study designs are overrepresented in Australian studies (75%), while weak study designs are overrepresented in studies from New Zealand. Weak study designs are also overrepresented among studies that have investigated the effects of DUI checkpoints during the first 6 months.

Table 14: Studies according to good vs. weak study design.

	Good study design		Weak study design		Total
	N	%	N	%	N
Australia	45	76 %	14	24 %	59
New Zealand	5	25 %	15	75 %	20
USA	15	56 %	12	44 %	27
Other	8	80 %	2	20 %	10
Publicity	18	64 %	10	36 %	28
Publicity campaign	44	59 %	31	41 %	75
Unspecified	11	85 %	2	15 %	13
< 6 months	9	35 %	17	65 %	26
> 6 months	60	70 %	26	30 %	86
Unspecified	4	100 %	0	0 %	4
Fatal	24	77 %	7	23 %	31
Injury	44	58 %	32	42 %	76
Unspecified	5	56 %	4	44 %	9
All accidents	21	70 %	9	30 %	30
Daytime accidents	3	50 %	3	50 %	6
Accidents involving alcohol	49	61 %	31	39 %	80
Total	73	63 %	43	37 %	116

4.4.2 Plot of studies and effects

A plot of studies and effect estimates is shown in Figure 4 and 5.

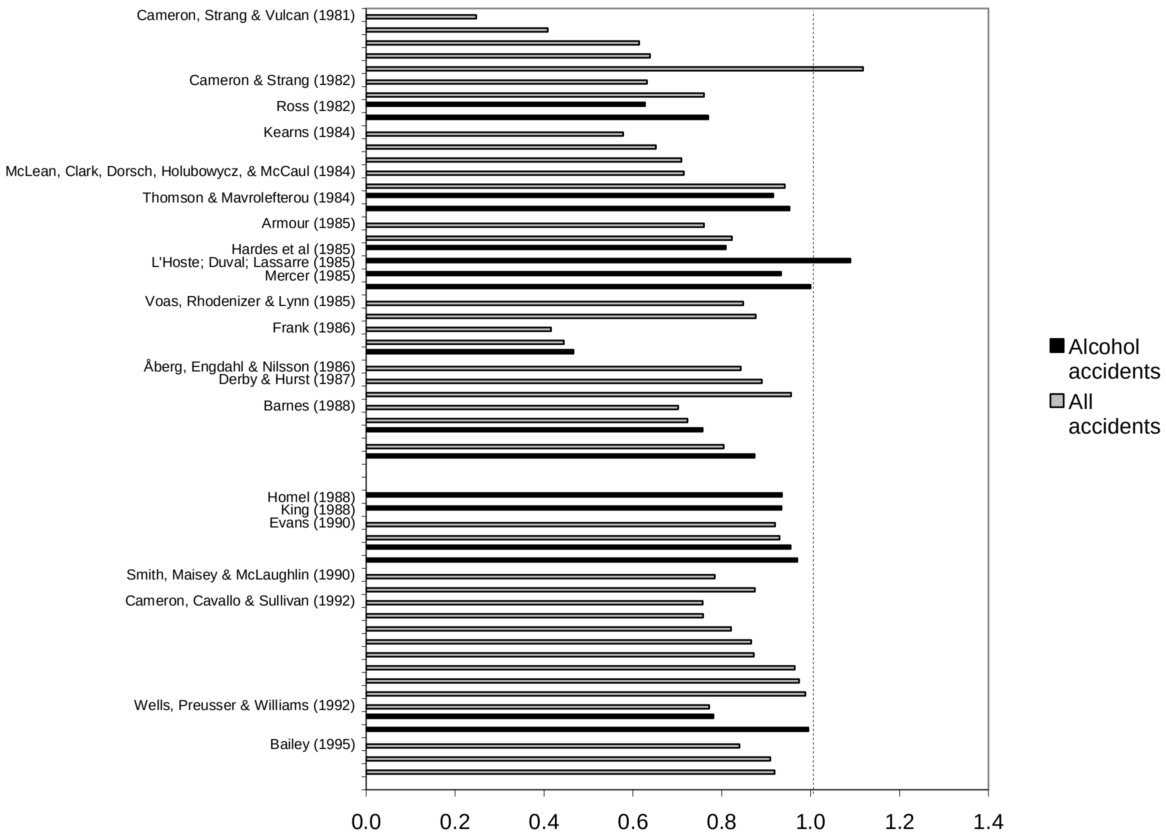

Figure 4: Plot of studies and effect estimates of patrolling studies.

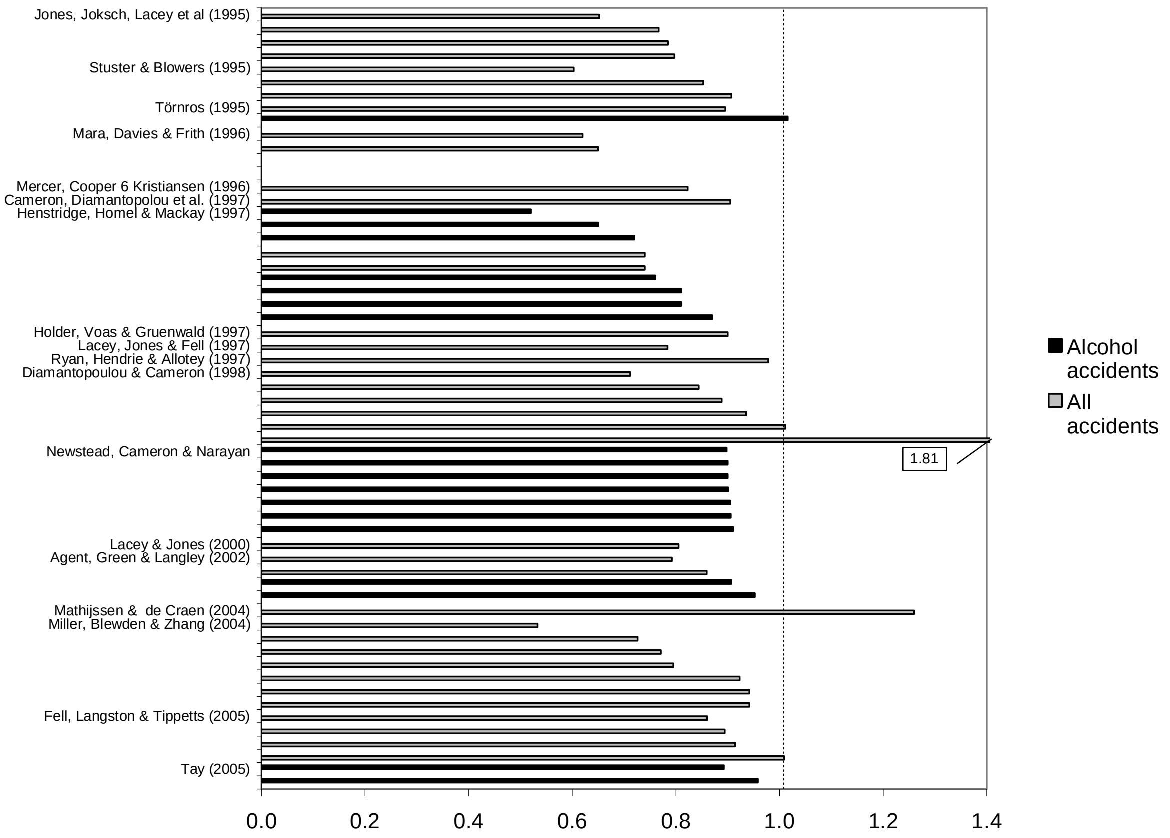

Figure 5: Plot of studies and effect estimates of patrolling studies.

4.4.3 Data inspection, publication bias

A scatterplot of all effect estimates for DUI checkpoints is shown in Figure 6. It is indicated in the figure if the effect estimates are from studies that have been classified as “good” or as “weak”. The summary effect (FE) for all effect estimates is indicated by a dotted line.

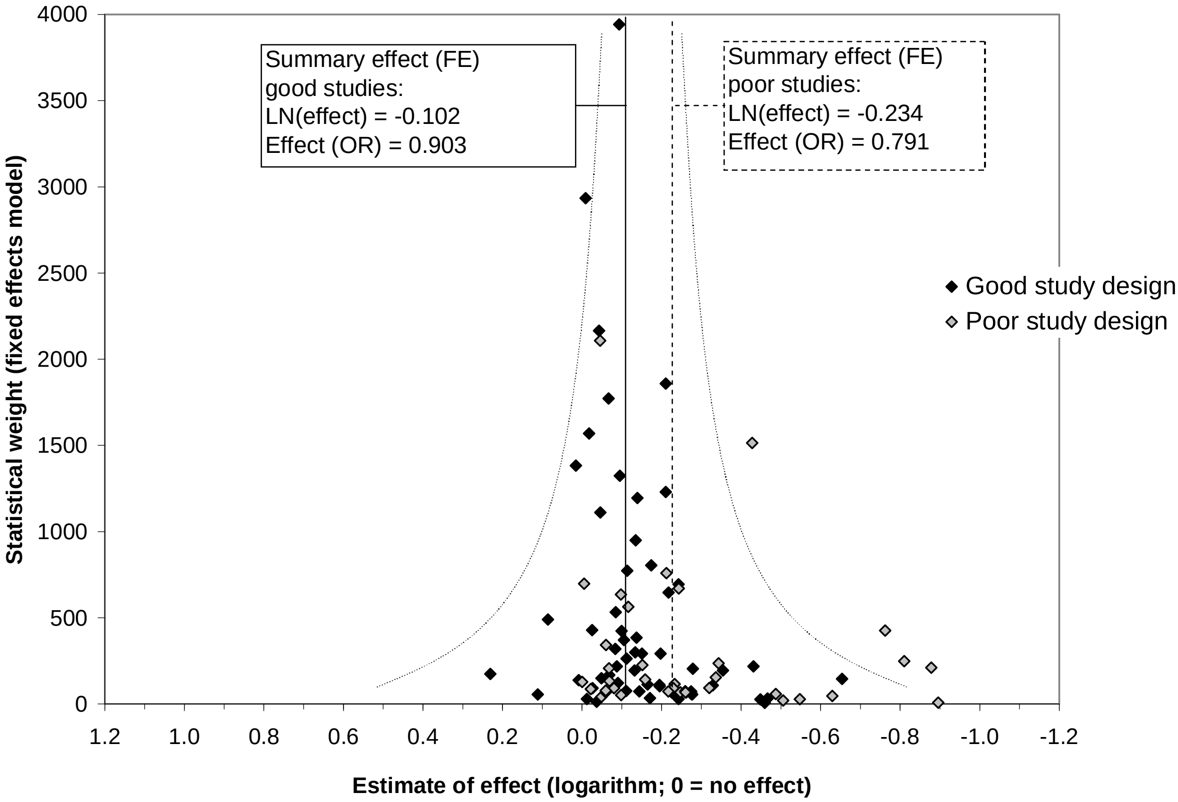

Figure 6: Scatterplot of effect estimates for DUI-checkpoints.

The effect estimates appear to be somewhat asymmetrically distributed. The effect estimates with the largest weights are further to the left than the summary effect, i.e. showing smaller accident reductions or increases of accident numbers. Among the effect estimates with the smallest statistical weights, large negative values, i.e. large accident reductions, are somewhat overrepresented. This is the pattern that would be expected when the results are affected by publication bias. The same asymmetry can be seen both among the effect estimates from good and from weak study designs.

Therefore, trim and fill analyses are conducted, first for all effect estimates, and then for the effect estimates from the good study designs only. The results of the trim and fill analysis that is based on the results from all studies are shown in Figure 7. In the trim and fill analysis 7 new effect estimates have been generated and the summary effect has a smaller negative value (smaller accident reduction) than the summary effect that is based on the original effect estimates only.

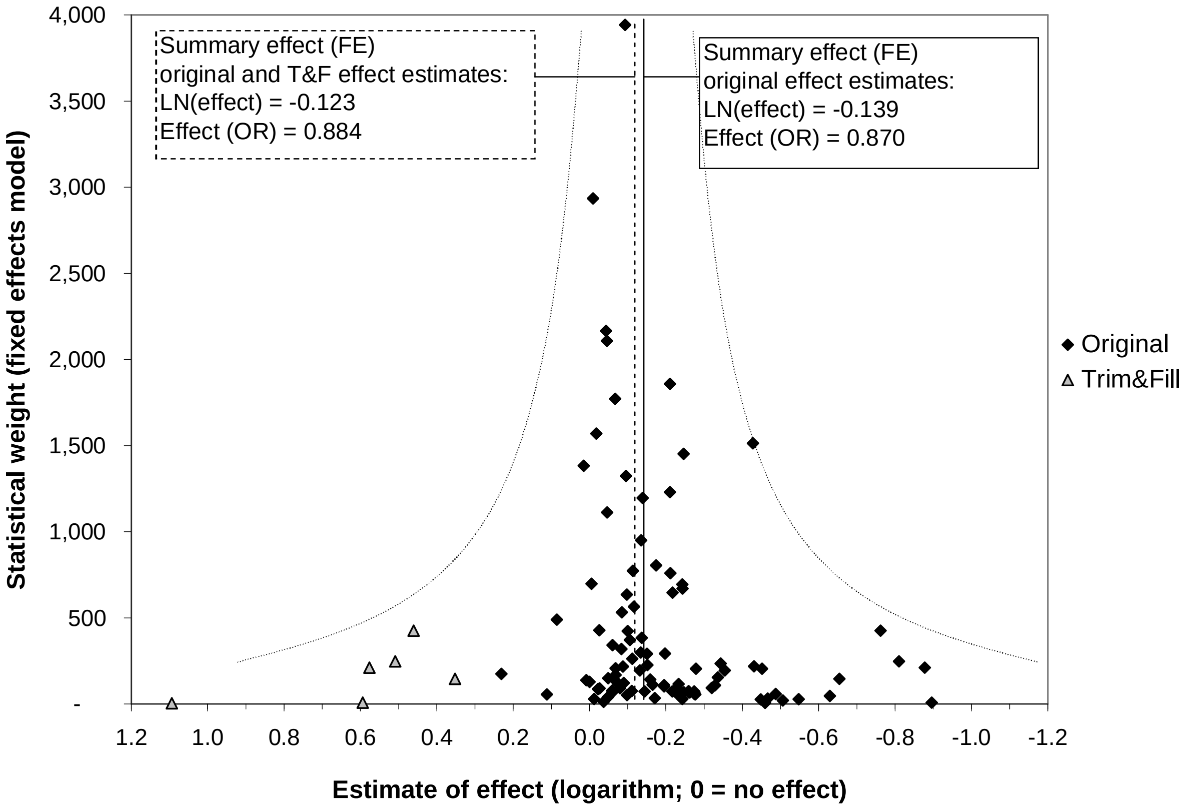

Figure 7: Scatterplot of effect estimates for DUI-checkpoints (all studies), trim and fill analysis.

The results from the trim and fill analysis that has been conducted based on the effect estimates from good study designs only are shown in Figure 8. In this analysis 12 new effect estimates have been generated. The summary effect lies further to the left and is closer to zero.

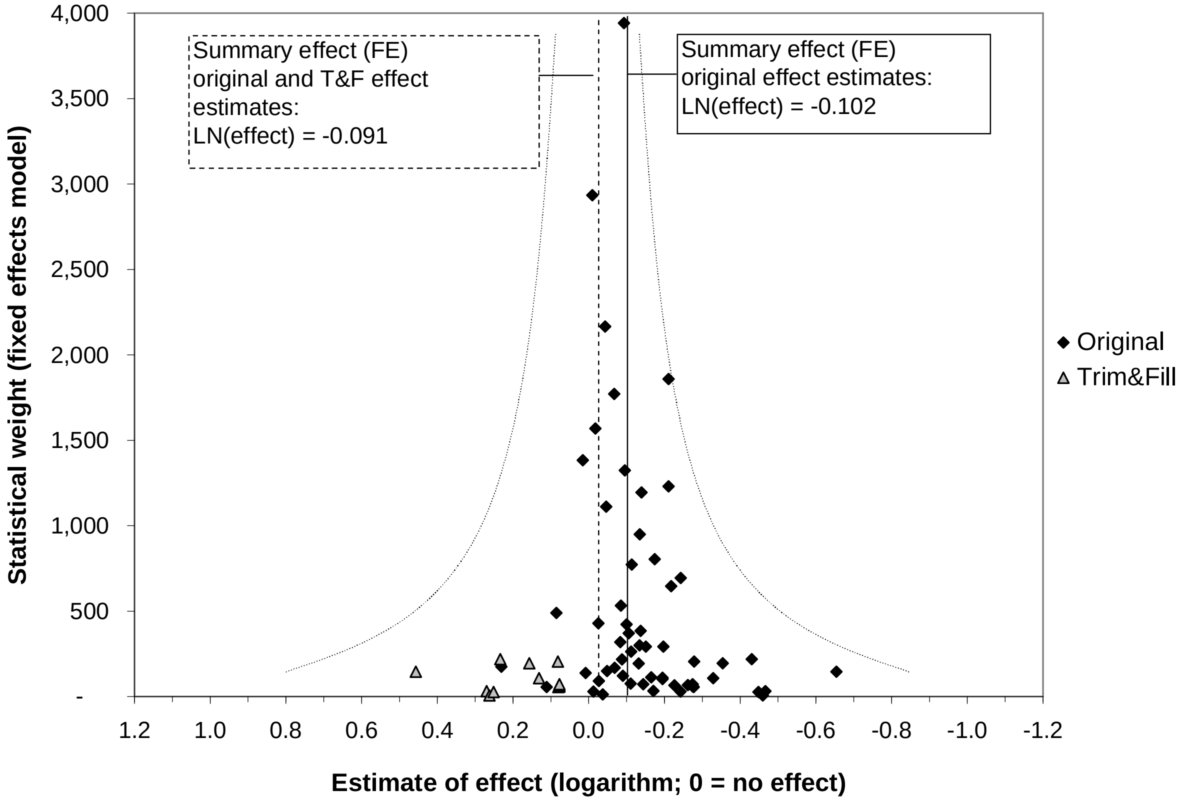

Figure 8: Scatterplot of effect estimates for DUI-checkpoints (good study designs only), trim and fill analysis.

These results from the trim and fill analysis indicate that the results for DUI checkpoints are affected by publication bias. The results that will be presented in the next section will therefore also be adjusted for publication bias.

4.4.4 Summary effects and tests for heterogeneity

Results from tests of heterogeneity and summary effects of the estimated effects of DUI checkpoints on accidents are shown in Table 15. The overall result i a significant reduction of the number of accidents. When the results are adjusted for publication bias in trim and fill analyses the summary effects become smaller, but all of them remain significant.

Table 15: Test of heterogeneity and summary effects (Random Effects models) of DUI checkpoints.

	Test of heterogeneity			Change of number of accidents (%)		Trim and fill analysis: Change of number of accidents (%)		
	Cochran's Q	df	p	Summary effect	95% confidence interval	Summary effect	95% confidence interval	Number of new effect estimates
All results	1007.221	96	0.000	-17	(-20; -14)	-15	(-18; -11)	7
Injury accidents	841.240	61	0.000	-19	(-22; -15)	-16	(-20; -11)	3
Fatal accidents	139.043	27	0.000	-15	(-19; -10)	-6	(-8; -5)	11
Unspecified severity	10.276	7	0.173	-11	(-16; -5)			

There are significant amounts of heterogeneity in all results. Moderator variables are therefore likely to be present.

4.4.5 Effects of potential moderator variables

Subgroup analyses: The effects of a number of potential moderator variables are investigated by comparing summary effects between subgroups of results. It has been found that the results are likely to be affected by publication bias. For all summary effects that are based on at least 10 effect estimates trim and fill analyses are therefore conducted. The results are summarized in Table 16a. The subgroup analyses have also been conducted based on injury and on fatal accidents. The results are shown in table 16b and 16c, respectively.

Table 16a: Test of heterogeneity and summary effects (RE models) of DUI checkpoints – moderator analyses: Results for all accident severities.

	Test of heterogeneity			Change of number of accidents (%)		Trim and fill analysis: Change of number of accidents (%)		
	Cochran's Q	df	p	Summary effect	95% confidence interval	Summary effect	95% confidence interval	Number of new effect estimates
Time period studied								
< 3 months	252.305	19	0.000	-29	(-39; -18)	-22	(-34; -7)	4
3 – 6 months	116.778	5	0.000	-21	(-36; -3)			0
6 – 12 months	164.274	35	0.000	-14	(-17; -11)	-13	(-16; -9)	4
12 – 24 months	92.7796	17	0.000	-13	(-19; -6)	-9	(-16; -1)	3
Country								
Australia	672.375	48	0.000	-22	(-26; -17)	-17	(-22; -11)	7
New Zealand	118.643	15	0.000	-14	(-19; -9)	-7	(-13; -1)	6
USA	38.1913	22	0.017	-12	(-15; -8)	-8	(-12; -4)	9
Other countries	30.8301	9	0.000	-4	(-13; +5)	+4	(-6; 14)	4
Good vs. weak study design								
Good design	556.387	38	0.000	-24	(-29; -18)	-13	(-20; -5)	13
Weak design	297.783	58	0.000	-13	(-15; -10)	-9	(-12; -7)	12
All accidents vs. accidents with alcohol involved								
Accidents with alcohol involved	495.591	65	0.000	-19	(-23; -15)	-17	(-21; -12)	4
All accidents	368.870	27	0.000	-15	(-20; -11)	-10	(-16; -4)	4
Testing of drivers stopped at the checkpoint								
Not all tested	849.275	72	0.000	-18	(-21; -14)	-15	(-19; -10)	6
All tested	140.017	24	0.000	-15	(-19; -10)	-13	(-17; -8)	2
Accompanying publicity with vs. without paid media								
With paid media	727.705	63	0.000	-20	(-24; -16)	-14	(-19; -10)	11
Without paid media	179.610	22	0.000	-15	(-20; -11)	-11	(-16; -6)	4

Table 16b: Test of heterogeneity and summary effects (RE models) of DUI checkpoints – moderator analyses: Results for injury accidents.

	Test of heterogeneity			Change of number of accidents (%)		Trim and fill analysis: Change of number of accidents (%)		
	Cochran's Q	df	p	Summary effect	95% confidence interval	Summary effect	95% confidence interval	Number of new effect estimates
Time period studied								
< 3 months	198.056	14	0.000	-29	(-41; -15)	-23	(-38; -4)	2
3 – 6 months	114.64	4	0.000	-18	(-34; +2)			
6 – 12 months	123.263	23	0.000	-14	(-17; -10)	-13	(-16; -10)	2
12 – 24 months	45.7203	8	0.000	-10	(-17; -1)			
Country								
Australia	549.785	29	0.000	-23	(-28; -18)	-20	(-26; -14)	2
New Zealand	112.48	14	0.000	-15	(-21; -10)	-8	(-14; -2)	6
USA	6.22813	6	0.398	-13	(-19; -8)			
Other countries	30.8301	9	0.000	-4	(-13; +5)			
Good vs. weak study design								
Good design	187.436	33	0.000	-12	(-15; -9)	-10	(-13; -7)	7
Weak design	492.795	27	0.000	-25	(-32; -18)	-16	(-24; -7)	7
All accidents vs. accidents with alcohol involved								
Accidents with alcohol involved	452.721	42	0.000	-20	(-25; -15)	-19	(-24; -14)	1
All accidents	261.348	15	0.000	-16	(-23; -9)	-13	(-21; -4)	1
Testing of drivers stopped at the checkpoint								
Not all tested	719.679	43	0.000	-19	(-24; -15)	-17	(-22; -12)	2
All tested	107.830	17	0.000	-16	(-22; -10)	-13	(-19; -6)	2
Accompanying publicity with vs. without paid media								
With paid media	562.718	38	0.000	-22	(-27; -16)	-19	(-25; -13)	2
Without paid media	150.978	15	0.000	-15	(-20; -10)	-14	(-19; -8)	1

Table 16c: Test of heterogeneity and summary effects (RE models) of DUI checkpoints – moderator analyses: Results for fatal accidents.

	Test of heterogeneity			Change of number of accidents (%)		Trim and fill analysis: Change of number of accidents (%)		
	Cochran's Q	df	p	Summary effect	95% confidence interval	Summary effect	95% confidence interval	Number of new effect estimates
Time period studied								
< 3 months	7.50882	2	0.023	-50	(-75; -1)			
3 – 6 months (fixed effects)	0	0		-42	(-60; -16)			
6 – 12 months	21.0118	6	0.002	-18	(-29; -4)			
12 – 24 months	42.0263	8	0.000	-16	(-27; -5)			
Country								
Australia	106.146	18	0.000	-19	(-26; -12)	-7	(-16; 2)	8
New Zealand	0	0		-4	(-8; 0)			
USA	19.9446	7	0.006	-11	(-17; -5)			
Good vs. weak study design								
Good design	108.484	20	0.000	-14	(-19; -9)	-17	(-30; -2)	2
Weak design	19.251	6	0.004	-21	(-32; -8)			
All accidents vs. accidents with alcohol involved								
Accidents with alcohol involved	37.559	16	0.002	-15	(-21; -9)	-13	(-20; -5)	4
All accidents	87.796	9	0.000	-16	(-23; -9)			
Testing of drivers stopped at the checkpoint								
Not all tested	117.130	25	0.000	-15	(-20; -10)	-7	(-14; -1)	9
All tested	20.550	1	0.000	-13	(-29; +6)			
Accompanying publicity with vs. without paid media								
With paid media	103.525	18	0.000	-16	(-22; -10)	-6	(-13; 2)	9
Without paid media	25.596	6	0.000	-15	(-25; -4)			

Comparing the summary effects between the subgroups of each of the potential moderator variables shows that larger accident reductions have been found

- when shorter time periods are studied, i.e. the largest accident reductions are found during the first half year;
- in Australia compared to New Zealand and USA, and New Zealand and USA compared to other countries;
- in studies with a weak study design compared to studies with a good study design;
- accidents involving alcohol compared to all accidents; this effect becomes even stronger when the results are controlled for publication bias;

- when not all drivers are tested at the checkpoints, compared to checkpoints where all drivers are tested, the difference is however not large;
- for injury accidents compared to fatal accidents, the difference becomes larger when the results are controlled for publication bias;
- when publicity involves paid media, compared to publicity with unpaid media only.

All results are significant, with the exception of the results for checkpoints in countries other than Australia, New Zealand and USA. Most results remain significant even when publication bias is controlled for.

The amount of heterogeneity remains significant in all subgroups. This indicates that the results within each of the subgroups are affected by further moderator variables.

4.5 Drink driving enforcement measures: Results from meta-regression

In the subgroup analyses in the preceding section, the effect of only one variable at a time is investigated. None of the investigated moderator variables has been found to be sufficient to reduce the heterogeneity in the results. In meta-regression analysis it is investigated how the potential moderator variables affect the effectiveness of DUI checkpoints when controlling for all other potential moderator variables at the same time. The same potential moderator variables are included as predictors that have been used as grouping variables in the subgroup analyses in the preceding section. In the meta-regression analysis all predictor variables are coded as dummy variables as shown in Table 17.

The meta-regression analysis is conducted in four steps:

Step (1) Stepwise forward analysis: The predictor variables are included in the regression model successively, one at a time. The first predictor is the one which produces the largest value of R^2 (variance explained by the model). The second predictor is the one which produces the largest value of R^2 when included in the model in addition to the first predictor. The third through seventh predictors are chosen accordingly. Table 17 shows the predictors in the order in which they were included in the model, and the values of R^2 that are obtained when the respective predictor and the preceeding predictors are included in the model.

Step (2) Complete regression model with all predictor variables, based on all studies: The regression coefficients are estimated for the complete model which contains all predictor variables.

Step (3) Complete regression model with all predictor variables, based on the studies with a good design only.

Step (4) In a fourth step, coefficients are estimated for a partial model which is based on all results. In the partial model only the predictor variables are included that have been identified as relevant in the preceding steps. The regression coefficients and p-values of the coefficients are then compared to the coefficients in the complete model with all predictor variables.

Step (3) and (4) are conducted in order to provide a test of the consistency ans robustness of the results. The results for step (1), (2) and (3) are shown in Table 17 (the results from step (4) are not shown in table). Table 17 shows how each of the predictor variables is dichotomized as a dummy variable with the values 0 and 1. The categories for which the larger accident reductions have been found are underlined. Table 18 shows the model summaries of the regression models that have been estimated in step (2) and (3).

Table 17: Effects of DUI checkpoints on accidents, results from meta regression.

		Step (1)	Step (2)		Step (3)	
			All results		Results from good study designs	
Predictor	(1; 0)	R^2 when predictor included into model	Coeff.	p(Coeff.)	Coeff.	p(Coeff.)
Time period	> 6 months (1) vs. < 6 months (0)	0.136	0.176	0.001	0.102	0.104
Country	Australia (1) vs. other countries (0)	0.225	-0.224	0.000	-0.180	0.003
Study design	Good design (1) vs. weak design (0)	0.277	0.120	0.009	-	-
Accident types	Involving alc. (1) vs. all accidents (0)	0.293	-0.080	0.052	-0.058	0.136
Testing of drivers	All tested (1) vs. not all tested (0)	0.311	-0.111	0.053	-0.085	0.171
Accident severity	Fatal accidents (1) vs. injury accidents (0)	0.314	-0.053	0.233	-0.061	0.159
Publicity	Paid media (1) vs. no paid media (0)	0.311	0.037	0.417	0.051	0.280

Table 18: Model summaries of the results from meta-regression.

Model	df	F	p(F)	R^2	Tau^2
Step (2) complete model based on all results	88	7.13	0.000	0.311	0.022
Step (3) complete model based on good study designs only	50	2.26	0.052	0.119	0.009

The first three variables that were included in the regression model are, in this order, time period studied, country and study design. All three variables have highly significant regression coefficients in the model that has been developed in step (2). The results are consistent with the findings from the subgroup analyses.

Time period: The results for the time period that has been studied show that larger accident reductions are found during the first half year, compared to longer time periods. The results indicate that the effectiveness decreases over time after introducing or changing a DUI programme. This does not mean that longer time periods reduce the effectiveness of checkpoints during the first half year after introduction, only that accident reductions are becoming smaller over time. One study has been found that has directly investigated the change of the effectiveness of DUI checkpoints over time (Newstead et al., 1998). In this study the actual numbers of accidents during each of eight consecutive years has been compared to the expected numbers of accidents, based on previous trends and when controlling for a number of other factors that affect accidents. The estimated accident reduction due to the DUI-checkpoint programme (Australian Random Breath Testing) has been found to be almost constant over time (ca. 9% in the first year and ca. 10% in the last year). A possible explanation for this finding is that checkpoints become more widely known and more drivers experience that getting away with drink-driving is unlikely once they pass a checkpoint. This may offset the effect of habituation and drivers getting acquainted to checkpoint locations, which makes it possible for experienced drink-drivers to avoid being caught at checkpoints.

Country: The countries in which DUI checkpoints have been studied also has a highly significant effect on the results. Country is included in the model earlier than any of the characteristics of the checkpoints. This suggests that there are properties of the checkpoint programmes which affect their effectiveness, but that could not be investigated (and controlled for) in the present meta-analysis. The largest accident reductions have been found in Australia in the subgroup analysis. Country was therefore coded as a dummy variable which compares Australia with all other countries in the meta-regression analysis. Among the Australian studies there are over 3 times more studies with a good quality, which tend to yield smaller effect estimates than weak study designs. This strengthens the conclusion that Australian DUI checkpoints are more effective than checkpoints in other countries. In the results from meta-regression the effect of study design is controlled for, i.e. the estimated coefficient for country is not affected by differences in study quality between countries.

Two obvious differences between Australian and other DUI checkpoint programmes are the use of highly visible “booze buses”, which were introduced in Australia in 1989 in the state of Victoria (Newstead et al., 1995), and the large amount of publicity accompanying the DUI checkpoint programmes in Australia. New Zealand is the only other country, where booze buses are being used. They were however introduced later (in 1996 on the north-island; Miller et al., 1996) and are to a lesser degree represented among the study results included in the meta-analysis. Both booze buses and a large amount of publicity have been found to be highly effective in increasing the effectiveness of DUI checkpoints in the study by Miller et al. (2004).

Study design: Study design is also highly significant predictor for the effects of DUI checkpoints. The largest accident reductions are found in studies with a weak study design, as has also been found in the subgroup analysis. The regression coefficients of the other predictor variables show the effects of these variables on the effectiveness of DUI checkpoints when controlling for study quality. These results are in other words not likely to be affected by study quality to a large degree. The classification of studies into “good” and “weak” studies is however only rough and study quality varies also within both categories.

The next two variables that were included in the regression model are accident type (alcohol vs. all accidents) and testing of drivers. Both variables have coefficients that are significant at the 95% level.

Accidents involving alcohol are according to the results in Table 16 more strongly reduced than other accidents. This result is consistent with the results from the subgroup analyses in the preceeding section. This is despite the fact that only few of the results that refer to accidents involving alcohol are actually based on accidents where the presence of illegal BAC limits is proven. Mostly, some substitute measure of alcohol accidents is used, such as weekend night accidents. Among such substitute types of accidents, alcohol is overrepresented, but not all of these accidents actually involve alcohol. The results are therefore likely to be closer to the results for all accidents than they would be if they were based only on accidents which are known to involve (or to be caused by) illegal BAC levels.

Testing of drivers seems to affect the effectiveness of DUI checkpoints. The negative regression coefficient indicates that checkpoints where all drivers are tested are more effective in reducing accidents. It should be taken into account that the grouping is mostly based on information about how checkpoints are supposed to work. In practice there may be differences between checkpoints within one programme. The effect of testing all drivers may therefore in practice be larger than the present results indicate. This result is contrary to the result of the subgroup analysis, where larger accident reductions were found of checkpoints where not all drivers are tested. This may be due to confounding factors, some of which are controlled for in the meta-regression analysis, but not in the subgroup analysis. One such confounding factor is country. The Australian checkpoints are classified as “not all drivers tested” and the Australian checkpoints have been found to be associated with larger accident reductions than other countries. This may explain the finding that more favourable effects are found when not all drives are tested and when country is not controlled for.

The last two variables that were included in the regression model are accident severity and accompanying publicity. None of these variables has a significant regression coefficient, and the amount of variance explained by the regression model does not increase noticeably when these variables are included in addition to the other predictor variables.

In the subgroup analysis larger accident reductions were found for injury accidents, compared to fatal accidents. The negative regression coefficient suggests larger effects on fatal accidents. However, the results from the regression analysis do not indicate that effects are significantly different between injury and fatal accidents. The results for accompanying publicity do not say anything about the effect of the amount or intensity of publicity, which could not be classified in a consistent manner.

4.5.1 Robustness of the results

A comparison between the complete model that is based on all studies and the complete model that is based on the good studies only does not show any large differences between the regression coefficients or the levels of significance. The direction of the effects and the relative importance of the predictors are comparable. A comparison between the complete model, which is based on all studies, and the partial model in step (4), which only includes the first five predictor variables, does not reveal any differences either. Both the regression coefficients and the significance levels are practically identical.

These findings indicate that the results from the meta-regression analysis are robust both with respect to the studies included in the analysis and the chosen regression model. This makes it less likely that the results for the moderator variables are arbitrary and due to randomness rather than “real” moderator effects.

Although 5 out of the 7 investigated moderator variables have been found to be significant predictors in the regression models, it is not unlikely that more moderator variables exist.

A limitation to the results from meta-regression is that publication bias is not controlled for. Under the assumption that all results are equally affected by publication bias, this would not limit the interpretation of the results in terms of moderator effects. If publication bias is not identical within all results however, the results may be biased.

In summary, the results from subgroup analyses and from meta-regression analysis have shown that DUI checkpoints are most effective in the first period after introduction and when all drivers who are stopped at a checkpoint are tested. The effects are larger on accidents involving alcohol than on other accidents. The effects are not different between accidents of different severity. Whether or not publicity is based on unpaid media only or includes paid media does not affect the effectiveness of DUI checkpoints either. This result does not say anything about the effect of the amount or intensity of publicity.

Larger effects have been found in Australia than in any other countries. In New Zealand and in USA the effects are larger than in other countries (except Australia). A comparison of characteristics of DUI checkpoint programmes between countries may therefore reveal further factors that improve the effectiveness of DUI checkpoints. Two such factors are, based on other studies, the amount and intensity of publicity and the used of highly visible “booze buses”.

Larger effects have been found in weak studies than in studies with a good study design. Studies that have not controlled for more than time trends at best have been classified as weak studies. The results are however robust against effects of study quality. In meta-regression it is controlled for study quality when al studies are included in the regression mode. Additionally, the results that are based on all studies could be replicated in a meta-regression model which is based on studies with a good study design only.

Almost all results are affected by publication bias, as indicated by the trim and fill analyses. Controlling for publication bias does not fundamentally change the results in the subgroup analyses. The effects of the moderator variables investigated can therefore be regarded as robust to publication bias as well. The size of the effects however is probably smaller than indicated by the results that are not corrected for publication bias.

4.5.2 Limitations of drink driving enforcement studies

The results from meta-analysis have shown that the results for both patrolling and DUI checkpoints are likely to be affected by study quality. Studies with a design that has been classified as “weak” in this analyses have consistently found larger accident reductions than studies with a “good” design. Studies without a comparison group and studies that have controlled for time trends only have been classified as “weak”. In these studies numerous factors that may affect accidents involving alcohol have not been controlled for. Studies with a “good” design are studies that have some kind of control group, either a type of accidents that is not assumed to be affected by DUI enforcement or accidents in a control area where no change is assumed in the type or intensity of DUI enforcement. Multivariate studies in which a number of factors are controlled for have also been classified as “good”. The quality of these studies varies, due to the limited number of results it is however not possible to make further tests of the effects of study quality on the estimated effects on accidents.

Trim-and fill analysis do not indicate that the results from studies of patrolling are affected by publication bias. The results are however so heterogeneous that publication bias would most likely not be detected. Results from studies of DUI checkpoints are likely to be affected by publication bias. Both the overall summary effects and the summary effects from subgroups of results (including results based on good study designs) change when publication bias is controlled for. When publication bias is controlled for, the estimated accident reductions diminish to between somewhat smaller and half as large as when publication bias is not controlled for.

Almost all results from meta-analysis show significant amounts of heterogeneity. This is true also for the results from the subgroup analyses, with the exception of the summary effect of patrolling on fatal accidents and the summary effect of patrolling on all types of accidents that is based on weak study designs. This indicates that the summary effects can not be regarded as representative of one “true” effect of one well-defined measure. The effect estimates from different studies rather seem to represent different “true” effects. All the same, the effects of further differences between studies have not been investigated. Sooner or later one would find some subgroups of results without significant amounts of heterogeneity. It would however be questionable if this would lead to the detection of relevant moderator variables. Most likely, the results would be mere products of chance.

5. EFFECTS OF SEAT BELT ENFORCEMENT MEASURES

5.1 Factors affecting the effectiveness of seat belt enforcement

In subgroup analyses studies are divided into subgroups according to a potential moderator variable. The summary effects of the subgroups of studies are then compared. In most cases there are only two subgroups. An important point in subgroup analysis is to keep the subgroups comparable with respect to other comparable moderator variables. Subgroups analyses are mostly only made when there is significant heterogeneity in the results for a group of effect estimates. Heterogeneity indicates that the results are affected by some factor or factors that are different between the studies. If there is no heterogeneity, all effect estimates can be regarded as representatives of the same outcome and no systematic differences between studies can be expected.

Type of enforcement measure: The following types of seat belt enforcement measures have been investigated:

- Stationary control at the roadside, checkpoints, mostly combined with speed or DUI enforcement
- Canadian and USA STEP program
- Combinations of checkpoints and mobile controls
- Educational enforcement of use of child restrains with leaflets and warnings instead of fines.

Most studies have investigated the effects of primary seat belt law enforcement on seat belt wearing. One study has investigated the effect of educational enforcement of the use of child restrains, where only warnings were given, but where no tickets were issued (Cox, 1981). These two studies are not included in the meta-analysis of the effects of seat belt enforcement.

Although all seat belt enforcement measures differ in several ways, there are no clearly distinguishable groups of different types of enforcement measures. For example, seat belt control at checkpoints may or may not be primary seat belt enforcement. The Canadian and USA STEP programs have been investigated in only one study each. Meta-analyses are conducted based on the studies of the effects of seat belt enforcement on seat belt wearing.

Drivers, front and back seat passengers: Effects of seat belt enforcement have been investigated on car occupants in different seating positions: drivers, front seat occupants, front seat passengers, and back seat passengers.

Daytime vs. night time: Most studies have not specified whether effects have been investigated at day or at night. Three studies have investigated effects on seat belt wearing at day and at night separately.

Visibility of enforcement: Most studies have investigated seat belt enforcement that was not signposted or have not specified whether or not signposting was used. Only two studies have specified that seat belt enforcement was signposted.

Randomization: In most of the studies randomization procedures are not clearly specified. Subgroup analyses have therefore not been conducted.

Change of type or amount of enforcement: All studies refer to either increased seat belt enforcement or to a combination of increased and changed seat belt enforcement.

Country: Most studies have been conducted in the Netherlands or USA. Some studies have also been conducted in Australia, Belgium and Canada. Subgroup-analyses are conducted for each of these countries.

Publicity: The studies have been classified according to the amount of accompanying publicity:

- No publicity
- Local publicity
- Publicity campaign, i.e. enforcement is accompanied by a campaign incl. a mass media component
- Seat belt enforcement is part of a wider enforcement programme, most likely including publicity

Study methodology: All studies have compared seat belt wearing rates before and after the amount of seat belt enforcement has been increased / changed. Six studies have used a control group and in the remaining studies no control group has been used.

5.2 Meta analysis of seat-belt enforcement studies

An overview of the available studies that are included in the meta-analysis is shown in Table 19.

Table 19: Studies of the effects of seat belt enforcement on seat belt wearing.

Authors	Year	Country	Type of measure	Study design	Sum of statistical weights[1]
Gundy	1986	Netherlands	Seat belt enforcement	With comparison	1,042
Beke	1990	Netherlands	Seat belt enforcement	Without comparison	11,616
Vissers	1989	Netherlands	Seat belt enforcement	Without comparison	4,809
Grant	1991	Canada	Seat belt enforcement	Without comparison	999
Geary	2005	USA	Seat belt enforcement	Without comparison	16,100
Nuyts	2006	Belgium	Seat belt enforcement	Without comparison	5,410
Gras	1987	Netherlands	Seat belt enforcement	Without comparison	1,043
Werkgroep actie autogordel Utrecht	1988	Netherlands	Seat belt enforcement	Without comparison	11,754
Kaye	1995	USA	Seat belt enforcement	With comparison	1,929
Mathijssen	1992	Netherlands	Seat belt enforcement	Without comparison	4,149
Reinfurt	1990	USA	Seat belt enforcement	Without comparison	25,000
Salzberg	2004	USA	Seat belt enforcement	Without comparison	5,750
Wells	1992	USA	Seat belt enforcement	With comparison	2,922
Streff	1992	USA	Seat belt enforcement	Without comparison	3,076
Cox	1981	USA	Seat belt enforcement	With comparison	1,549
Hagenzieker	1991	Australia	Enforcement of use of child restraints	With comparison	93
Lund	1989	Netherlands	Secondary seat belt law enforcement	Without comparison	6,147

[1] Statistical weights in fixed effects model

Data inspection, publication bias: A scatterplot of all effect estimates for seat belt enforcement is shown in Figure 9 (studies of use of child restraints are not included). It is indicated in the figure for all effect estimates whether they refer to a before-during comparison or a before-after comparison. The natural logarithm of the overall summary effect that is based on all before-during effect estimates is 0.244, which corresponds to an odds ratio of 1.276 (Fixed Effects), i.e. an increase of the use of seat belts by ca. 28%. The natural logarithm of the overall summary effect that is based on all before-after effect estimates is 0.129, which corresponds to an odds ratio of 1.138 (Fixed Effects), i.e. an increase of the use of seat belts by ca. 14%.

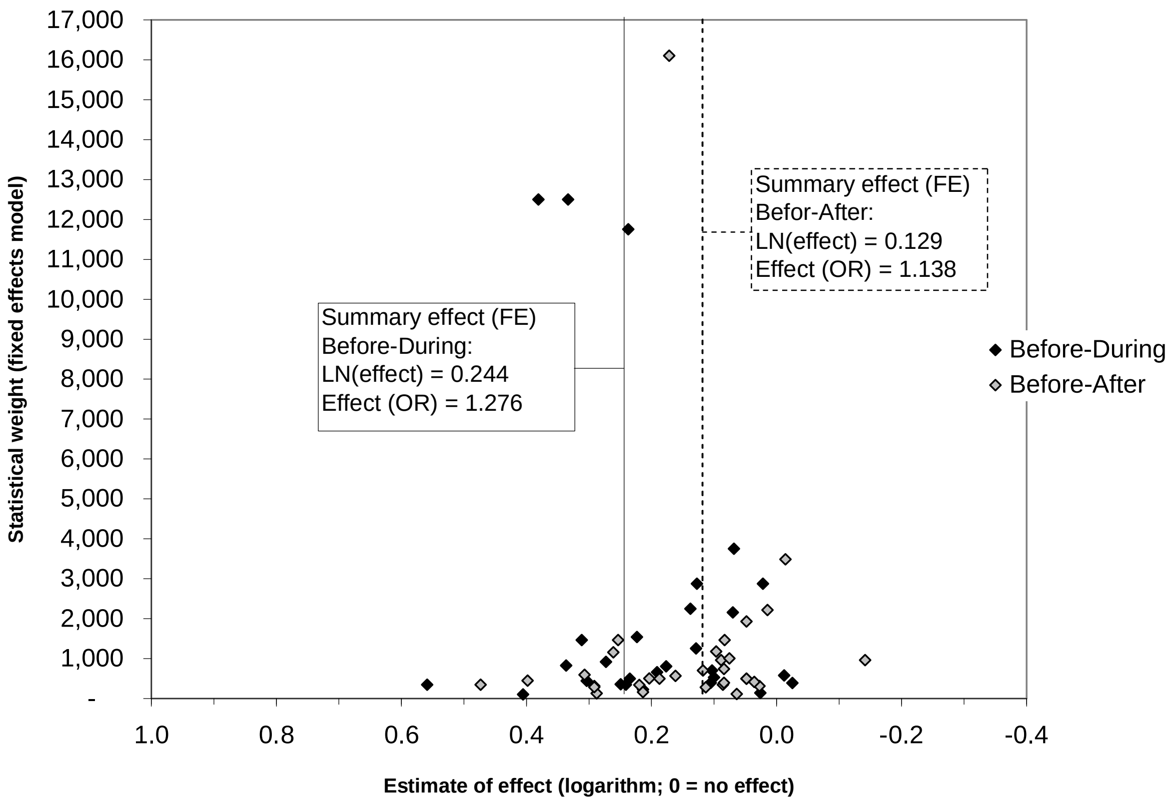

Figure 9: Scatterplot of effect estimates for seat belt enforcement.

The effect estimates in the lower part of Figure 9 are distributed in a funnel-like pattern, which is however quite asymmetrical. Larger increases of the use of seat belts (higher positive values of the logarithm of the effect estimates) have been found in smaller studies. This pattern might indicate the presence of publication bias. Four effect estimates are associated with far larger statistical weights (above 11,000) than all other effect estimates. These effect estimates are based on three different studies, which do appear to be systematically different from the other studies. Trim and fill analysis is conducted with all effect estimates, with the effect estimates that refer to Before-During comparisons, and with effect estimates that refer to Before-After comparisons. No new effect estimates have been generated in any of these analyses, which indicates that the results are not affected by publication bias.

5.2.1 Seat belt enforcement: Summary effects, tests for heterogeneity, and subgroup analyses

Results from tests of heterogeneity and summary effects of the estimated effects of seat belt enforcement on seat belt wearing rates are shown in Table 20. The overall result is a significant increase in the seat belt usage rates. Larger increases of the use of seat belts have consistently been found in comparisons of seat belt use before and during the implementation of enforcement measures

Table 20: Test of heterogeneity and summary effects (Random Effects models) of seat belt enforcement.

	Before – During					Before - After				
	Test of heterogeneity			Change of seat belt usage(%)		Test of heterogeneity			Change of seat belt usage(%)	
	Cochran's Q	df	p	Summary effect	95% confidence interval	Cochran's Q	df	p	Summary effect	95% confidence interval
All	924.28	29	0.000	+21	(+16; +27)	391.37	29	0.000	+15	(+10; +20)
Belgium	1.46	1	0.227	+17	(+11; +22)	5.67	1	0.017	+10	(-1; +23)
Canada	0.00	0		+26	(+16; +38)	0.00	0		+21	(+10; +32)
The Netherlands	271.25	18	0.000	+20	(+14; +27)	267.48	19	0.000	+19	(+11; +26)
USA	452.64	7	0.000	+24	(+13; +36)	104.89	6	0.000	+7	(-1; +16)
With comparison	36.19	2	0.000	+38	(+10; +73)	101.68	7	0.000	+11	(+1; +23)
Without comparison	886.07	26	0.000	+20	(+14; +26)	267.95	21	0.000	+17	(+11; +22)
Drivers	693.83	20	0.000	+23	(+16; +30)	299.29	17	0.000	+18	(+11; +24)
Front seat passengers	13.60	4	0.009	+20	(+10; +32)	6.65	4	0.155	+17	(+10; +26)
All front seat occupants	23.38	2	0.000	+11	(+2; +21)	43.00	5	0.000	+5	(-3; +14)
All seat belts wearing rates	903.20	27	0.000	+21	(+15; +27)	390.01	27	0.000	+15	(+10; +21)
Daytime wearing rates	0.00	0		+11	(+3; +19)	0.00	0		+10	(+4; +17)
Nighttime wearing rates	0.00	0		+40	(+31; +50)	0.00	0		+12	(0; +26)
Increased enforcement	161.28	8	0.000	+30	(+18; +44)	326.36	13	0.000	+19	(+11; +28)
Increased and changed enforcement	686.63	20	0.000	+18	(+12; +25)	56.20	15	0.000	+12	(+7; +16)
Signposted	23.38	2	0.000	+11	(+2; +21)	0.00	0		+9	(+1; +17)
Not signposted	272.38	19	0.000	+21	(+14; +27)	289.17	21	0.000	+19	(+13; +25)
Unspecified signposting	181.40	6	0.000	+28	(+19; +37)	49.09	6	0.000	+6	(-1; +14)
No publicity						0.00	0		-13	(-19; -8)
Local publicity	195.22	7	0.000	+21	(+12; +31)	30.89	5	0.000	+17	(+7; +28)
Publicity campaign	285.84	13	0.000	+24	(+17; +31)	266.40	14	0.000	+20	(+13; +28)
Enforcement program	39.52	7	0.000	+17	(+10; +25)	1.55	7	0.980	+9	(+6; +12)

There are significant amounts of heterogeneity in most results. Moderator variables are therefore likely to be present. Those potential moderator variables for which subgroup analyses could be conducted are country, use of comparison group, seating position, time of seat belt usage, increase vs. change of enforcement, signposting of enforcement, and publicity. The subgroup analyses show that

- seat belt enforcement is more effective in the USA than in other countries when regarding before-during comparisons, and least effective in Belgium;
- in before-during comparisons larger increase of seat belt use have been found in studies which have not applied a control group; however, in before-after comparisons the difference between studies with and without control group is only small,
- there are no systematic differences in the effects on drivers and front seat passengers,
- seat belt enforcement is more effective in increasing seat belt use at night; this result refers only to before-during comparisons and is based on only one study,
- an increase of seat belt enforcement is more effective than a simultaneous increase and change of the type of enforcement
- seat belt enforcement that is conducted without signposting is more effective that signposted enforcement,
- local publicity and a publicity campaign increase the effectiveness of seat belt enforcement compared to no publicity or enforcement programmes (it is likely that enforcement programmes include at least some components of a publicity campaign).

The amount of heterogeneity remains significant in almost all subgroups. This indicates that the results within each of the subgroups are affected by further moderator variables.

5.2.2 Enforcement of the use of child restraints

Only one study has been found that has investigated the effect of enforcement on the use of child restraints. In this study (Cox, 1981) a non-significant increase of the use of child restraints by 15% has been found (95% confidence interval [-13%; +25%]). This study refers not strictly speaking to enforcement, since no tickets have been issued but only information has been provided and drivers not properly using child restraints have got warnings.

6. DISCUSSION

Speed enforcement: In previous reviews, as with Elvik and Vaa's (2004), the literature on speed enforcement have justified a separation of different methods of speed enforcement as methods differ in their effectiveness of creating halo effects in time and distance. Specifically, a distinction should be made between stationary methods and "mobile" methods, as with police patrols (Elvik and Vaa, 2004). This distinction is relevant because halo effects have been found in time and space for stationary enforcement, but not for mobile patrols (Shinar and McKnight 1985; Vaa 1993). A further distinction is the selection of radar/laser "American type" and "Composite other". The former is the situation where the same unit unobtrusively or hidden observes the driver and also initiate a pursuit in case of violations. This speed enforcement category is used in Australia as well and possibly also in other countries. "Composite other" indicates that speed enforcement is done with a range of different measures. In addition to marked cars, motorcycles, helicopter, aircraft or other may be used either as an observation unit and/or in apprehension or pursuit. The essence of this category is that there is some element of visible and stationary element, thus resembling the category "Stationary manual", but differs in terms of the number of measures that were available. The speed studies also enable a distinction between "Mobile cameras" and "Fixed speed cameras", a distinction that was not applicable in the previous meta-analysis of speed enforcement measures.

Summing up the effects, the meta-analysis states that the overall effect of speed enforcement measures is a reduction in the number of accidents (all severity levels included) of 18% (-23;-13). The effects of injury accidents and fatal accidents after correction for publication bias are reductions of 14% (-20;-7) and 29% (-34; -23), respectively. Serious injury accidents is also significantly reduced by 27% (-36;-16), while slight injuries are not. This makes sense as speed enforcement act on speeding behaviour and it is a well-known fact that the level of injuries increase when the level of speed increases (Elvik, 2004). Considering the overall effects on all accident severities of different speed enforcement measures, speed cameras are the only measure that reduces accidents significantly by 30% (-38;-23) and fixed cameras are significantly better than mobile cameras. There is a tendency that stationary, manual speed enforcement, and "Composite other" also reduce accident, but this tendency is insignificant. Patrolling and radar/laser "American type" have no effect on the number of accidents. Considering injury accidents ("unspecified" and property-damage-only accidents excluded), the "Composite other" speed enforcement here reduces the number of accidents by 18% (-30:-4). Regarding fatal accidents, the number of results is too limited to perform meta-analyses, but some single-standing studies show that stationary, manual, mobile speed cameras and fixed speed cameras reduce fatal accidents significantly.

Considering potential predictors that may explain why speed enforcement reduces accidents, these enforcement attributes seem important: Visibility (of enforcement), signposting, local publicity, and increasing the amount of enforcement by more than 200%.

Drink driving enforcement: The present study allows for a distinction between observing and apprehending drink drivers by patrolling and by using DUI checkpoints where one or more police cars are standing beside the road (not driving) and where police officers pull out drivers in order to check whether or not a driver is has an illegal BAC level. This distinction between the two methods is a step forward compared to a previous meta-analysis of drink driving enforcement (Elvik and Vaa, 2004). In the present meta-analysis the overall effect of patrolling is estimated to a reduction in the number of accidents (all severities) of 8% (-12;-3). However, when excluding studies with a weak study design the effect is down to a marginal 4% (-8;+0), while an estimate based on weak designs only is as high as -24% (-28;-20). However, this result is found when there is no control for confounding variables. The overall impression from patrolling is that there is some reduction in the number of accidents, but the effect is rather small. Excluding studies with “Unspecified accident types”, i.e. studies that may comprise property-damage-only accidents, the cultivated effect on injury accidents only is a reduction of 11% (-16;-4) while there is no effect on fatal accidents.

Considering DUI checkpoints, however, there are significant effects on injury accidents by -16 (-20;-11) as well as fatal accidents by -6% (-7;-5). The effect on injury accidents is significantly higher than estimated by Elvik and Vaa (2004) when it was estimated to be -7% (-8;-6). One explanation for this change could be that the present estimate is stripped of less effective patrolling enforcement. However, the estimate on fatal accident is lower in the present analysis (-6% (-8;-5)) than found by Elvik and Vaa. (-9% (-11; -8)). These estimates are, however, not significantly different.

Seat belt enforcement: The enforcement of seat belt use is very consistent. The overall estimate as well as estimates in all sub-groups tested, are significant, ranging from a low + 11% to a maximum of +40% increase in seat belt use in the during-period of enforcement. For the after-period the effects are somewhat lower, than in the during-period, but also most of the sub-group estimates are statistically significant. The finding that seat belt enforcement is more effective in the USA than in other countries when regarding before-during comparisons, might be rooted in a higher potential of improvement in the usage rates in the USA than in other countries.

7. CONCLUSIONS

7.1 Effects of speed enforcement measures

A total of 45 evaluation studies on speed enforcement measures were identified and found to be of an acceptable quality to be included in the database. The 45 studies have been published in 14 countries and comprise a total of 129 results. USA, Australia, UK and Sweden are the ones which have had the largest numbers of results from speed enforcement evaluations.

The following methods of speed enforcement are distinguished:

- **Stationary speed enforcement using laser or radar** that measures speed from one, usually unobtrusive or hidden observation site, or instruments that measure mean speed between two fixed observation sites and clearly visible apprehension sites staffed by uniformed police officers and marked cars.
- **Stationary radar enforcement "American type".** The police observer (sometimes one officer alone in a car) measures speed by a radar mounted on the window and then pursues offending vehicles straightaway in order to apprehend and sanction the speeding driver. This technique is also used in Australia and the observation unit may be overt (marked car) as well as covert (unmarked car).
- **Patrolling**: Mobile police patrols with uniformed cars or motorcycles.
- **Composite police controls with stationary and visible elements**: This term is used to signify that speed enforcement may utilize a whole range of different techniques and methods, but also that it comprises some element that is stationary and that some of the activity is visible to the drivers passing by.
- **Speed cameras:** Several deployment patterns can be distinguished: 1) Fixed speed cameras, most often visible, on fixed locations/poles with a mobile camera moving around, or 2) Mobile cameras, less obtrusive or even hidden cameras used on different locations, and 3) section control of speed (where the average speed between two fixed sites is calculated and enforced if the speed limit is violated).

The overall result is a significant reduction of 18 % of the number of accidents. There are large differences between the estimated effects of the different types of speed enforcement. The trim-and-fill analyses indicate that the results for mobile speed cameras are affected by publication bias.

Table 21 . Summary effects of speed enforcement measures

	Test of heterogeneity			Change of number of accidents (%)	
	Cochran's Q	df	p	Summary effect	95% confidence interval
All measures	5307.82	128	0.000	**-18**	(-23; -13)
Stationary manual	1854.17	22	0.000	**-11**	(-22; +1)
Patrolling	62.7573	10	0.000	**-6**	(-16; +4)
Radar/laser US/AUS	22.3372	30	0.841	**-0**	(-3; 4)
Speed cameras (all types)	1693.9	42	0.000	**-30**	(-38; -23)
- Subgroup: Mobile speed cameras	168.476	12	0.000	**-17**	(-34; 4)
- Subgroup: Fixed speed cameras	1513.02	27	0.000	**-34**	(-42; -25)
Composite Other	454.306	20	0.000	**-18**	(-33; +1)

Table 22 shows the results from meta-regression:

Table 22: Effects of speed enforcement on accidents, results from meta regression.

Predictor	Dummy variables (1; 0)	Step (1) R^2 when predictor included into model	Step (2) All results Coeff.	Step (2) All results p(Coeff.)	Step (3) Results from good study designs Coeff.	Step (3) Results from good study designs p(Coeff.)
Type of enforcement	(Radar US/AUS; <u>other</u>)	0.096	0.224	0.050	0.142	0.026
	(<u>Speed cameras</u>; other)		-0.030	0.831	-0.036	0.674
	(<u>Composite</u>; other)		-0.090	0.442	-0.089	0.245
Signposted	(<u>Signposted</u>; other)	0.169	-0.231	0.130	-0.126	0.261
Accident severity.	(<u>Fatal/serious</u>; other)	0.207	-0.166	0.020	-0.166	0.001
Increase	(Incr.; <u>change of enf.</u>)	0.230	0.124	0.150	0.084	0.203
Method	(Good st.; <u>weak study</u>)	0.237	0.054	0.523		
Visibility	(<u>Visible</u>; other)	0.238	-0.133	0.117	-0.162	0.003
Country	(Australia; other)	0.240	-0.055	0.602	0.054	0.405
	(UK; other)		-0.051	0.679	0.162	0.050
	(US A; other)		-0.191	0.076	0.093	0.195
Publicity	(<u>Local publicity</u>; other)	0.230	-0.085	0.328	-0.004	0.949
	(<u>Pub. campaign</u>; other)		-0.028	0.807	-0.034	0.616

- The American / Australian type of radar / laser enforcement is less effective than the other methods. This result is consistent between the subgroup and the meta-regression analysis.
- **Signposting**: Speed enforcement has been found to be more effective when signposted than when not signposted. The regression coefficient is not significant in the complete model with all predictor variables. In the partial models (step 4) however, which are based only on

those predictor variables that have been identified in the stepwise procedure (step 1), the coefficient is somewhat larger (-0.337) and significant (p = 0.002).

- **Accident severity**: The effects of speed enforcement have been found to be greater for more serious accidents. The coefficient is significant both when all studies are included in the study and when the regression analysis is based on good studies only. In the partial models (step 4) the result is similar as in the complete model (coefficient = -0.184; p = 0.007).

- **Increase vs. change of enforcement**: The effectiveness of enforcement has been found to be greater when a changed form of enforcement is introduced, compared to an increase of the amount of enforcement. When the type of enforcement is changed the intensity (e.g. in terms of person hours) if often increased as well. This result is consistent between the subgroup analysis and the meta-regression analysis. The regression coefficients are however not significant.

- **Method**: Greater effects of speed enforcement have been found in studies with a weak study design, compared to studies with a good study design. The regression coefficient is however not significant and the increase in R^2 when including this predictor variable in the model in the stepwise procedure is only small.

- **Visibility**: The visibility of speed enforcement has not been found to be relevant for the effectiveness of speed enforcement in the subgroup analysis. In the meta-regression analysis no significant effect has been found and R^2 is not improved when this variable is introduced in the model in the stepwise procedure. All the same, the visibility of enforcement may not be irrelevant since signposting has been fount to be a significant predictor for the effectiveness. Additionally, in the subgroup analysis, the effect of non-visible enforcement is far smaller than the effect of visible enforcement when it is controlled for publication bias.

- **Country**: The results for the effects of speed enforcement in different countries are inconsistent. In the meta-regression analysis the coefficients have the opposite sign in the model based on all studies and in the model based on studies with a good study design. It is also curious that all three dummy variables have the same sign, although the subgroup analysis would indicate that the effectiveness is smaller in Australia than in other countries and greater in the UK than in other countries. When adding the country dummy variables in the stepwise regression analysis, the predictive power of the model does not improve markedly. It is therefore concluded that the effectiveness of speed enforcement is not systematically different between different countries.

- **Publicity**: Publicity is, according to the results from meta-regression not a relevant predictor variable. In the subgroup analysis local publicity has been found to increase the effectiveness of enforcement, while publicity campaigns have been found to decrease its effectiveness. All except one studies of enforcement that is accompanied by a publicity campaign have used a “good” study design, which has been found to reduce the effectiveness that is found in a study. This may explain the findings that publicity campaigns seem to reduce the effectiveness of enforcement, but that this effect disappears when study design is controlled for.

7.2 Effects of drink driving enforcement measures

A comparison between the results from good and weak study designs shows that a significant accident reduction only was found in studies with a weak design, not in studies with a good design. The rest of table 23 shows the results that are based on studies with a good design only. For both groups of effect estimates there are significant amounts of heterogeneity and the summary effects are almost identical. For fatal accidents, no significant effect on accidents has been found.

Table 23. Test of heterogeneity and summary effects (Random Effects models) of patrolling.

	Test of heterogeneity			Change of number of accidents (%)	
Type of accidents affected	**Cochran's Q**	**df**	**P**	**Summary effect**	**95% confidence interval**
	All results				
All accidents	156.41	26	0.000	-8	(-12; -3)
	Good vs. weak study design				
All accidents, good study design	54.38	20	0.000	-4	(-8; +0)
All accidents, weak study design	2.680	5	0.749	-24	(-28; -20)
	Injury, fatal and night time accidents (good study designs only)				
Injury accidents / unspecified severity; all types of accidents	39.62	6	.000	-6	(-11; 0)
Injury accidents / unspecified severity; all types of accidents – outlier omitted[1]	25.98	5	.000	-7	(-14; -1)
Injury accidents / unspecified severity; night time accidents	22.97	4	.000	-9	(-17; 0)
Fatal accidents; all types of accidents	14.71	13	.325	-1	(-7; +5)

Effects of DUI-checkpoints

The results from the trim-and-fill analysis indicate that the results for DUI checkpoints are affected by publication bias. The results presented in table 24 are therefore also adjusted for publication bias. These show to which groups of injury severity, type of checkpoint, time period, and accompanying publicity the effect estimates refer. The overall result is a significant reduction of 15% of the number of accidents when the results are adjusted for publication bias in trim-and-fill analyses.

Table 24. Test of heterogeneity and summary effects (Random Effects models) of DUI checkpoints.

	Test of heterogeneity			Change of number of accidents (%)		Trim-and-fill analysis: Change of number of accidents (%)		
	Cochran's Q	df	p	Summary effect	95% confidence interval	Summary effect	95% confidence interval	Number of new effect estimates
All results	1007.221	96	0.000	-17	(-20; -14)	-15	(-18; -11)	7

DUI: Results from meta-regression

Table 25 presents the results from meta-regression of the effects of DUI-checkpoints.

Table 25. Effects of DUI checkpoints on accidents, results from meta-regression.

Predictor	(1; 0)	Step (1) R^2 when predictor included into model	Step (2) All results		Step (3) Results from good study designs	
			Coeff.	p (Coeff)	Coeff.	p (Coeff.)
Time period	> 6 months (1) vs. < 6 months (0)	0.136	0.176	0.001	0.102	0.104
Country	Australia (1) vs. other countries (0)	0.225	-0.224	0.000	-0.180	0.003
Study design	Good design (1) vs. weak design (0)	0.277	0.120	0.009	-	-
Accident types	Involving alc. (1) vs. all accidents (0)	0.293	-0.080	0.052	-0.058	0.136
Testing of drivers	All tested (1) vs. not all tested (0)	0.311	-0.111	0.053	-0.085	0.171
Accident severity	Fatal accidents (1) vs. injury accidents (0)	0.314	-0.053	0.233	-0.061	0.159
Publicity	Paid media (1) vs. no paid media (0)	0.311	0.037	0.417	0.051	0.280

The results for the *time period* that has been studied show that the largest accident reductions are found during the first half year. When the time period after introducing or changing a DUI programme is longer, the effectiveness seems to decrease. This does not mean that longer time periods reduce the effectiveness of checkpoints during the first half year after introduction, only that accident reductions are becoming smaller over time

The *country* in which DUI checkpoints have been studied also has a highly significant effect on the results. The largest accident reductions have been found in Australia. Two obvious differences between Australian and other DUI checkpoint programmes are the use of highly visible “booze buses”, which were introduced in Australia in 1999, and the large amount of publicity accompanying the DUI checkpoint programmes in Australia.

Study design is also a highly significant predictor for the effects of DUI checkpoints, but study design is not related to any of the characteristics of DUI enforcement.

Accidents involving alcohol are according to the results more strongly reduced than other accidents. Mostly, some substitute measure of alcohol accidents is used, such as weekend night accidents.

Testing of drivers seems to affect the effectiveness of DUI checkpoints. The negative regression coefficient indicates that checkpoints where all drivers are tested are more effective in reducing accidents.

Accident severity and use of media: The results from the regression analysis do, however, not indicate that effects are significantly different between injury and fatal accidents. The results for accompanying publicity do not say anything about the effect of the amount or intensity of publicity, which it was not possible to classify in a consistent way for all studies.

7.3 Effects of seat belt enforcement measures

The following types of seat belt enforcement measures have been investigated:

- Stationary control at the roadside, checkpoints, mostly combined with speed or DUI enforcement
- Canadian and American STEP program
- Combinations of checkpoints and mobile controls
- Educational enforcement of use of child restrains with leaflets and warnings instead of tickets

Most studies have investigated the effects of primary seat belt law enforcement on seat belt wearing. Although all seat belt enforcement measures differ in several ways, there are no clearly distinguishable groups of different types of enforcement measures. The Canadian and USA STEP programs have been investigated in only one study each. Meta-analysis is conducted based on the studies of the effects of seat belt enforcement on seat belt wearing.

Publication bias: Trim and fill analysis is conducted with all effect estimates. No new effect estimates have been generated in any of these analyses, which indicate that the results are not affected by publication bias.

Results from tests of heterogeneity and summary effects of the estimated effects of seat belt enforcement on seat belt wearing rates are shown in Table 26. The overall result is a significant increase in the seat belt usage rates. Larger increases of the use of seat belts have consistently been found in comparisons of seat belt use before and during the implementation of enforcement measures.

Table 26: Overall summary effects (Random Effects models) of all seat belt enforcement studies

Before – During					Before - After				
Test of heterogeneity			Change of seat belt usage(%)		Test of heterogeneity			Change of seat belt usage(%)	
Cochran's Q	df	P	Summary effect	95% confid-ence interval	Cochran's Q	df	p	Summary effect	95% confide-ence interval
924.28	29	0.000	+21	(+16; +27)	391.37	29	0.000	+15	(+10; +20)

The subgroup analyses show that:

- seat belt enforcement is more effective in the USA than in other countries when regarding before-during comparisons,
- in before-during comparisons larger increase of seat belt use have been found in studies which have not applied a control group; however, in before-after comparisons the difference between studies with and without control group is only marginal,
- there are no systematic differences in the effects on drivers and front seat passengers,
- seat belt enforcement is more effective in increasing seat belt use at night; this result refers only to before-during comparisons and is based on only one study,
- an increase of seat belt enforcement is more effective than a simultaneous increase and change of the type of enforcement
- seat belt enforcement that is conducted without signposting is more effective than signposted enforcement,
- local publicity and a publicity campaign increase the effectiveness of seat belt enforcement compared to no publicity or enforcement programmes (it is likely that enforcement programmes include at least some components of a publicity campaign).

The amount of heterogeneity remains significant in almost all subgroups. This indicates that the results within each of the subgroups are affected by further moderator variables.

Enforcement of the use of child restraints

Only one study has been found that has investigated the effect of enforcement on the use of child restraints. In this study a non-significant increase of the use of child restraints by 15% has been found (-13; +25). This study refers not strictly speaking to enforcement, since no tickets have been issued. Only information was provided and drivers not properly using child restraints got warnings but no fines.

8. REFERENCES

Elvik, R; Vaa, T (2004): The Handbook of Road Safety Measures. Oxford, Elsevier.

Elvik, R (2004): Speed and road accidents: An evaluation of the Power model. Oslo, Institute of Transport Economics. TØI-report no 740/2004.

Fleiss, J.L. (1981). *Statistical Methods for Rates and Proportions.* Second Edition. New York, NY, John Wiley and Sons.

Rocakova, P.F. (2008): "Evaluation of the 0.0 BAC limit for car drivers in Czech Republic, Slovakia, Hungary and Croatia". Working paper 41 from the PEPPER project

van Schagen, Erke, A; Ewert, U; Kallberg, V-P. (2008). Good practice in data, data collection, and data use for monitoring and evaluating Traffic Law Enforcement. Deliverable 4a from the PEPPER project.

Shinar D, & McKnight AJ: The effects of enforcement and public information on compliance. I Schwing RC og EVANS L (eds): Human Behaviour and Traffic safety, New York, Plenum Press, 1985.

Kallberg, V-P.; Ewert, U (2008): Method for the prediction of the effects on safety of traffic enforcement measures. Deliverable 4b from the PEPPER project.

Vaa, T: Politiets trafikkontroller: Virkning på atferd og ulykker. En litteraturstudie. Oslo, Transportøkonomisk institutt, 1993 (TØI-rapport nr 204/1993).

9. STUDIES USED IN META-ANALYSIS OF SPEED ENFORCEMENT MEASURES

Amann, R. (1990). 20 Jahre Radarkontrollen in Nürnberg - Geschwindigkeitsüberwachung, Einflußfaktor auf das Unfallgeschehen Periodikum E35: BR 0722-5962. Vol. 1, pp. 1-4.

Chen, G., Wilson, J., Meckle, W. and Cooper, P. (2000). Evaluation of photo radar program in British Columbia. Accident Analysis and Prevention, 32, pp. 517-526.

Chen, G., Meckle, W., and Wilson J. (2002). Speed and safety effect of photo radar enforcement on a highway corridor in British Colombia. Accident Analysis and Prevention, 34(2), pp. 129-38.

Christie, S.M., Lyons, R.A., Dunstan, F.D., Jones, S.J. (2003). Are mobile speed cameras effective? A controlled before and after study. Injury Prevention, 9, pp. 302-306.

Dobrotka, D. (1998) Speed enforcement program. "Its our town, please slow down" Description and Analysis. Submitted to: The Police Executive Research Forum to be considered for The Herman Goldstein Award for Excellence in Problem-Oriented Policing. Submitted by: Chief David Dobrotka Glendale Police Department, Glendale, Arizona. Accessed 16 May 2008: http://www.popcenter.org/library/awards/goldstein/1998/98-23.pdf

Fitzharris, M., Gelb, K.R. Harrison, W.A. Newstead, S.V., Diamantopoulou, K and Cameron, M.H. (1999). Evaluation of the Effect of the Deployment of Hand-held Laser Speed-detection Devices in the Melbourne Metropolitan Area. Road Safety: Research, Policing & Education Conference: Handbook and proceedings, pp. 709-720.

Gains, A., Norstrom, M., Heydecker, B., and Shrewsbury, J. (2004). The national safety camera progamme. Four-year evaluation report. December 2005. PA Consulting Group, London.

Goldenbeld, C. and I.N.L.G van Schagen (2005). The effects of speed enforcement with mobile radar on speed and accidents. An evaluation study on rural roads in the Dutch province Friesland. Accident Analysis and prevention, 37, 1135-1144. **()**

Ha, T-J.; Kang, J-G.; Park, J-J. (2003). The effects of automated speed enforcement systems on traffic flow characteristics and accidents in Korea. ITE Journal, 73(2), pp. 28-31. **()**

Hakkert, A.S., Gitelman, V., Cohen, A., Doveh, E., Umansky, T. (2001). The evaluation of effects on driver behavior and accidents of concentrated general enforcement on interurban roads in Israel. Accident Analysis and Prevention, 33(1), pp. 43-63. (SWOV)

Hess, S. and J. Polak (2003). Effects of speed limit enforcement cameras on accident rates. Transportation Research Board, 1830, 25-33.

Hess, S. (2004). An analysis of the effects of speed limit enforcement cameras with differentiation by road type and catchment area, Transportation Research Record, 1865, 28-34.

Jorgensen, N.O. (1973). Experimental research concerning the effects of police enforcement on traffic accidents. Proceedings of the seminar on accidents analysing, University of Sussex.

Kang,J-G. (2002). Changes of speed and safety by automated speed enforcement. IATSS review. Vol. 26, No. 2, pp. 38-44.

Lamm R, and Kloeckner J H. (1984). Increase of Traffic Safety by Surveillance of Speed Limits with Automatic Radar Devices on a Dangerous Section of a German Autobahn: A Long-Term Investigation. Transportation Research Record, 974, pp. 8-16.

Leggett, L .M. W.(1988). The effect on accident occurrence of long-term, low-intensity police enforcement. Proceedings of the 14th ARRB Conference, Part 4, 92-104, 1988.

Machemer, E.; Runde, B.; Wolf, U.; Büttner, D.; Tücke, M.; Kley, H.; Pieper, M.; Schophuis, R.; Timmer, H.; Trecksel, H.; Wundram, A. (1996). Delegierte belohnung und intensivierte verkehrsüberwachung im vergeleich: eine empirische untersuchung zur beeinflussung des geschwindigkeitsverhaltens. Bericht zum forschungsprojekt 2.9129 'Reduzierung von geschwindigkeiten im ländlichen raum durch delegierte belohnung und/oder intensivierte verkehrsüberwachung' der bundesanstalt für strassenwesen BASt. Berichte der bundesanstalt für strassenwesen: Mensch und Sicherheit; Heft M 48. Bundesanstalt für Strassenwesen BASt, Bergisch Gladbach.

McCartt, A.T. and D.H. Rood (1989). Evaluation of the New York State Police 55 mph Speed Enforcement Project. Final Report. State University of New York, National Highway Traffic Safety Administration, August 1989, p. 137 .

Newstead S , and Cameron M. (2003). Evaluation of crash effects of the Queensland speed camera program. Proceedings 2003 Road Safety Research, Policing and Education Conference, 24–26 September, Sydney, Roads and Traffic Authority, NSW, CD-ROM 2003.

Newstead, S.V., Cameron, M.H. and Leggett, M.W. (2001). The crash reduction effectiveness of a network-wide traffic police deployment system. Accident Analysis and Prevention, 33, pp. 393-406.

Nuyts, E. (2006). Effectiviteit van onbemande camera's. Data uit vijf politiezones. RA-2006-90. Diepenbeek, Belgium: Steunpunt Verkeersveiligheid (in literature list more Belgium studies are mentioned that are not yet part of the current list).

Oei, H.L. and Polak, P.H. (1992). Effect van automatische waarschuwing en toezicht op snelheid en ongevallen. Resultaten van een evaluatie-onderzoek in vier provincies. Rapport 92-23. SWOV, Leidschendam.

Pez, P. (2002). Weniger Unfälle durch Öffentlichtkeitsarbeiet und Verkehrsüberwachung. Zeitschrift für Verkehrssicherheit, 48, pp. 58-64.

Pigman, J.G.,Agent. K.R., Deacon, J.A., and Kryscio, R.J. (1987). Evaluation of unmanned radar installations. Research Report UKTRP-87-34. University of Kentucky/ Kentucky Transportation Research Program/ College of Engineering/ Federal Highway Administration FHWA/ U.S.Department of Transportation, Lexington, Kentucky.

Roszbach, R. en A. Blokpoel (1991). Veiligheidseffecten van de invoering van 100 en 120 km/uur snelheidslimieten op autosnelwegen. Rapport 91-95. SWOV, Leidschendam.

Saunders, C.M. (1977). A study on increased intensity of traffic law enforcement as a means of reducing accidents. Report no 8. Research and Statistics Division, Road Traffic Authority, Perth, Australia.

Shumate, R.P. (1958). The effects of increased patrol on accidents, diversion and speed. Traffic Institute, Northwestern University, Evanston, Illinois.

Stefan, C. (2006). Section control – automatic speed enforcement in the Kaisermülen tunnel (Vienna, A22 Motorway). Austrian Road Safety Board (KvF), Vienna.

Swali, L. N. (1994). The effect of speed cameras in West London. Manchester, University of Manchester, 1993 (PTRC 21st Summer Annual Meeting, Traffic Management and Road Safety, Proceedings of Seminar C, pp249-260).

Tay R. (2000). Do speed cameras improve road safety? In: Traffic and transportation studies: international conference on traffic and transportation studies. Beijing, China, July 2000: 44-57.

University of Durham (1968a). The joint University of Durham/ Durham Constabulary research project on trunk road patrolling: 1967-1968. Second Report: Part I Main Experiment. University of Durham, North Carolina.

University of Durham (1968b). The joint University of Durham/ Durham Constabulary research project on trunk road patrolling: 1967-1968. Part III: developments from the 1967/1968 experiments. University of Durham, North Carolina.

University of Durham (1968c). The joint University of Durham/ Durham Constabulary research project on trunk road patrolling: 1967-1968. Part II: Other investigations arising from the main experiment. Part II. University of Durham, North Carolina.

Veling, I.H. (1992). Verkeersveiligheidscampagne N(. Samenvattend eindrapport. Rapport TT92-38. Traffic Test Bv, Veenendaal.

Wildervanck, C. (1993). 4e/5e evaluatie S2-project. Conceptverslag december 1993. ROV Groningen, Groningen.

Winnet, M. (1994). A review of speed camera operations in the UK. PTRC 22nd Summer Annual Meeting, Traffic Management and Road Safety, Proceedings of Seminar J. pp. 265-276.