



Universiteit
Leiden
The Netherlands

Using machine learning techniques to characterize sleep-deprived driving behavior

Wall, H.E.C. van der; Doll, R.J.; Westen, G.J.P. van; Koopmans, I.; Zuiker, R.G.; Burggraaf, J.; Cohen, A.F.

Citation

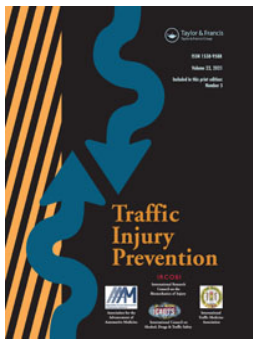
Wall, H. E. C. van der, Doll, R. J., Westen, G. J. P. van, Koopmans, I., Zuiker, R. G., Burggraaf, J., & Cohen, A. F. (2021). Using machine learning techniques to characterize sleep-deprived driving behavior. *Traffic Injury Prevention*, 22(5), 366-371. doi:10.1080/15389588.2021.1914837

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3249267>

Note: To cite this publication please use the final published version (if applicable).






Using machine learning techniques to characterize sleep-deprived driving behavior

H. E. C. van der Wall, R. J. Doll, G. J. P. van Westen, I. Koopmans, R. G. Zuiker, J. Burggraaf & A. F. Cohen


To cite this article: H. E. C. van der Wall, R. J. Doll, G. J. P. van Westen, I. Koopmans, R. G. Zuiker, J. Burggraaf & A. F. Cohen (2021) Using machine learning techniques to characterize sleep-deprived driving behavior, Traffic Injury Prevention, 22:5, 366-371, DOI: [10.1080/15389588.2021.1914837](https://doi.org/10.1080/15389588.2021.1914837)

To link to this article: <https://doi.org/10.1080/15389588.2021.1914837>

 View supplementary material 

 Published online: 07 May 2021.

 Submit your article to this journal 

 Article views: 163

 View related articles 

 View Crossmark data 



Using machine learning techniques to characterize sleep-deprived driving behavior

H. E. C. van der Wall^{a,b}, R. J. Doll^a, G. J. P. van Westen^b, I. Koopmans^{a,c}, R. G. Zuiker^a, J. Burggraaf^{a,b,c}, and A. F. Cohen^{a,b,c}

^aCentre for Human Drug Research, Leiden, the Netherlands; ^bLeiden Academic Centre for Drug Research, Leiden, the Netherlands; ^cLeiden University Medical Center, Leiden, the Netherlands

ABSTRACT

Objective: Sleep deprivation is known to affect driving behavior and may lead to serious car accidents similar to the effects from e.g., alcohol. In a previous study, we have demonstrated that the use of machine learning techniques allows adequate characterization of abnormal driving behavior after alprazolam and/or alcohol intake. In the present study, we extend this approach to sleep deprivation and test the model for characterization of new interventions. We aimed to classify abnormal driving behavior after sleep deprivation, and, by using a machine learning model, we tested if this model could also pick up abnormal driving behavior resulting from other interventions.

Methods: Data were collected during a previous study, in which 24 subjects were tested after being sleep-deprived and after a well-rested night. Features were calculated from several driving parameters, such as the lateral position, speed of the car, and steering speed. In the present study, we used a gradient boosting model to classify sleep deprivation. The model was validated using a 5-fold cross validation technique. Next, probability scores were used to identify the overlap of driving behavior after sleep deprivation and driving behavior affected by other interventions. In the current study alprazolam, alcohol, and placebo are used to test/validate the approach.

Results: The sleep deprivation model detected abnormal driving behavior in the simulator with an accuracy of $77 \pm 9\%$. Abnormal driving behavior after alprazolam, and to a lesser extent also after alcohol intake, showed remarkably similar characteristics to sleep deprivation. The average probability score for alprazolam and alcohol measurements was 0.79, for alcohol 0.63, and for placebo only 0.27 and 0.30, matching the expected relative drowsiness.

Conclusion: We developed a model detecting abnormal driving induced by sleep deprivation. The model shows the similarities in driving characteristics between sleep deprivation and other interventions, i.e., alcohol and alprazolam. Consequently, our model for sleep deprivation may serve as a next reference point for a driving test battery of newly developed drugs.

ARTICLE HISTORY

Received 10 August 2020
Accepted 5 April 2021

KEYWORDS



Safety; machine learning; automobile driving; driving under the influence

Introduction

Research into abnormal driving behavior is needed as car-drivers have a potential risk to become involved in a crash and compromise traffic safety of others and themselves. The risk on abnormal driving behavior and ensuing car accidents depends on numerous factors, such as predisposing driving style and individual characteristics (e.g., age, gender) of the car-driver and intake of alcohol (Irwin et al. 2015; Sagberg et al. 2015). CNS-active medicines and recreational substances may also negatively affect car-driving behavior (Arnedt et al. 2000; Houwing et al. 2012; Mets et al. 2011; Robertson et al. 2017).

To quantify abnormal driving behavior many researchers have used the standard deviation of the lateral position (SDLP) of the car on the road as a valid measure (Darby et al. 2009; Mets et al. 2011; Verster and Roth 2011). In a

recent study we have shown that by using machine learning, a more sensitive driving measure based on multiple driving features can be created (van der Wall et al. 2020). In that study, two models were developed that were able to classify driving behavior affected by either alcohol or by alprazolam. Moreover, our results suggested that a series of these machine learning models could evolve to a test battery, allowing a more precise and accurate evaluation of abnormal driving behavior in the process of new drug development. However, the generalizability of such a model is still unknown. At the moment, it has only been shown that such a model can recognize solely the drug that has been used for the development of the model. However, the ultimate goal would be to test new drugs or interventions with (a selection of) these models. A model for detection of sleep-deprived driving would be a good first in a battery of

CONTACT H. E. C. van der Wall  hvdwall@chdr.nl  Centre for Human Drug Research, Zernikedreef 8, 2333 CL, Leiden, the Netherlands

 Supplemental data for this article is available online at <https://doi.org/10.1080/15389588.2021.1914837>.

Associate Editor Jessica S. Jermakian oversaw the review of this article.

© 2021 Taylor & Francis Group, LLC

tests that can evaluate the effect of new drugs on driving behavior, as sleep deprivation can serve as a surrogate of sedation caused by sedative drug effects (Van Steveninck et al. 1999).

In the current study, we attempted to create a model to evaluate the effect of sleep deprivation on driving behavior as sleepiness is also known to affect driving behavior (Gaspar et al. 2017; Koopmans et al. 2020; Schwarz et al. 2019; Soares et al. 2020). Although drowsy drivers are as dangerous as drivers with unlawful blood alcohol levels they cannot be caught in a police checkpoint, but only in case of a perceived dangerous driving situation (Haraldsson and Akerstedt 2001). Such a model, when sufficiently accurate could be used to detect drug or food induced sleepiness, allowing either dose adjustment or adequate warning notes.

The aim of the current study was twofold: 1) to develop a new model allowing to characterize sleep-deprived driving behavior, and 2) to demonstrate how driving behavior after intake of alcohol or alprazolam is similar to sleep-deprived driving behavior, in order to validate the use of the model for characterization of a new drug.

Methods and materials

Data collection for the model

All data used in the present study were collected during two previous studies (Huizinga et al. 2019; Koopmans et al. 2020). In both studies subjects were healthy adults who were in possession of a valid drivers license. They were active and skilled drivers with a minimum mileage of 3,000 km per year. Subjects were instructed to drive in a driving simulator (Green Dino BV, Wageningen, The Netherlands) with a steady lateral position in the right-hand lane of a 30 min dual-carriageway highway scenario similar to the one being used during on-road tests; overtaking other vehicles was allowed. The simulators have a nonmoving base and consist of a mockup car with three pedals (clutch, brake and gas), manual shift, steering wheel, safety belt, indicators and hand brake. The controls are linked to a dedicated graphics computer that simulates road environment and dynamic traffic. The driving simulators have a wide view display, made with three LCD (24") flat panel monitors positioned side by side. The total LCD monitor surface is 0.48 m².

Data used to create the model that allows the characterization of sleep-deprived driving behavior were collected during a previous study (Koopmans et al. 2020). In short, this was an exploratory single-center cross-over study in 24 healthy male subjects, 23 to 35 years of age, to investigate the effects of sleep deprivation on driving.

Subsequently, this model was used to demonstrate sleep-deprived driving characteristics in subjects after intake of alprazolam or alcohol. The effect of alprazolam and alcohol on driving was previously studied in our institution by Huizinga et al. (2019). (Huizinga et al. 2019) In short, this was a single-center, randomized, double-blind, double-dummy, placebo-controlled, four-way crossover-study on alcohol and alprazolam in 24 healthy subjects (12 males, 12 females, age range 20–43 years), while performing

neurocognitive and psychomotor tests on the NeuroCart® – a comprehensive battery that can test all functional domains of the central nervous system (CNS) – and a driving simulator (Green Dino BV, Wageningen, The Netherlands). The interventions consisted of intravenously administered alcohol to obtain steady state concentrations of 0.5 g L⁻¹ and 1.0 g L⁻¹, oral administration of 1 g alprazolam, or placebo (Zoethout et al. 2012). Driving and laboratory tests were performed at regular time intervals during a study day. As the pharmacodynamic effects for alcohol and alprazolam varied during one single study period, measurements at 2- and 4-h post-dose were used for the characterization of alprazolam. Measurements at 5 and 6 h post dose were used for the characterization of alcohol. All used parameters are listed in Table A1.

Feature pre-processing

All measurements were corrected for baseline, by subtracting the mean of all baseline values of all treatment arms in the alcohol and alprazolam datasets, of the subject from the values after drug ingestion.

Due to the nature of the study the sleep-deprivation dataset had no baseline measurement shortly before the intervention. Therefore, the morning measurement in the well-rested occasion was used as baseline. The feature values of this measurement were subtracted from the sleep-deprived measurement and the afternoon well-rested measurement (which was used as control).

The features required to develop a model were created in a similar way as described in the article by van der Wall et al. (2020). In short, the mean, the standard deviation, and the mean absolute difference between consecutive time points were calculated for all parameters. In addition, the minima and maxima for the speed, steer-speed and distance to the car in front were calculated. Additionally, the maximum intensity of the power spectrum of a Fourier transform at low frequency (<0.05 Hz) and high frequency (>0.05 Hz) were calculated for all parameters.

Finally, some of the features were obtained from the original study, which were calculated after cleaning the data (including removal of lane switches): the standard deviation of the lateral position (SDLP), the mean lateral position (MLP), the mean speed (MS) and the standard deviation of speed (SDS). A list of all features is shown in Table A2.

Feature selection

When two features had a high correlation (>0.9 or < -0.9), only the most important one – based upon the feature importance of fitting the model on the training set was used for final validation (van der Wall et al. 2020).

Machine learning

In our previous study two linear and two non-linear models were tested on driving simulator data after intake of alcohol or alprazolam. The linear models showed accuracies of 67%

and 54% for the alprazolam training set (logistic regression and Support Vector Machine, respectively), and 60% and 52% for the alcohol training set. The non-linear models gave the best performances. Random forest and gradient boosting models both showed an accuracy 81% for the alprazolam training set, and 65% and 68% for the alcohol training set, respectively (van der Wall et al. 2020). Since it was shown to work best overall on this type of driving simulator data, a gradient boosting classifier with a subsample rate of 0.5 was used as a model, which was obtained from the scikit learn module version 0.23.1 in python 3.7.3.

A cross validation was performed to evaluate the performance of the algorithm. The features of 80% of randomly selected subjects were used five times to train the model. Subsequently, the model was tested on the other 20%. The features were standardized based on the data in the training set. The model performance was evaluated by assessing accuracy, specificity, sensitivity, positive predictive values (PPV), and negative predictive values (NPV) (Wong and Lim 2011). Data were presented as mean \pm SD. The performance when using all driving features was compared with the performance of the model when using SDLP only.

To demonstrate sleep-deprived driving characteristics in subjects after intake of alprazolam or alcohol, the model was trained on the entire sleep deprivation data set (all measurements and all features) and tested on the alprazolam and alcohol data sets.

Since a continuous “sleep-deprived” score was preferred to get an indication of how similar the effects of the drug was to the sleep deprivation effects, a simple (binary) prediction from a classifier did not seem useful. Therefore, the measurements in the alcohol and alprazolam datasets were given probability predictions.

Probability scores indicate the likelihood that the outcome of the model is positive for a particular intervention, in this case sleep-deprivation. Therefore, these values may provide information about the similarity to the sleep-deprived effects. A higher score means a higher probability of being sleep-deprived and therefore the driving pattern is more similar to a sleep-deprived driving pattern. In the current study, probability scores can be used to demonstrate the similarity between the effects of sleep deprivation and alcohol/alprazolam intake on driving behavior. A high probability score means that a test subject using these compounds acting on the central nervous system shows driving characteristics very similar to those of sleep deprived subjects.

Finally, the model was also fitted on the alprazolam and alcohol sets to collect the feature importance for detection of these interventions, whereupon it was compared with the feature importance of the sleep-deprivation model. The importance of the features derived from the same parameter was added together because of the high correlations between these features.

Statistical analysis

Treatment effects on the probability score were analyzed with a mixed model analysis of variance (ANOVA) with

treatment, measurement, occasion in the cross-over experiment and treatment by measurement as fixed factors and subject, subject by treatment and subject by measurement as random factors. This was similar to the statistical analysis of the SDLP in the study of Huizinga et al. For this analysis, R version 3.6.1 was used. The difference was considered significant if the p -value was below 0.05.

Results

From the sleep-deprived driving dataset in 24 subjects, one subject did not complete the “baseline” measurement and therefore the data of this subject were not taken into consideration. Additionally, one subject did not finish the control measurement. Therefore, 45 baseline corrected measurements, consisting of 23 sleep-deprived and 22 controls, were used for our analyses.

Model performance

The model had an accuracy of $77 \pm 9\%$, a specificity of $77 \pm 23\%$, and a sensitivity of $76 \pm 9\%$. The PPV and NPV of the model were $83 \pm 17\%$ and $74 \pm 5\%$. [Supplementary Figure 1](#) shows the average feature importance of the repeated model fitting. The most important feature for predicting whether a subject was driving sleep-deprived was the SDLP. The maximum steering speed was only of minor importance for predicting sleep-deprived driving behavior.

When using only SDLP the accuracy of the model was $70 \pm 10\%$, the sensitivity $80 \pm 14\%$ and the specificity $58 \pm 13\%$. In this case the PPV and NPV were $69 \pm 7\%$ and $74 \pm 18\%$.

Characterization of alprazolam or alcohol affected driving behavior

For the alcohol and alprazolam data, the same measurements were used as in our previous study in which the models for detection of abnormal driving was introduced (van der Wall et al. 2020). This concerned 80 placebo measurements (40 to compare with alprazolam and 40 to compare with alcohol), 44 measurements after alprazolam intake and 36 measurements after 1 g/l alcohol intake. Additionally, in the current study also 42 measurements after 0.5 g/l alcohol were used to compare with placebo on sleep-deprived driving characteristics.

In [Figure 1](#) violin plot is shown containing the probability scores of the model on the 1.0 g/l alcohol dataset. It demonstrates the similarity between the effects of sleep deprivation and alcohol intake on driving behavior. The mean probability score for the measurements after 1.0 g/l alcohol intake was 0.63 ± 0.37 , substantially higher than the score of 0.30 ± 0.36 for the measurements after placebo intake. This finding indicates a substantial similarity in the driving characteristics after sleep deprivation and those after 1.0 g/L alcohol intake. Statistical analysis of the scores revealed that there was a significant effect of treatment ($p = 0.0014$).

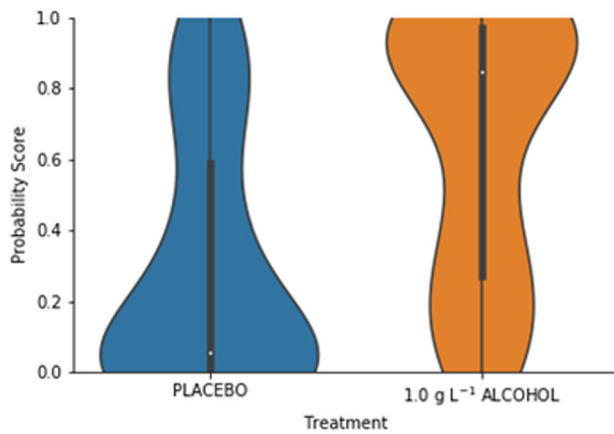


Figure 1. Sleep deprived probability score for placebo and 1.0 g/L alcohol. Violin plot of the probability scores of the measurements in the 1.0 g/L alcohol data set, indicating the distribution of the probability scores. The left violin plot shows the scores for the placebo measurements. The right violin plot shows the scores for the 1.0 g/L alcohol measurements. The width of the violin reflects the relative number of measurements with that score.

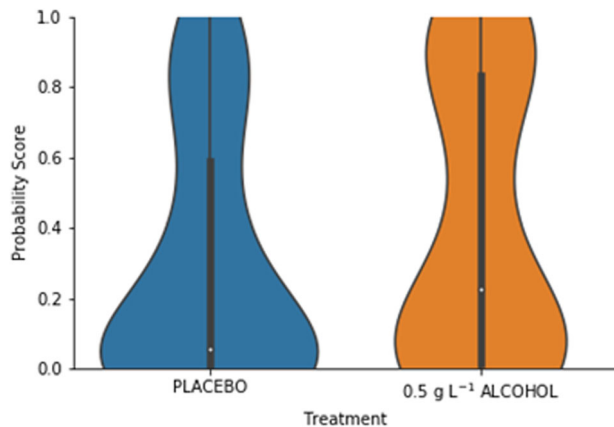


Figure 2. Sleep deprived probability score for placebo and 0.5 g/L Alcohol. Violin plot of the probability scores of the measurements in the 0.5 g/L alcohol data set, indicating the distribution of the probability scores. The left violin plot shows the scores for the placebo measurements. The right violin plot shows the scores for the 0.5 g/L alcohol measurements. The width of the violin reflects the relative number of measurements with that score.

The violin plot in Figure 2 shows the probability scores of the model on the 0.5 g/l alcohol dataset. The mean probability score for the measurements after 0.5 g/l alcohol intake was 0.41 ± 0.40 , compared to 0.30 ± 0.36 of the placebo measurements. The differences in these scores were insignificant ($p = 0.2434$).

The probability scores on the alprazolam data set are shown in the violin plot of Figure 3. The mean probability score for the measurements after alprazolam intake was as high as 0.79 ± 0.32 , versus 0.27 ± 0.32 for the placebo measurements. So, after alprazolam intake subjects showed car-driving characteristics which were quite similar to sleep-deprived driving characteristics. As expected, the ANOVA yielded a significant effect of treatment ($p < 0.0001$).

In Figure 4, the sum of all feature importances for all parameters is shown for the alprazolam, 1.0 g/l alcohol and sleep-deprived dataset. For all datasets, the sum of the importances of the features derived from the lane position

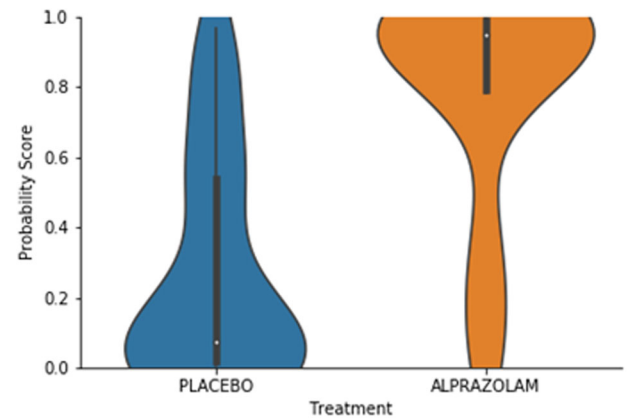


Figure 3. Sleep deprived probability score for placebo and alprazolam. Violin plot of the probability scores of the measurements in the alprazolam data set, indicating the distribution of the probability scores. The left violin plot shows the scores for the placebo measurements. The right violin plot shows the scores for the alprazolam measurements. The width of the violin reflects the relative number of measurements with that score.

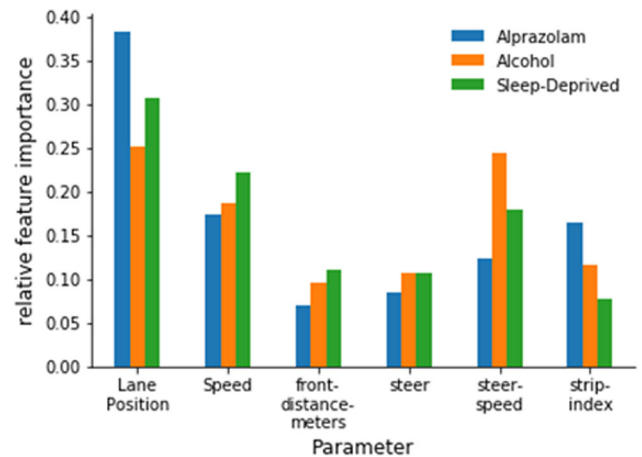


Figure 4. Relative feature importance. Sum of the relative feature importances grouped by type for all parameters (sum of all feature importances is 1). An overview of all features is shown in Table A2.

was the highest. For the 1.0 g/l alcohol set, the steer-speed features were also of high importance.

Discussion

Sleep deprivation is known to impair driving performance (Gaspar et al. 2017; Peters et al. 1999; Philip et al. 2005; Schwarz et al. 2019; Soares et al. 2020). The current study has shown that, by using machine learning, the effect of sleep deprivation on driving behavior can be classified. The created model performed with an average accuracy of 77%. The current study did also show that driving behavior after sleep deprivation had a great similarity with driving behavior after alprazolam intake and to a lesser extent after alcohol intake of 1 g/l. For the first time, we developed a model to characterize abnormal driving behavior for a single intervention, which can also be used for characterization of other/new interventions. Using this model, the effects of a newly developed drug on driving behavior can be compared with the effects of sleep deprivation. In this way, a series of

these machine learning models could evolve to a test battery, which allows a more precise and accurate evaluation of abnormal driving behavior by creating a predictive effect profile for a medicine.

The performance of the created model is similar to what we have shown for the classification of alcohol and alprazolam in our previous study (van der Wall et al. 2020). These earlier models, which performed with an accuracy of around 80%, could characterize abnormal driving behavior solely for the drug that had been used for the development of that particular model. In the present study we have shown that such a model may serve to characterize driving behavior after a diversity of interventions. Given the large number of influencing factors that can affect driving behavior the accuracy of 77% of our model seems satisfying. Larger number of subjects may improve future performance of the model.

When using all driving features the performance of the model is, although not significantly, higher than when using only SDLP. This is in line with the results of previous study (van der Wall et al. 2020).

Also, when using SDLP only, the specificity is much lower and thereby susceptible to false positive results.

The SDLP remains the most important feature for distinguishing abnormal driving behavior and has rightfully been used as a standard measure. However, by combining all driving features in a model the way in which driving behavior deviates can be determined, so that a more informative assessment can be made.

The probability scores for the placebo measurements in the alprazolam dataset (3 and 4 h after intake) and alcohol datasets (5 and 6 h after intake) were very similar. This means that the average control measurement would get a probability score somewhat below 0.3. An average sleep-deprived score of 0.3 can be considered high for a subject driving under “normal” circumstances but considering the relatively small training set this score is reasonable. When more data will become available, the recognition of control measurements will improve and the probability scores for these measurements will decrease.

The current study has also shown that driving behavior after sleep deprivation shows characteristics, which are quite like those after alprazolam intake and, but to a lesser extent, after alcohol intake of 1 g/l. An explanation for this can be found in the analysis of feature importance. The features derived from the lane position, which also includes the SDLP, are most important. However, for assessing the effects of alcohol, the features calculated from the steer-speed parameter are also of high importance, while these are of minor importance in the detection of alprazolam and sleep deprivation.

As alprazolam is a sedative, great similarities with drowsiness were expected, which was confirmed in our study by the highest sleep deprived probability scores after intake of this drug. After 1.0 g/L alcohol intake also a substantial similarity of the abnormal driving characteristics was found, but to a lesser extent. After 0.5 g/l alcohol intake, which is just under the legal limit, there was almost no sleep-deprived

effect in driving behavior. The probability scores seem to match the relative drowsiness expected from the interventions, indicating that the model shows the degree of sleep deprivation in other interventions.

Limitations

Although PPV and NPV of the model were quite high (around 75%), these predictive values were still suboptimal, possibly due to the relatively small number of subjects. This information can be derived from the violin plots, where still a great variation of scores can be observed. Therefore, more subjects are needed to be tested in order to obtain a reliable characterization of driving behavior of a drug, and to neutralize the errors of the model. Part of the variation can also be explained by inter-subject differences. The effect of an intervention may vary substantially for each subject and the number of subjects in the datasets is relatively small. Moreover, the change in driving performance may also differ for subjects in a training set compared to subjects in a test set. Currently, the training set only contains male subjects. As we don't know the difference of the effect of sleep-deprivation on driving behavior between males and females, we cannot say anything about the effect of sleep-deprivation in female subjects. However, female subjects have been involved in several previous car driving studies (Åkerstedt et al. 2010; Banks et al. 2004).

It must be kept in mind that baseline measurements are still required for an accurate evaluation. In this way, the inter-subject variation may be reduced by correction for the baseline variation in driving style. The accuracy might improve when individual normal driving behavior could be learned based on multiple control measurements in one subject.

It is difficult to estimate how dangerous driving behavior is with sleep-deprived driving characteristics. Because sleep deprivation may lead to abnormal driving and may cause accidents, it seems reasonable to assume that with a higher score on sleep-deprived driving, driving performance is more impaired. On the other hand, a small overlap with driving after sleep deprivation does not mean that driving is safe. Driving behavior might be negatively influenced in a different way. In the current study the alcohol intake well over the legal dose (1.0 g L^{-1}) appears to have less impact on the driving performance than alprazolam, while previous studies have shown that alcohol greatly impairs driving behavior (Arnedt et al. 2000; Bunn et al. 2019; Huizinga et al. 2019; Irwin et al. 2017; Mets et al. 2011). In our previous study we have shown that driving behavior after alcohol intake could be assessed with an accuracy of more than 80% (van der Wall et al. 2020). Therefore, a test battery with multiple models is needed to give a good indication of the way driving is affected by drugs. This test battery consists of a set of models, characterizing the effect of CNS-active medicines on driving behavior.

The order of sleep deprivation for the interventions tested in the current study was as could have been anticipated (alprazolam highest, then 1.0 g/l alcohol, then 0.5 g/l alcohol,

then placebo). Therefore, the model developed in the current study can be used to identify how interventions such as the intake of drugs may influence the driving performance in a sleep-deprived way. Therefore, this can serve as a benchmark in a test battery to characterize how drugs affect driving performance. In this way psychopharmacological drugs could be tested for effects on driving behavior in an early stage of development.

Acknowledgments

The authors would like to thank Dr Karen Broekhuizen, who provided medical writing services on behalf of the Centre for Human Drug Research.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Åkerstedt T, Ingre M, Kecklund G, Anund A, Sandberg D, Wahde M, Philip P, Kronberg P. 2010. Reaction of sleepiness indicators to partial sleep deprivation, time of day and time on task in a driving simulator—the drowsi project. *J Sleep Res.* 19(2):298–309.
- Arnedt JT, Wilde G, Munt PW, MacLean AW. 2000. Simulated driving performance following prolonged wakefulness and alcohol consumption: separate and combined contributions to impairment. *J Sleep Res.* 9(3):233–241.
- Banks S, Catchside P, Lack L, Grunstein RR, McEvoy RD. 2004. Low levels of alcohol impair driving simulator performance and reduce perception of crash risk in partially sleep deprived subjects. *Sleep.* 27(6):1063–1067.
- Bunn T, Singleton M, Chen I-C. 2019. Use of multiple data sources to identify specific drugs and other factors associated with drug and alcohol screening of fatally injured motor vehicle drivers. *Accid Anal Prev.* 122:287–294.
- Darby P, Murray W, Raeside R. 2009. Applying online fleet driver assessment to help identify, target and reduce occupational road safety risks. *Saf Sci.* 47(3):436–442.
- Gaspar JG, Brown TL, Schwarz CW, Lee JD, Kang J, Higgins JS. 2017. Evaluating driver drowsiness countermeasures. *Traffic Inj Prev.* 18(sup1):S58–S63.
- Haraldsson P, Åkerstedt T. 2001. Drowsiness—greater traffic hazard than alcohol. Causes, risks and treatment. *Lakartidningen.* 98(25): 3018–3023.
- Houwing S, Mathijssen R, Brookhuis K. 2012. In search of a standard for assessing the crash risk of driving under the influence of drugs other than alcohol; results of a questionnaire survey among researchers. *Traffic Inj Prev.* 13(6):554–565.
- Huizinga CR, Zuiker RG, de Kam ML, Ziagos D, Kuipers J, Mejia Y, van Gerven JM, Cohen AF. 2019. Evaluation of simulated driving in comparison to laboratory-based tests to assess the pharmacodynamics of alprazolam and alcohol. *J Psychopharmacol.* 33(7):791–800.
- Irwin C, Iudakhina E, Desbrow B, McCartney D. 2017. Effects of acute alcohol consumption on measures of simulated driving: a systematic review and meta-analysis. *Accid Anal Prev.* 102:248–266.
- Irwin C, Monement S, Desbrow B. 2015. The influence of drinking, texting, and eating on simulated driving performance. *Traffic Inj Prev.* 16(2):116–123.
- Koopmans I, Heima H, Doll RJ, van der Wall HEC, de Kam ML, Groeneveld GJ, Cohen AF, Zuiker RG. 2020. Sensitivity and validity of on-the-road and simulated driving test to measure impaired driving behaviour: effect of sleep deprivation. *Sleep Med.* (in preparation).
- Mets MA, Kuipers E, de Senerpont Domis LM, Leenders M, Olivier B, Verster JC. 2011. Effects of alcohol on highway driving in the stimis driving simulator. *Hum Psychopharmacol Clin Exp.* 26(6):434–439.
- Peters RD, Wagner E, Alicandri E, Fox JE, Thomas ML, Thorne DR, Sing HC, Balwinski SM. 1999. Effects of partial and total sleep deprivation on driving performance. *Public Roads.* 62(4):2–6.
- Philip P, Sagaspe P, Moore N, Taillard J, Charles A, Guilleminault C, Bioulac B. 2005. Fatigue, sleep restriction and driving performance. *Accid Anal Prev.* 37(3):473–478.
- Robertson RD, Hing MM, Pashley CR, Brown SW, Vanlaar WG. 2017. Prevalence and trends of drugged driving in Canada. *Accid Anal Prev.* 99:236–241.
- Sagberg F, Selpi B, Piccinini GF, Engström J. 2015. A review of research on driving styles and road safety. *Hum Factors.* 57(7): 1248–1275.
- Schwarz C, Gaspar J, Miller T, Yousefian R. 2019. The detection of drowsiness using a driver monitoring system. *Traffic Inj Prev.* 20(sup1):S157–S161.
- Soares S, Ferreira S, Couto A. 2020. Driving simulator experiments to study drowsiness: a systematic review. *Traffic Inj Prev.* 21(1):29–37.
- van der Wall H, Doll R, van Westen G, Koopmans I, Zuiker R, Burggraaf J, Cohen A. 2020. The use of machine learning improves the assessment of drug-induced driving behaviour. *Accid Anal Prev.* 148(105822).
- Van Steveninck A, Van Berckel B, Schoemaker R, Breimer D, Van Gerven J, Cohen A. 1999. The sensitivity of pharmacodynamic tests for the central nervous system effects of drugs on the effects of sleep deprivation. *J Psychopharmacol.* 13(1):10–17.
- Verster JC, Roth T. 2011. Standard operation procedures for conducting the on-the-road driving test, and measurement of the standard deviation of lateral position (SDLP). *Int J Gen Med.* 4(359):359–371.
- Wong HB, Lim GH. 2011. Measures of diagnostic accuracy: sensitivity, specificity, PPV and NPV. *Proc Singapore Healthcare.* 20(4):316–318.
- Zoethout RW, de Kam ML, Dahan A, Cohen AF, van Gerven JM. 2012. A comparison of the central nervous system effects of alcohol at pseudo-steady state in Caucasian and expatriate Japanese healthy male volunteers. *Alcohol.* 46(7):657–664.