Accident Analysis and Prevention 53 (2013) 9–16

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Comparison of driving simulator performance with real driving after alcohol intake: A randomised, single blind, placebo-controlled, cross-over trial

Arne Helland[a,b,*], Gunnar D. Jenssen[c], Lone-Eirin Lervåg[c], Andreas Austgulen Westin[a], Terje Moen[c], Kristian Sakshaug[c], Stian Lydersen[d], Jørg Mørland[e], Lars Slørdal[a,b]

[a] Department of Clinical Pharmacology, St. Olav University Hospital, Trondheim, Norway
[b] Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway
[c] SINTEF Technology and Society, Department of Transport Research, Trondheim, Norway
[d] The Regional Centre for Child and Adolescent Mental Health, Norwegian University of Science and Technology, Trondheim, Norway
[e] National Institute of Public Health, Oslo, Norway and Institute of Clinical Medicine, University of Oslo, Norway

ARTICLE INFO

Article history:
Received 19 September 2012
Received in revised form
20 December 2012
Accepted 28 December 2012

Keywords:
Automobile driving
Traffic accidents
Ethanol

ABSTRACT

The purpose of this study was to establish and validate a driving simulator method for assessing drug effects on driving. To achieve this, we used ethanol as a positive control, and examined whether ethanol affects driving performance in the simulator, and whether these effects are consistent with performance during real driving on a test track, also under the influence of ethanol. Twenty healthy male volunteers underwent a total of six driving trials of 1 h duration; three in an instrumented vehicle on a closed-circuit test track that closely resembled rural Norwegian road conditions, and three in the simulator with a driving scenario modelled after the test track. Test subjects were either sober or titrated to blood alcohol concentration (BAC) levels of 0.5 g/L and 0.9 g/L. The study was conducted in a randomised, cross-over, single-blind fashion, using placebo drinks and placebo pills as confounders. The primary outcome measure was standard deviation of lateral position (SDLP; "weaving"). Eighteen test subjects completed all six driving trials, and complete data were acquired from 18 subjects in the simulator and 10 subjects on the test track, respectively. There was a positive dose–response relationship between higher ethanol concentrations and increases in SDLP in both the simulator and on the test track ($p < 0.001$ for both). In the simulator, this dose–response was evident already after 15 min of driving. SDLP values were higher and showed a larger inter-individual variability in the simulator than on the test track. Most subjects displayed a similar relationship between BAC and SDLP in the simulator and on the test track; however, a few subjects showed striking dissimilarities, with very high SDLP values in the simulator. This may reflect the lack of perceived danger in the simulator, causing reckless driving in a few test subjects. Overall, the results suggest that SDLP in the driving simulator is a sensitive measure of ethanol impaired driving. The comparison with real driving implies relative external validity of the simulator.

1. Introduction

Impaired driving caused by ethanol and/or drugs is a major cause of traffic accidents, and thus a major public health problem (Blomberg et al., 2009). The relationship between blood ethanol concentrations (BAC) and accident risk is well established in large epidemiological studies (Borkenstein et al., 1974; Blomberg et al., 2009). With the exception of cannabis (Ramaekers et al., 2004), similar relationships have not been demonstrated for other psychoactive drugs and drugs of abuse. Case–control studies on non-alcohol drugs require screening and quantification of a large number of potentially impairing drugs, as well as a large number of cases, as each drug has a relatively low prevalence of detection in car crash drivers. Such studies have seldom been performed, leaving the relation between blood drug concentrations and crash risk largely unknown. Also, blood sampling for drug testing of controls – as compared to simple breath tests in ethanol studies – is necessary, and makes the recruitment of controls more difficult (Verster et al., 2009a). Furthermore, post-mortem drug concentration changes occur to a larger degree in non-alcohol drugs, making interpretation of toxicological data from studies of killed drivers difficult.

Epidemiological approaches cannot establish causal relationships, and are fraught with methodological difficulties, including

* Corresponding author at: Department of Clinical Pharmacology, St. Olav University Hospital, Pb 3250 Sluppen, NO-7006 Trondheim, Norway. Tel.: +47 72 82 91 00; fax: +47 72 82 91 10.
E-mail addresses: arne.helland@legemidler.no, arnehelland@yahoo.no (A. Helland).

0001-4575
http://dx.doi.org/10.1016/j.aap.2012.12.042

the possibility of confounding factors. Thus, experimental studies are crucial to investigate the impairing effects of drugs and the relationship between drug concentrations, impaired performance and possible accident risk. All experimental settings are *a priori* artificial, and may thus have limited external validity when applied to real driving conditions. For instance, laboratory testing of cognitive and psychomotor functioning may measure some skills that are considered essential to safe driving, but can never fully reproduce the complexity of actual driving. Real on-road driving with measurements of standard deviation of lateral position (SDLP) has come to be considered the method of reference for assessing driving impairment from CNS depressant drugs (Verster et al., 2004), although this measure reflects mainly one (i.e., automatic behaviour) of the three "core levels" of driving (Walsh et al., 2008). Much of the on-road experiments have so far been conducted in The Netherlands on flat, straight multi-lane motorways; a driving scenario that may not reflect conditions elsewhere. Also, legal issues and safety considerations may hinder on-road experiments, and the costs of such experiments may be prohibitive.

Experimental studies utilising driving simulators may avoid some of the problems listed above. However, even very sophisticated simulators cannot fully replicate real driving conditions (Verster et al., 2004; Shechtman et al., 2009). Driving simulator studies of effects of depressant drugs on driving ability frequently yield inconclusive results due to the lack of validation against a known positive control; in practice, ethanol. The positive control is necessary to ensure that correlations between drug intake and driving related outcome measures actually reflect a drug related impairment of driving ability, and not simply randomly observed correlations with no relevance to impairment (Walsh et al., 2008). Ethanol as a positive control also ensures that the experimental design is sufficiently sensitive to the impairing effects of depressant drugs. Another common limitation of driving simulators is the lack of validation against a real driving scenario; i.e., the external validity. This leaves doubt as to whether test subject performance in the simulated scenario may predict performance in real driving situations.

We wanted to develop a valid and functional tool for assessing drug effects on driving performance, taking into account the recommendations made in the guidelines for research on drugged driving. To achieve this, we conducted a validation study of the SINTEF driving simulator. The purpose of the study was to establish a driving simulator test battery that is sensitive to ethanol effects, and to validate the test battery by comparing performance in the simulator with actual driving performance on a closed-circuit test track resembling rural driving conditions. Even though both simulator and closed circuit driving constitute experimental conditions, which do not fully reproduce the real life driving experience, both are widely used for assessing driving performance, and real driving is generally considered to be the reference methodology as far as validity is concerned. In this paper we present results from the primary outcome measure SDLP, measured in the simulator and on the test track.

2. Materials and methods

2.1. Test subjects

Twenty healthy, Caucasian, male volunteers aged 25–35 years (mean 28.7 years) who had been in possession of a driver's license for at least 5 years (mean 10.6 years), were included in the study. They were all recreational users of alcohol, and as a group drove slightly less and had a somewhat higher educational level than the general population. Women and non-Caucasians were excluded because of the teratogenic risk associated with ethanol use in the former group, and the possibility of deviant ethanol metabolism in the latter. The other exclusion criteria were previous or present drug or alcohol abuse or atypical reactions to alcohol, previous history of driving under the influence, significant adverse reactions to previous blood sampling, regular (daily) intake of any prescribed drug, or high likelihood of motion sickness as assessed with a modified version of the Apfel risk score for postoperative vomiting (Apfel et al., 1998). Each participant underwent a screening for eligibility, received written and oral information about the study and provided a written consent to participate. The study was approved by the Regional Ethics Committee, and was registered as a clinical trial in the ClinicalTrials.gov database. All participants received a gift certificate worth NOK 1000 (approx. USD 150) upon completion of the study.

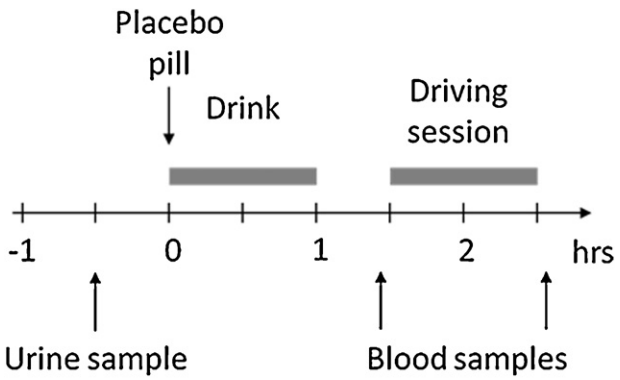

Fig. 1. Outline of trial test design.

2.2. Trial design

The experiment was designed as a randomised, placebo-controlled, single blind, crossover study. Only the necessary personnel were informed about which interventions were given. An outline of the design is presented in Fig. 1. Each participant underwent three driving tests of 1 h duration, both on a closed-circuit test track and in an advanced driving simulator, on six different test days with washout periods of minimum two days between test days to allow the dissipation of any learning or fatigue effects. The driving scenario in the simulator was modelled to mimic the test track, as illustrated in Fig. 2, to ensure that the driving experience would be as similar as possible in the two test conditions. Before testing commenced, the study subjects undertook a training session, both on the test track and in the simulator, in order to familiarise themselves with the testing scenario and minimise the impact of possible learning effects. On test days, the participants were obliged to deliver a urine sample on arrival at the test site to exclude the presence of drugs. The subjects' weight was registered each test day, after which they were administered a weight-adjusted dose of ethanol (0, 0.7 and 1.05 g per kg body weight), calculated to obtain an intended blood alcohol concentration (BAC) during testing of 0, 0.5 and 0.9 g/L on the three different test days both in the simulator and on the test track, respectively. The Widmark equation (Andréasson and Jones, 1995), was used to estimate the ethanol doses, assuming a total body water to total body mass ratio of 0.68, a bioavailability of 75%, and a metabolic rate for ethanol of 0.15 g/L per hour. We used vodka mixed with fruit extracts, orange and lime juice to make the drinks palatable. The placebo drinks were spiked with non-alcoholic vodka flavour in water to mimic the vodka taste. The drinks were served in closed plastic containers, from which the participants were instructed to sip the drink through a straw. To avoid an obvious ethanol taste, no drinks were stronger than 10% (v/v) ethanol, and they were kept cold by the addition of ice.

Fig. 2. Example of the driver's visual impression on the closed-circuit test track (left) and in the driving simulator (right).

The participants were allowed 1 h to finish their drinks, after which they waited another 30 min before the driving test started, to allow for absorption of the administered ethanol. The order in which the participants were tested at different BAC levels was randomised by use of a counterbalanced, multi-condition design. The same order of BAC levels was used for each participant both on the test track and in the simulator. As an additional confounder to enhance blinding, the study subjects were administered a placebo pill, which they were told may or may not contain a sedative drug, with the drink. Venous blood samples were drawn immediately before and after each driving session, and the mean value was used as the best estimate of the mean BAC during testing.

2.3. Real driving on test track

The test track driving was undertaken during a frost-free period of six weeks in the autumn. All study sessions were done after nightfall, between 20:00 and 01:00 h. The test track circuit was 1.37 km long, closed to ordinary traffic, and laid out in hilly terrain, with both gentle and sharper curves. The track was hard-surfaced, with two lanes each approx. 2.75 m wide, and had midline and side markings similar to standard Norwegian road markings. Thus, the test track closely resembled roads typical of rural Norway. Surprise obstacles (1 m^3 foam rubber cubes) were placed in two locations on two occasions, one at the beginning and one towards the end of each driving trip, and were to be avoided by the test subjects. Stoplights present in two locations turned red on one occasion during each trip. The participants drove an instrumented car (Volvo V70 2.4s) with automatic transmission, fitted with a double set of pedals. They were instructed to drive as they would normally do on a regular road. A professional driving instructor was present in the front passenger seat during all sessions of test track driving, in order to intervene if necessary. A physician was present on the site at all times during test drives. Permission to carry out the test track driving was granted from the local police. To enable continuous recording of lateral position in the road lane, the test car was equipped with an infrared wide-angle camera fixed to the roof of the car, and pointing at a downward angle to the rear of the car. The data were stored in a database and analysed in a program for photo analysis (Open Source Computer Vision Library). A filtering algorithm (Hough transformation) was used to identify roadside markings. The car also featured other equipment for recording the location of the car on the test circuit (global positioning system; GPS), speed, pedal use and steering wheel movements.

2.4. Driving simulator

Testing in the driving simulator took place in late autumn after the test track driving tests were completed. Test sessions were done at the same times during the evening and night as on the test track, using a virtual model of the test track and a nighttime scenario (Fig. 2), to ensure comparable results and eliminate differences in circadian influences. In addition to obstacles and stoplights, the simulator scenario also included two incidents (a car abruptly entering the road and a pedestrian crossing the road in front of the driver) that each occurred once at the end of the driving session. The simulator had the appearance of a regular car (Renault Scenic) with automatic transmission and original controls (Fig. 3). Information from the use of steering wheel, pedals, transmission etc. was fed into a dedicated driving scenario graphics computer. The driving scenario was depicted on screens covering 180° of the driver's forward field of vision and 90° of the rear field of vision, and synchronously in internal and external mirrors. The vertical field of view was 47° both to the front and to the rear. The simulator reproduced realistic motion, vibration and sound through a three-axis moving platform, a vibration system in the chassis and a four-channel sound system. Data on lateral position, speed, pedal use and steering wheel movements over the entire duration of the test sessions were extracted directly from the simulator computer and logged 20 times per second. A detailed description of the SINTEF simulator can be found in Engen (2008).

2.5. Measurements

The predefined primary outcome measure was the standard deviation of lateral position (SDLP), which is a measure of the degree of weaving of the car on the road. SDLP has been shown to correlate with BAC levels in a dose dependent manner, and is a thoroughly validated measure of the degree of driving impairment (Verster et al., 2004). Secondary outcome measures were number of brake pedal pressures per lap, number of accelerator

Fig. 3. Setup of the driving simulator. Vehicle and surrounding frontal screens.

pedal pressures per lap, steering wheel movement speed, steering wheel movement per lap, steering wheel reversals per lap, steering wheel reversal frequency, average speed, standard deviation of speed (measured continuously throughout the driving sessions), driving behaviour at unexpected incidents, and driving against red light. We aim to present the secondary outcome measures in a subsequent article.

Before and after each driving session, the participants completed a questionnaire, with items covering their feelings of intoxication, mastery, safety, sleepiness, alertness, whether they thought the drink had contained ethanol, and whether they thought the pill had contained a sedative drug. At the test track, driving instructors were also asked to rate the test subjects' degree of intoxication and driving performance.

Blood ethanol concentrations were quantified using a headspace gas chromatography–mass spectrometry (GC–MS) method. In brief, 200 μL blood was mixed with 50 μL internal standard (d6-ethanol). Samples were left for 30 min to achieve equilibrium before the gas fraction was aspirated into an Agilent HP 6890-5973 GC–MS system (Agilent, Palo Alto, CA). Separation was performed on a J&W Scientific 123-9134 DB-ALC1 (30 m × 1.2 mm) column with a helium mobile phase and a run time of 0.90 min. Ethanol was monitored at m/z 31 and the internal standard at m/z 33. The level of quantification (LOQ) was 2 mmol/L (approx. 0.09 g/L). Between-day coefficient of variation (CV) calculated from quality control samples was 4.5% at 5 mmol/L (0.22 g/L) and 1.8% at 50 mmol/L (2.2 g/L).

2.6. Statistical analyses

An *a priori* sample size estimation performed with one-tailed, paired *t*-tests indicated that a total sample size of $n = 11$ would be sufficient to detect significant differences in BAC level influence on SDLP with significance level (α) of 0.05 and power ($1-\beta$) of 0.95. Although theoretically 11 subjects would suffice, we chose to include 20 subjects in the study, to allow for the uncertainty in the underlying assumptions of the sample size estimation, as well as the possibility of dropouts, for instance due to simulator sickness.

In the results analyses, we used a linear mixed model with SDLP as dependent variable, measured BAC as covariate, and participant as random effect. Separate analyses were performed for test track and simulator. Reported results are from restricted maximum likelihood estimation. The maximum likelihood estimation did not always converge. The independent variables tested for significance were BAC level, curved/straight section and part of trip driven (each trip was divided in four equal parts of 15 min). To identify possible learning effects that could interfere with the results, the impact of the number of trips driven before the actual one was also analysed. Two-sided *p*-values <0.05 were considered significant. The analyses were performed in SPSS 18 and Stata 12.

3. Results

Of the 20 participants enrolled in the study, all completed three driving sessions on the test track, while 18 out of 20 completed all three sessions in the driving simulator. Two subjects did not complete the simulator testing; one because of intolerable nausea, and the other because of a surgical procedure unrelated to the study. On the test track, 10 out of the 60 driving sessions did not yield sufficient SDLP data to be included in the analyses. The car-mounted camera was out of position in eight sessions, the camera was not switched on in one instance, and one participant in his first session misinterpreted the instructions to drive in lane. Thus, a complete set of outcome data was obtained from 10 participants on the test track and 18 participants in the simulator. Data from the valid driving sessions of all subjects were included in the analyses.

Table 1
Measured blood ethanol concentrations (BAC) in simulator driving and on test track at the three designated BAC levels of 0, 0.5 g/L and 0.9 g/L among all test subjects with samples.

Test scenario	Intended BAC	Mean BAC (±SD)
Simulator ($n = 19$)	0	0
	0.5 g/L	0.38 (±0.10) g/L
	0.9 g/L	0.82 (±0.19) g/L
Test track ($n = 20$)	0	0
	0.5 g/L	0.42 (±0.09) g/L
	0.9 g/L	0.88 (±0.12) g/L

3.1. Safety and adverse events

No safety violations or serious or unexpected adverse events occurred during the study. The most common adverse event in the simulator was nausea, which is a known disadvantage of driving simulators. Six subjects (four at BAC 0 and two at BAC 0.5) had to terminate their first simulator session early because of this, but five of them were eventually able to complete all three sessions. Thus, only one subject had to withdraw from the study due to nausea. Prior experience suggests that ethanol may protect against simulator sickness, and repeated exposures to the simulator tend to attenuate the nausea. Therefore, in order to prevent dropouts, all participants who terminated their sessions early due to nausea were tested at the highest BAC level in the subsequent session. The random order was also modified in an additional three subjects due to other practical causes. These modifications to the randomisation did not affect concealment of the interventions, and did not appear to introduce systematic bias, since there was no statistically significant correlation between BAC level and the number of previous test sessions (Pearson correlation 0.241 ($p = 0.080$) in simulator and 0.094 ($p = 0.477$) on test track).

3.2. Blood alcohol concentrations

The ethanol concentrations are presented in Table 1. Ethanol concentrations were slightly lower than intended both in the simulator and on the test track, with concentrations closer to 0.4 g/L at the intended level of 0.5 g/L. The BAC also tended to be slightly lower in the simulator than on the test track. Paired sample *t*-test showed a statistically significant difference between the BAC levels in simulator and on test track for the designated BAC level of 0.5 g/L ($p = 0.041$); however, the mean difference was only 0.039 g/L. For the designated BAC level of 0.9 g/L, there was no statistically significant difference between BAC levels in simulator and on test track ($p = 0.21$). In the following, ethanol levels are referred to as the intended levels (BAC 0, BAC 0.5 and BAC 0.9, respectively).

3.3. Questionnaires

After each driving session, the participants were asked whether they thought the drink and the pill had contained alcohol and a sedative drug, respectively. Most subjects correctly identified the drink as containing/not containing ethanol (in 32 of 38 placebo trials, 35 of 38 BAC 0.5 trials and 37 of 38 BAC 0.9 trials, respectively). However, a few misidentified their drinks, and quite a few wrongly identified the pill as containing a sedative drug (in 15 of 38 placebo trials, 3 of 38 BAC 0.5 trials and 7 of 38 BAC 0.9 trials, respectively).

There were significant correlations between higher BAC levels and subjective (self reported) ratings of poorer driving performance both in the simulator ($R = 0.35$, $p = 0.013$) and on the test track ($R = 0.63$, $p < 0.001$). Likewise, there was a strong correlation between higher BAC levels and objective (driving instructor

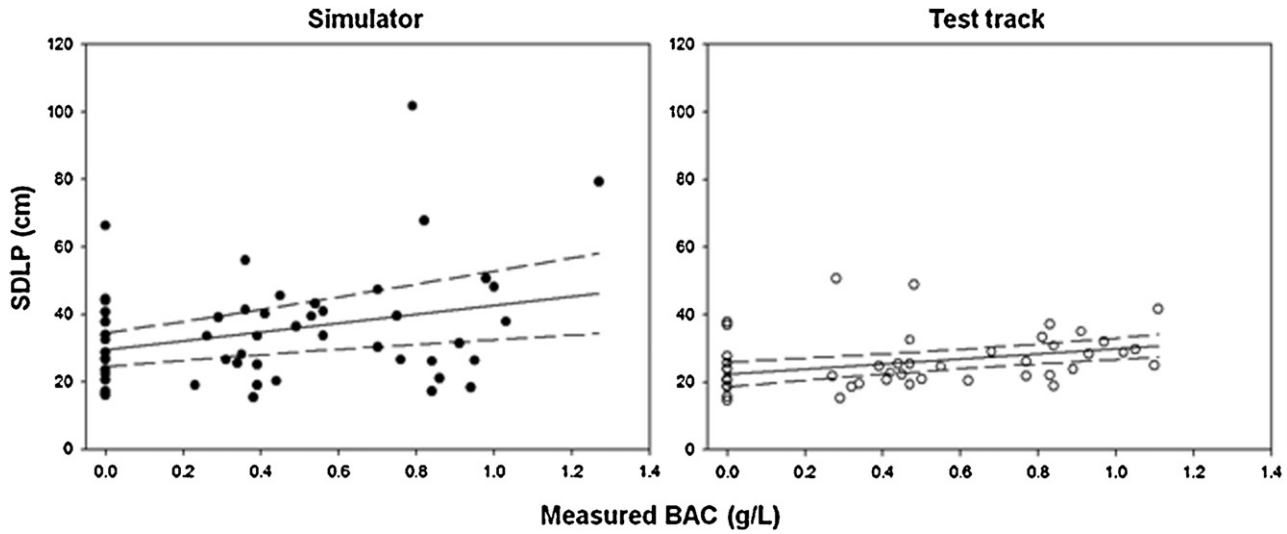

Fig. 4. Regression analysis of the relationship between blood alcohol concentration (BAC) and standard deviation of lateral position (SDLP) in simulator (left; filled circles) and on test track (right; open circles). The circles represent individual BAC and the corresponding SDLP value. The regression lines and their 95% confidence intervals are shown as continuous lines and broken lines, respectively.

reported) ratings of poorer driving performance on the test track ($R = 0.52$, $p < 0.001$).

3.4. SDLP

Fig. 4 shows the individual SDLP values at the corresponding BAC, with the estimated regression line and its 95% confidence interval. Both in the simulator and on the test track, there were significant positive correlations between BAC and SDLP (positive regression slope with $p < 0.001$). The estimated regression lines for the simulator (Eq. (1)) and the test track (Eq. (2)) are as follows, with standard errors for the estimates in parentheses:

$$(\text{simulator}) : \text{SDLP}(\text{cm}) = 29.43\,(\pm 2.57) + 13.20\,(\pm 3.61) \times \text{BAC} \quad (1)$$

$$(\text{test track}) : \text{SDLP}(\text{cm}) = 22.30\,(\pm 1.89) + 7.61\,(\pm 1.91) \times \text{BAC} \quad (2)$$

SDLP values were higher in the simulator than on the test track at baseline (placebo) conditions (29.4 cm vs. 22.3 cm, respectively), and showed a steeper increase with increasing BAC, as seen from Eqs. (1) and (2), as well as Fig. 4. As evident from Fig. 4, SDLP variance was also larger in simulator driving than in test track driving.

The relationship between BAC levels and SDLP results show a dose–response effect, as quantified by the slopes 13.20 and 7.61 in Eq. (1) and (2). Furthermore, a visual comparison of SDLP results in the simulator and on the test track in each of the 20 individual subjects shows similar, positive slopes in most subjects (Fig. 5).

To identify possible differential effects of test duration and curved/straight sections on SDLP, the SDLP results were analysed with respect to time intervals (four equal intervals of 15 min each), and performance on curved and straight sections of the driving scenario. In the simulator, mean SDLP values were significantly higher in curved sections than in straight sections ($p = 0.047$), whereas there were no such differences on the test track ($p = 0.17$). In the simulator, statistically significant differences in SDLP between BAC levels were seen in all four time intervals. On the test track, the differences in SDLP were similar but less pronounced, and mostly did not reach significance during the first half hour of the test. In the simulator, there was a trend towards higher SDLP values with longer test duration, especially at the highest BAC level. No such tendency was evident on the test track.

To identify possible learning effects that would be expected to reduce SDLP with the number of prior test sessions, the number of trips driven before the actual one was also analysed as an independent variable. However, this had no statistically significant correlation with SDLP results either in the simulator ($p = 0.70$) or on the test track ($p = 0.66$).

4. Discussion

4.1. SDLP

Our results show a positive dose–response correlation between BAC and SDLP in the simulator and on the test track, both for individual and mean data. A high degree of intra-individual similarity in the BAC-correlated increase in SDLP in the simulator and on the test track, suggests that SDLP is a valid and sensitive measure of ethanol-induced driving impairment in the simulator.

Absolute values of SDLP were higher in the simulator than on the test track, with mean SDLP at BAC 0 (sober state) of 29 cm and 22 cm, respectively. SDLP values during placebo conditions in the simulator were also considerably higher than those seen in Dutch on-road driving tests, where mean baseline SDLP is approx. 19 cm (range 9–30 cm) (Verster and Roth, 2011). The relatively demanding driving scenario that was used in our experiment may account for the slightly higher SDLP values on the test track than those seen during previous on-road tests. Higher absolute SDLP values in the simulator compared to real driving may be explained by unfamiliarity with the driving experience in the simulator, a lack of perceived danger, and lack of gravitational cues and feedback that will normally adjust steering. This notion is also supported by the observation that SDLP values were higher in curved sections than in straight sections in the simulator, whereas such a difference was not observed on the test track. Together with the more demanding driving scenario in our experiment, this may account for the considerably higher SDLP values than those seen for instance in the Dutch STISIM simulator employing a monotonous highway scenario (Mets et al., 2011b).

Most test subjects showed similar SDLP increases in the simulator and on the test track. However, from the individual SDLP data shown in Fig. 5, a few subjects behave differently, evidenced by excessive SDLP values in the simulator. For instance, test subject no. 15 had a mean SDLP exceeding 1 m at the highest BAC level. This would correspond to the car being located mostly out of lane during the trip, which is in accordance with the actual observations made

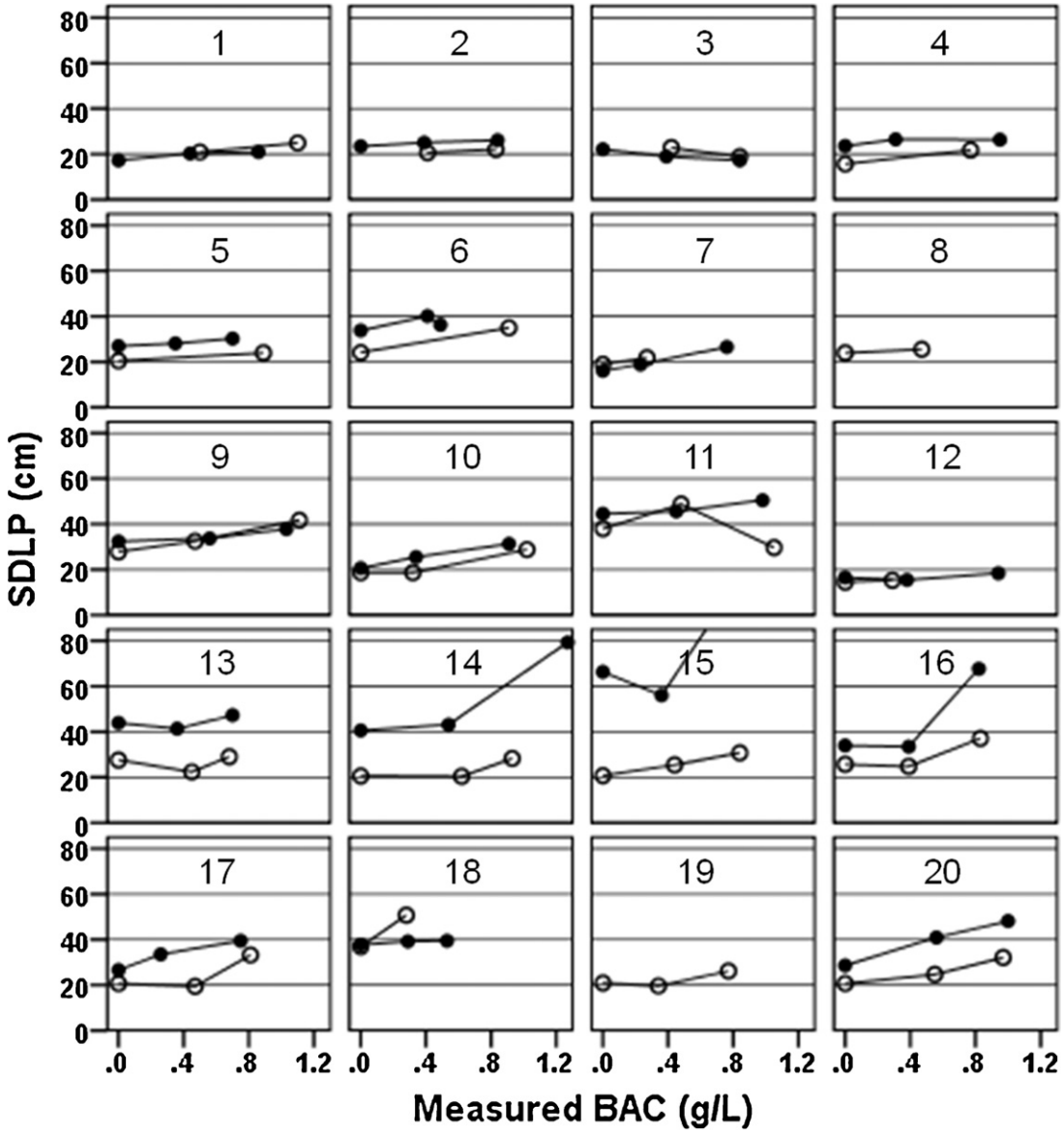

Fig. 5. Individual SDLP data at actual BAC levels in simulator (filled circles) and on test track (open circles). For test subject 15, the BAC and SDLP values at the highest BAC level in the simulator were 0.79 g/L and 102 cm, respectively.

during this individual's simulator driving. It is our experience from the present and earlier simulator experiments that some participants regard the simulator as a kind of game and behave more like virtual rally drivers instead of following the instructions to drive appropriately according to conditions. This can explain the large discrepancies in SDLP between test track and simulator seen in a few of the subjects. Subject no. 14 attained an unexpectedly high BAC at his highest BAC level in the simulator (1.25 g/L), which may explain the high SDLP observed in that driving session. Also, we cannot exclude the possibility that some participants' SDLP scores were influenced by simulator sickness.

4.2. BAC

Mean subjective and objective ratings of intoxication and driving performance correlated with BAC level in the expected manner. The somewhat lower BAC levels in simulator than on test track may be due to a possible conditioned nausea response in the simulator that could have caused retention of stomach content with delayed ethanol absorption. One participant (subject no. 6) was unable to finish his drink at the intended BAC 0.9 level in the simulator, and consequently acquired a low BAC.

Most participants correctly identified their drink as containing/not containing ethanol and the pill as containing/not containing a sedative drug, although quite a few of the participants misidentified the placebo pill, especially in the BAC 0 trials. This probably reflects an expectation bias in some subjects, and indicates that the use of placebo pills to enhance blinding of the intervention in experimental trials with ethanol may be worthwhile. Previous experience suggests that concealment of ethanol is difficult in blinded studies due to the distinctive taste and smell and the characteristic and familiar effects of ethanol.

4.3. Comparison with other driving simulator studies and on-road tests

To date, there are few other studies validating the use of driving simulators for drug and/or ethanol impairment research. A simulator validation study published in 2009 used data from two separate previous studies (on-road and in simulator). The

description of the simulator they used suggests that it was similar to the SINTEF simulator, but the driving scenario and outcome measure were different (urban traffic and number of driving errors at intersections as assessed by a driving instructor, respectively). No ethanol or other drugs were used. Results indicated relative validity for the simulator, and suggested absolute validity for the type of errors pertaining lane maintenance, adjustment to stimuli and visual scanning (Shechtman et al., 2009).

There are few previous simulator studies using SDLP as outcome measure. Only one study has validated SDLP as an indicator of unsafe driving in the simulator that was used. Mets et al. published a validation study in 2011 showing the ability of the STISIM driving simulator to differentiate between different BAC levels based on SDLP results. In this study, 27 healthy volunteers underwent a simulator adaptation of the standardised Dutch on-road test scenario (multi-lane highway driving for 1 h). BAC levels of 0 g/L, 0.5 g/L, 0.8 g/L and 1.1 g/L yielded mean SDLP values of 28.0 cm, 29.7 cm, 33.8 cm and 36.3 cm, respectively. This study did not validate the simulator results against a real driving test (Mets et al., 2011b). Apart from this, only two simulator studies concerning driving performance after drug intake have been published using SDLP as an outcome measure. Mets et al. have investigated the effects of caffeine (given in the form of the energy drink Red Bull® and coffee, respectively) on driving performance in healthy volunteers in two studies in the Dutch STISIM simulator, and found small but significant reductions in SDLP after caffeine administration in both studies (Mets et al., 2011a, 2012).

In 2009, a validation study with ethanol in a divided-attention steering simulator (DASS) was published. As the name suggests, the simulator is designed to measure ability of divided attention. Accordingly, it employs a rather artificial test scenario, where subjects must keep the car in lane and simultaneously respond to peripheral visual stimuli. Also, the simulator used did not resemble a normal car. Dose-dependent impairment was found with higher ethanol levels (Verster et al., 2009b).

The standardised on-road driving test with SDLP as the outcome measure developed in The Netherlands remains the method of reference to examine driving impairment from drugs. In such testing, BAC levels of 0.5 g/L and 0.8 g/L on average increases SDLP from placebo conditions with 2.4 cm and 4.3 cm, respectively (Verster and Roth, 2011). Our results from the test track show slightly larger increases in SDLP, whereas the BAC-related increases in the simulator were considerably larger. Again, the discrepancy between our results and the Dutch on-road results may be explained by the more demanding driving scenario employed in our validation study.

4.4. Implications for the validity and further use of the simulator

External validity of a driving simulator refers to the test scenario's ability to invoke similar reactions in the drivers as a real driving scenario. Validity is specific for the particular type of scenario and simulator, test, and population used in the validation experiments, and will not necessarily be transferable to other driving scenarios, simulators, tests, or populations. External validity is absolute if the same effect is invoked to the same extent both in the simulator and in the real driving environment. Relative external validity implies that there exists a trend of change in the same direction both in the simulator and in the real driving environment, but the magnitude of change is different (Shechtman et al., 2009).

There was a large degree of similarity in the relationship between SDLP and BAC levels in the simulator and on the test track. However, the absolute values of SDLP in the simulator were consistently higher than on the test track. Thus, the relative (but not the absolute) external validity of the SINTEF simulator has been established when validated against test track driving in a driving scenario that is representative of the demanding rural driving conditions in Norway, using ethanol as a positive control. We believe that this validation may be extended to real driving under similar conditions; however, this assumption has not been proven.

In the simulator, we found consistent and significant BAC-related increases in SDLP in all time intervals when the hour-long test was divided into four 15-min time intervals. This suggests that the duration of the simulator test in order to reach significant results may be shortened in future studies.

4.5. Limitations of the study

In our study, all test subjects were healthy young male volunteers, who are not representative for the general driving population. Our results may therefore give a somewhat inaccurate estimation of the impact of BAC on SDLP in the general population.

There are three levels of behaviour relevant to traffic safety: automatic, control and executive planning behaviour (Michon, 1985; Walsh et al., 2008). SDLP as the primary outcome measure in this study is mainly representative for the effect of ethanol on automated actions at a behavioural control level. Outcome measures of driving behaviour at manoeuvring and strategic levels will be reported in a separate publication. Driving simulators may be especially suitable to test higher behavioural levels like hazard avoidance, dual attention, risk taking and impulsivity, both for ethical (risk of injury) and practical (ease and reliability of measurements) reasons.

We employed a single blind design, keeping the intervention concealed from the test subjects but not from the study personnel or those responsible for analysing the outcome data.

Unlike some of the most advanced simulators in use, the SINTEF simulator allows only limited tilting (three degrees of freedom). Motion-based simulators with full tilting technology might increase the realism of the driving experience, and thus heighten the external validity of the simulator.

Several of the test subjects experienced nausea in the simulator, which caused one subject to withdraw from the study, and may have affected driving behaviour in others. This is a general drawback of driving simulators, which may to some extent be unavoidable, even when using screening procedures including test drives before enrolment. We also employed a rather challenging driving scenario, with many curves and long duration, which may have exacerbated the problems related to nausea.

The validation against real driving was done on a closed test track. The length (approx. 1.4 km) and layout (curvy, hard-top road approx. 5.5 m wide with midline and side markings) of the test track ensured that the driving experience resembled real driving on rural Norwegian roads. However, it may be impossible to fully eliminate the feeling of an artificial situation when driving on a closed test track. For safety reasons, a driving instructor was present in the passenger seat at all times on the test track, as well as a police officer on the test track site. This may have constituted a restraining effect as well as heightened the attention of test subjects, causing them to drive more carefully and attentively than they would otherwise have done.

Finally, our study had a limited sample size, which generally increases the risk of type II errors (i.e., failing to detect real differences). Also, missing data from 10 of 60 driving sessions on the test track may have limited the statistical significance of our findings. The missing data occurred due to random incidents, and we have no reason to believe this introduced systematic bias.

5. Conclusions

In healthy volunteers, SDLP as a measure of drug-impaired driving shows qualitatively similar outcomes during test track driving and in a driving simulator designed to mimic the test track, both sober and under the influence of ethanol. However, SDLP is amplified in the simulator as compared to real driving. Although closed circuit driving is an experimental situation and thus of limited external validity, the quantitative and qualitative similarities between simulator and test track driving nevertheless imply external validity of the simulator. In conclusion, the SINTEF driving simulator is a sensitive and valid tool to assess driving impairment from ethanol, and this may be extended to include other CNS depressant drugs.

Acknowledgements

The research project was funded by a grant from the Research Council of Norway.

The authors would like to thank the following employees at SINTEF Technology and Society and the Department of Clinical Pharmacology at St. Olav University Hospital who made valuable contributions to the research project: Thomas Engen, Dagfinn Moe, Kenneth Sørensen, Terje Giæver, Cato Mausethagen, Nina Husby, Marianne Nordtømme, Odd Hjelkrem, Guro Haltbrekken, Anita Skogholt, Trude Kristine Vaag Gjerde.

We would also like to thank Prof. Dr. Jan Ramaekers at Maastricht University, The Netherlands, for valuable advice and helpful discussions during the preparations of the study.

Finally, we would like to express our gratitude to the study subjects who uncomplainingly underwent the rather demanding trials.

References

Andréasson, R., Jones, A.W., 1995. Erik M. P. Widmark (1889–1945): Swedish pioneer in forensic alcohol toxicology. Forensic Science International 72 (1), 1–14.

Apfel, C.C., Greim, C.A., Haubitz, I., Goepfert, C., Usadel, J., Sefrin, P., Roewer, N., 1998. A risk score to predict the probability of postoperative vomiting in adults. Acta Anaesthesiologica Scandinavica 42 (5), 495–501.

Blomberg, R.D., Peck, R.C., Moskowitz, H., Burns, M., Fiorentino, D., 2009. The Long Beach/Fort Lauderdale relative risk study. Journal of Safety Research 40 (4), 285–292.

Borkenstein, R.F., Crowther, R.F., Shumante, R.P., Ziel, W.B., Zylman, R., 1974. Die Rolle des alkoholisierten Fahrers bei Verkehrsunfällen (the Grand Rapids study). Blutalkohol 11, 1–7.

Engen, T., 2008. Use and validation of driving simulators. Thesis. Norwegian University of Science and Technology.

Mets, M.A., Baas, D., Van Boven, I., Olivier, B., Verster, J.C., 2012. Effects of coffee on driving performance during prolonged simulated highway driving. Psychopharmacology (Berlin) 222 (2), 337–342.

Mets, M.A., Ketzer, S., Blom, C., Van Gerven, M.H., Van Willigenburg, G.M., Olivier, B., Verster, J.C., 2011a. Positive effects of Red Bull® energy drink on driving performance during prolonged driving. Psychopharmacology (Berlin) 214 (3), 737–745.

Mets, M.A., Kuipers, E., Senerpont Domis, L.M., Leenders, M., Olivier, B., Verster, J.C., 2011b. Effects of alcohol on highway driving in the STISIM driving simulator. Human psychopharmacology. [Epub ahead of print].

Michon, J.A., 1985. A critical view of driver behavior models: what do we know, what should we do? In: Evans, L., Schwing, R.C. (Eds.), Human Behavior and Traffic Safety. Plenum Press, New York, pp. 485–520.

Ramaekers, J.G., Berghaus, G., Van Laar, M., Drummer, O.H., 2004. Dose related risk of motor vehicle crashes after cannabis use. Drug and Alcohol Dependence 73 (2), 109–119.

Shechtman, O., Classen, S., Awadzi, K., Mann, W., 2009. Comparison of driving errors between on-the-road and simulated driving assessment: a validation study. Traffic Injury Prevention 10 (4), 379–385.

Verster, J.C., Pandi-Perumal, S.R., Ramaekers, J.G., De Gier, J.J., 2009a. Drugs, Driving and Traffic Safety. Birkhäuser Verlag AG, Basel, Switzerland.

Verster, J.C., Roth, T., 2011. Standard operation procedures for conducting the on-the-road driving test, and measurement of the standard deviation of lateral position (SDLP). International Journal of General Medicine 4, 359–371.

Verster, J.C., Veldhuijzen, D.S., Volkerts, E.R., 2004. Residual effects of sleep medication on driving ability. Sleep Medicine Reviews 8 (4), 309–325.

Verster, J.C., Wester, A.E., Goorden, M., Van Wieringen, J.P., Olivier, B., Volkerts, E.R., 2009b. Novice drivers' performance after different alcohol dosages and placebo in the divided-attention steering simulator (DASS). Psychopharmacology (Berlin) 204 (1), 127–133.

Walsh, J.M., Verstraete, A.G., Huestis, M.A., Mørland, J., 2008. Guidelines for research on drugged driving. Addiction 103 (8), 1258–1268.