The use of machine learning improves the assessment of drug-induced driving behaviour

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

Citation

The use of machine learning improves the assessment of drug-induced driving behaviour

H.E.C. van der Wall [a,b,*], R.J. Doll [a], G.J.P. van Westen [b], I. Koopmans [a], R.G. Zuiker [a], J. Burggraaf [a,b,c], A.F. Cohen [a,b,c]

[a] *Centre for Human Drug Research, Leiden, the Netherlands*
[b] *Leiden Academic Centre for Drug Research, Leiden, the Netherlands*
[c] *Leiden University Medical Centre, Leiden, the Netherlands*

ARTICLE INFO	ABSTRACT
	Rationale: Car-driving performance is negatively affected by the intake of alcohol, tranquillizers, sedatives and sleep deprivation. Although several studies have shown that the standard deviation of the lateral position on the road (SDLP) is sensitive to drug-induced changes in simulated and real driving performance tests, this parameter alone might not fully assess and quantify deviant or unsafe driving. *Objective:* Using machine learning we investigated if including multiple simulator-derived parameters, rather than the SDLP alone would provide a more accurate assessment of the effect of substances affecting driving performance. We specifically analysed the effects of alcohol and alprazolam. *Methods:* The data used in the present study were collected during a previous study on driving effects of alcohol and alprazolam in 24 healthy subjects (12 M, 12 F, mean age 26 years, range 20–43 years). Various driving features, such as speed and steering variations, were quantified and the influence of administration of alcohol or alprazolam was assessed to assist in designing a predictive model for abnormal driving behaviour. *Results:* Adding additional features besides the SDLP increased the model performance for prediction of drug-induced abnormal driving behaviour (from an accuracy of 65 %–83 % after alprazolam intake and from 50 % to 76 % after alcohol ingestion). Driving behaviour influenced by alcohol and alprazolam was characterised by different feature importance, indicating that the two interventions influenced driving behaviour in a different way. *Conclusion:* Machine learning using multiple driving features in addition to the state-of-the-art SDLP improves the assessment of drug-induced abnormal driving behaviour. The created models may facilitate quantitative description of abnormal driving behaviour in the development and application of psychopharmacological medicines. Our models require further validation using similar and unknown interventions.

1. Introduction

Road safety research indicates that car-drivers differ in the risk to become involved in a crash and that this depends on numerous factors. (Sagberg et al., 2015; Shinar, 2017) Besides the predisposing behaviour (e.g., driving style) and predisposing characteristics (e.g., age) of the car-driver, intake of alcohol, psychoactive medicines and recreational drugs also negatively affect car-driving behaviour.(Arnedt et al., 2000; Mets et al., 2011; Robertson et al., 2017; Vanlaar, 2005; Vanlaar and Robertson, 2010).

Drive simulators provide a good opportunity to study car-driving and

driving safety associated with drug effects. However, the assessment of deviant driving behaviour remains challenging. Many researchers use the Standard Deviation of the Lateral Position of the car on the road (SDLP) as a measure to quantify driving safety. (Liguori, 2009; Verster and Roth, 2011) Although several studies have shown that the SDLP is sensitive to drug-induced changes in driving behaviour (Darby et al., 2009; Guo and Fang, 2013; Mets et al., 2011; Verster and Roth, 2011), it is uncertain that it is able to distinguish between numerous different aspects of driving. To date, it is unclear whether a combination of several features such as the mean lateral position (MLP), mean speed (MS), and the standard deviation of speed (SD-Speed), amongst many

other features, could improve the assessment of driving behaviour and safety.

In recent years, machine learning models have aided in detecting predictive factors in various fields of behavior and engineering. (Deo, 2015; Hegde and Rokseth, 2020; Obermeyer and Ezekiel, 2016; Paltrinieri et al., 2019) Tango and Botta (2013) for example, proposed a number of factors classifying driver distractions in relation to driving safety. (Tango and Botta, 2013) Machine learning models may not only detect deviant driving behaviour but may also explain how driving behaviour is affected by the intake of drugs. To date, there is no evidence on the association between the intake of drugs and combined driving parameters. Such a model could improve the early recognition of how new drugs could affect driving behaviour.

Current automotive and simulator technology allows the extraction of multiple features other than lateral movement. The question is if these extra features could lead to a better detection of drug-induced effects on driving. While many different driving studies have been conducted in recent decades, including the testing of drugs, cognitive disorders and other diseases on the road and in a driving simulator (Brunnauer et al., 2016, 2009; Grabe et al., 1998; Jauhar et al., 1993; Leufkens et al., 2014a, 2014b; Moore, 1977; O'Hanlon et al., 1995; Soyka et al., 2005; Van Laar et al., 1992; Wylie et al., 1993), none of these sufficiently validated other evaluation indices in addition to SDLP. (Iwata et al., 2018)

In this study, we applied Machine Learning methods to build models describing which parameters can best distinguish driving while under the influence of different substances (i.e., alprazolam and alcohol). We investigated if including multiple simulator-derived parameters rather than the SDLP alone would provide a more accurate assessment of the effect of substances affecting driving performance. We specifically analysed the effects of alcohol and alprazolam on car driving behaviour as these effects have shown to have the highest frequencies among fatally injured drivers. (Bunn and Chen, 2019)

2. Material & methods

2.1. Data collection

Data used in this study were collected during a previous study, which was approved by the Independent Ethics Committee of 'Foundation Evaluation of Ethics in Biomedical Research BEBO'. (Huizinga et al., 2019) In short, this was a single-centre, randomized, double-blind, double-dummy, placebo-controlled, four-way crossover-study with 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® and a driving simulator (Green Dino BV, Wageningen, The Netherlands). Subjects were instructed to drive 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. Subjects were instructed to maintain a steady speed with a maximum speed of 100 km/h; overtaking other vehicles was allowed. The interventions consisted of intravenously administered alcohol using a protocol (Zoethout, 2012) to obtain continuous concentrations of 0.5 g L^{-1} and 1.0 g L^{-1}, and alprazolam which was given orally in a dose of 1 mg. Driving tests and laboratory tests were done at regular time intervals during a study day. In the current analysis the driving parameters from the study days with 1.0 g L^{-1} alcohol, alprazolam and placebo were considered. As the pharmacodynamical effects for alcohol and alprazolam varied during one single occasion, the measurements were selected based on the highest pharmacodynamic effects of both alprazolam and alcohol. The largest effect of alprazolam on driving was observed between 2 and 4 h after drug intake. The largest response of alcohol was observed between 5 and 6 h. We used 2 measurements in the period of the largest effects instead of only the one with the highest effect to optimally train the models.

Data from all 24 subjects were used for our analysis when available.

The data set comprised a total of 80 test drives from 20 study days with placebo treatment; 40 of these placebo tests were used to create and validate the model for alprazolam and the other 40 test drives for optimization and validation of the alcohol model. The effect of alprazolam was assessed using 44 test drives from 22 study days and for the evaluation of the effect of alcohol 35 test drives from 18 study days were available. We ensured that the external test set only contained subjects who joined both the drug administration days and the placebo study days.

All used parameters are listed in Table 1.

2.2. Feature preprocessing

To create a model, features for every observation were required. The first 5 and the last 10 min of each measurement were removed from the dataset which left 15 min (from 5−20 min) of driving data per measurement. Contrary to the previous analyses as reported by Huizinga et al. (2019), lane switches were included in the dataset. For each parameter time series, the following features were calculated:

- Mean: mean of the whole time series
- Std: standard deviation of the whole time series
- Diff: average absolute difference between successive time points in a time series
- Intensity: highest intensity of the power spectrum of the Fourier transform (sampling frequency of 10 Hz) of the time series corrected with the mean value of the time series
- Frequency: frequency with the highest intensity of the power spectrum of a Fourier transform (sampling frequency of 10 Hz) of the time series corrected with the mean value of the time series

For the speed, steer-speed and front-distance-meters, also the following was calculated:

- Min: minimum of the time series
- Max: maximum of the time series

In addition, the following features obtained from the original study of Huizinga et al. (2019) - after cleaning the data (including removal of lane switches) - were used: the standard deviation of the lateral position (SDLP), the mean lateral position (MLP), the mean speed (MS), the standard deviation of speed (SDS). A list of all features is shown in Table 2.

To obtain baseline corrected values, the mean of all baseline values (of all treatment arms) of the subject was subtracted from the values after drug ingestion.

2.3. 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 our final validation.

Table 1
List of used driving parameters with their descriptions.

Parameters	Description
Strip-index	lateral position on the entire highway
Lane Position	lateral position in the lane
Speed	speed
Steer	steer-position
Steer-speed	speed of steering to the right
Front-distance-meters	distance to the car in front in meters

Table 2
Overview of all features.

Feature Name	mean	std	diff	intensity	freq (f)	min	max
Strip-index	x	x	x	x	x		
Lane Position	x	x	x	x	x		
Speed	x	x	x	x	x	x	x
Steer	x	x	x	x	x		
Steer-speed	x	x	x	x	x	x	x
Front-distance-meters	x	x	x	x	x	x	x
Features from original study:							
GD_SDLP2							
GD_lane_mean							
GD_spd_mean							
GD_SDSpeed							

2.4. Machine learning

Pipeline Pilot 2018 was used for all analyses and calculations performed in the present study. (BIOVIA, 2018) All machine learning was performed using scikit-learn version 0.21.2 in python 3.6.7.

We aimed for a uniform model that could be applied for a variety of pharmaceutical drugs.

Data sets were randomly split into a training set, consisting of 80 % of the subjects, and a test set containing the other 20 %. Prior to modelling, the features were normalized based on the data in the training set.

Accuracy scores of internal cross-validation on the training sets both with all features and using SDLP only were used in order to select a machine learning model. We compared two linear and two non-linear models. The linear models we used were: 1) Logistic regression and 2) linear support vector machines. Logistic regression is a widely used multivariable method for modeling dichotomous outcomes, which converts linear regression to a binary classifier with sigmoid function. (Bagley et al., 2001) Support Vector Machines separate classes using hyperplanes that split the classes, using a flat plane, within the predictor space. (Subasi and Gursoy, 2010) The two non-linear models we used were: 1) Random Forest, and 2) Gradient Boosting. Random Forest models have been reported as excellent classifiers with the following advantages: simple theory, fast speed, stable and insensitive to noise, little or no overfitting, and automatic compensation mechanism on biased sample numbers of groups. (T. Chen et al., 2013) The gradient boosting model is an ensemble of decision trees - which categorizes data by setting up decision rules - with boosting algorithm, and has been successfully used to predict cardiovascular events, development of sepsis and delirium. (Du et al., 2020; Zhang et al., 2019) In the current study we used a subsample rate of 0.5 to mitigate overfitting.

2.5. Final validation

For the final validation of the model, it was first studied whether the administration of alprazolam or alcohol could be distinguished from placebo treatment using only the SDLP obtained from the original analysis. Next, it was studied if this could be performed and improved upon using all features.

The training and testing of the model were repeated five times for both the data set with only SDLP and the full data set with all features. Model performance was evaluated by assessing accuracy, specificity, sensitivity, positive predictive values (PPV), negative predictive values (NPV), and area under the Receiver Operating Characteric (ROC) curves. Data were presented as mean ± SD. Finally, the probability/continuous scores of the predictions – ranging from 0 (placebo) to 1 (intervention) - were extracted to show how the models could be used for distinguishing abnormal from normal driving behaviour.

3. Results

First, the average accuracies of the internal cross validation for the

suggested models are presented. Based on this outcome, a model is chosen for the final validation. Subsequently, the performance of the model using both SDLP only and all parameters for the discrimination of Alprazolam and alcohol consecutively is presented.

3.1. Model selection

For both the alcohol and alprazolam training set with all features, the non-linear models gave the best performances. Random Forest and Gradient boosting models both showed an accuracy of 81 % for the alprazolam training set, and 65 % and 68 % for the alcohol training set, respectively. The linear models showed lower accuracy with values of 67 % and 54 % for the alprazolam training set (logistic regression and SVM, respectively), and 60 % and 52 % for the alcohol training set. (Table 3) Most models showed lower performance on the SDLP-only training set. Since the gradient boosting model scored overall best for the internal validation of the training set, this model was used for final validation.

3.2. Alprazolam

Fig. 1 shows the accuracy, specificity, and sensitivity for alprazolam usage versus placebo of SDLP alone (black bars), and of the models using all driving features on predicting alprazolam ingestion (grey bars). These driving features have been listed in Table 2. Fig. 1 clearly shows that the addition of other driving features considerably improved the prediction model compared to the performance of the model using SDLP only. The accuracy improved from $65 \pm 0\%$–$83 \pm 4\%$, the specificity from $50 \pm 0\%$–$82 \pm 7\%$, and the sensitivity from $80 \pm 0\%$–$83 \pm 6\%$. For the models using all features, the PPV and NPV were $83 \pm 6\%$ and $84 \pm 4\%$, respectively, versus $62 \pm 0\%$ and $71 \% \pm 0\%$, respectively, for the model using SDLP only. Supplementary Figs. 1 and 2 show the ROC curves. The area under the curve improved from $77 \pm 1\%$ for the models using SDLP only to $91 \pm 3\%$ for the models using all features.

Fig. 2 shows the average feature importance of the models based on all features included in our analyses. The most important feature for predicting whether a subject had used alprazolam was the SDLP, which represents the standard deviation in lateral position after removal of lane switches. By contrast, the maximal speed was only of minor importance in predicting the usage of alprazolam.

In Fig. 3, boxplots are shown containing the continuous (probability) predictions of one of the repetitions for both alprazolam models. The differences in prediction score between alprazolam and placebo is significantly larger when using multiple features.

3.3. Alcohol

Fig. 4 shows the accuracy, specificity, and sensitivity for alcohol usage versus placebo of both SDLP alone (black bars) and the models using all driving features on predicting alcohol intake (grey bars). In terms of performance, the accuracy improved from $50 \pm 0\%$–$76 \pm 4\%$, the specificity from $60 \pm 0\%$–$82 \pm 7\%$, and the sensitivity improved

Table 3
The internal cross-validation accuracy scores on all training sets for the suggested models.

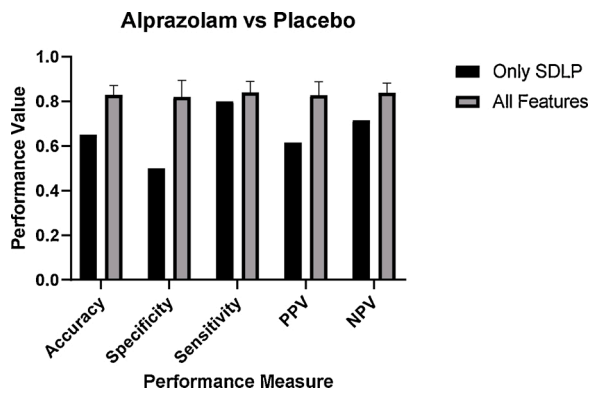| Model	Alprazolam all features	Alprazolam SDLP-only	Alcohol all features	Alcohol SDLP-only
Logistic Regression	67 %	61 %	60 %	69 %
Support Vector Machine	54 %	58 %	52 %	65 %
Random Forest	81 %	58 %	65 %	61 %
Gradient Boosting	81 %	58 %	68 %	65 %

Fig. 1. Performances with standard deviation of the models using only the standard deviation of the lateral position (SDLP, black bars) and the models using all features on predicting ingestion of alprazolam (grey bars). PPV, positive predictive values. NPV, negative predictive values.

from $40 \pm 0\%$–$70 \pm 0\%$. For the models using all features, the PPV and NPV were $80 \pm 7\%$ and $73 \pm 2\%$, respectively, versus $50 \pm 0\%$ and $50 \pm 0\%$, respectively, for the model using SDLP only. Supplementary Figs. 3 and 4 show the ROC curves. The area under the curve improved from $57 \pm 4\%$ for the models using SDLP only to $82 \pm 2\%$ for the models using all features. Similar to the results observed with alprazolam, the addition of driving features substantially improved the performance of the model predicting alcohol ingestion. In Fig. 5 the relevance of the various

features that were used in the analyses on alcohol intake is shown. The most important feature for predicting the presence of alcohol was - like the results for alprazolam prediction - the SDLP. Conversely, the mean speed (after removal of lane switches, MS) was only of minor importance. In Fig. 6, boxplots are shown containing the continuous (probability) predictions of one of the repetitions for both alcohol models. It is clearly shown that the difference in prediction score between alcohol and placebo is significantly larger when using multiple features.

4. Discussion

Sedative drugs and alcohol are well known to significantly influence driving behaviour, which can be evaluated by driving parameters such as SDLP (Verster and Roth, 2011). Accurate knowledge of these (side) effects is of crucial importance in the development and application of new psychoactive medicines. This study did create a model using machine learning to detect driving impairment due to the use of alprazolam and alcohol, with inclusion of multiple driving features rather than the SDLP alone. These models provided improved insight into the way driving behaviour was affected by alcohol and alprazolam. These models may serve as a new benchmark for analysis of newly developed drugs for improving driving safety evaluation, but this can only be shown after further work.

Alprazolam and alcohol significantly affected the main parameters of driving in the simulator and affected scores of safe driving (Huizinga et al., 2019). The current study showed that, if only the SLDP was considered, machine learning models could predict the intake of
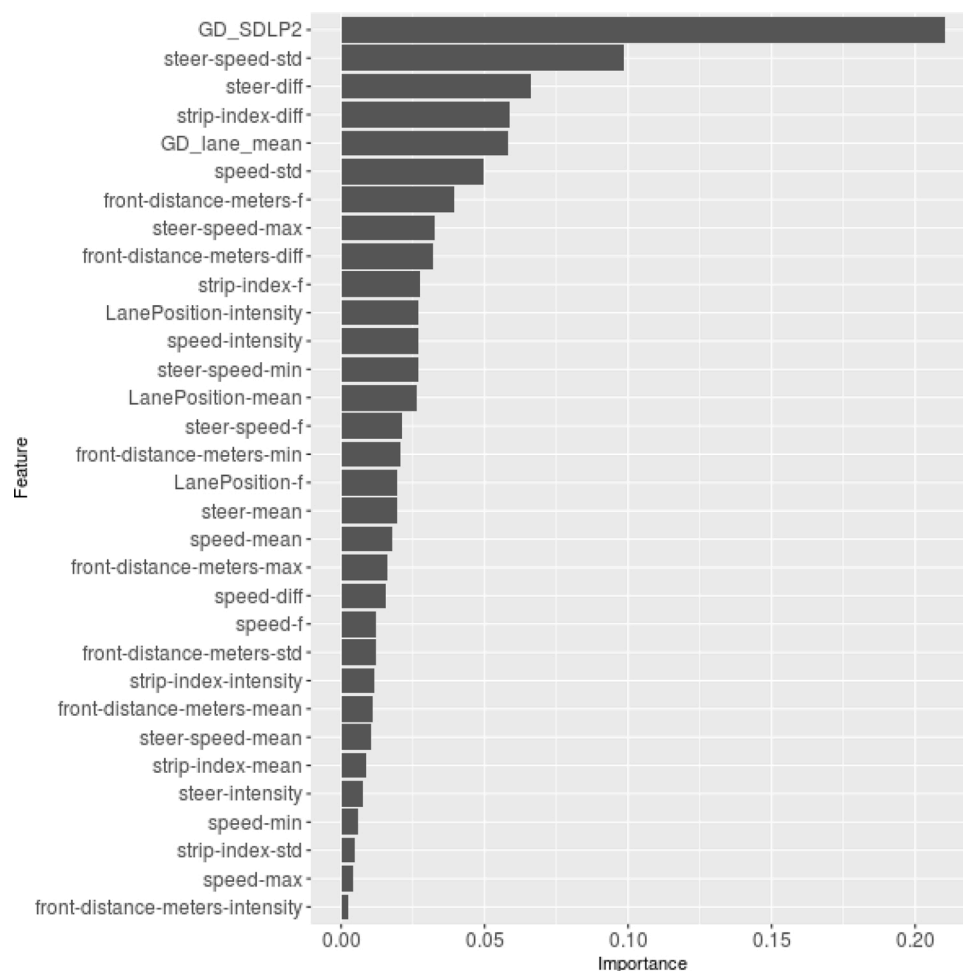

Fig. 2. Average feature importance of the models using all parameters predicting ingestion of alprazolam using all features. GD_SDLP, Standard deviation of the lateral position on the read (SDLP). GD_lane_mean, Mean Lane Position (MLP).
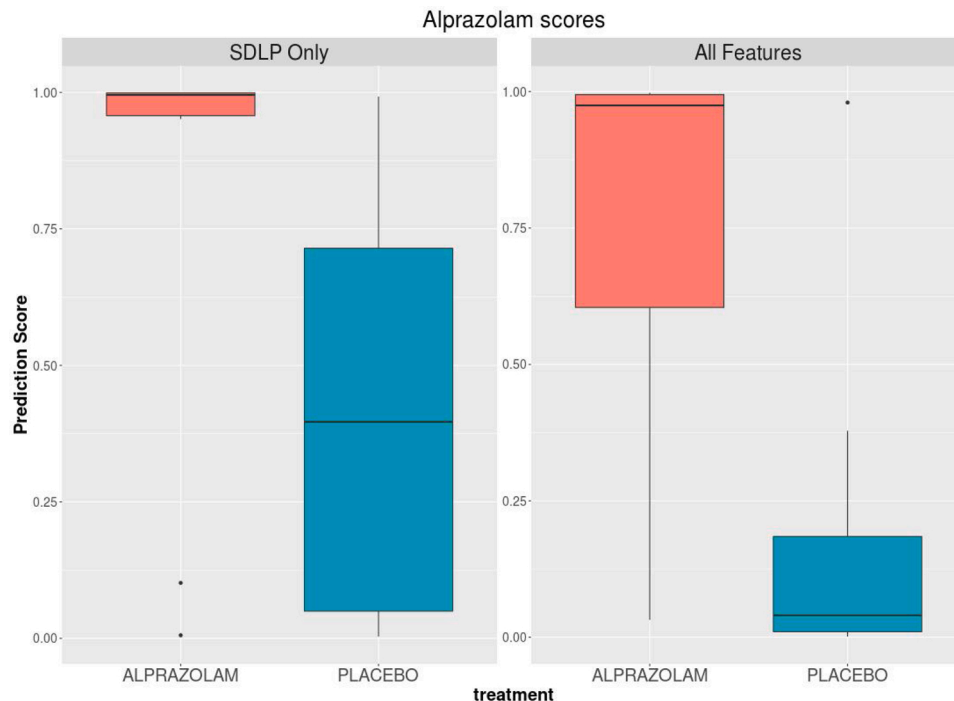
Fig. 3. Boxplots (indicating minimum, maximum, median, first and third quartiles) of the probability predictions of one of the repetitions of alprazolam intake. Left: the model using standard deviation of the lateral position (SDLP) only. Right: the model using all features.
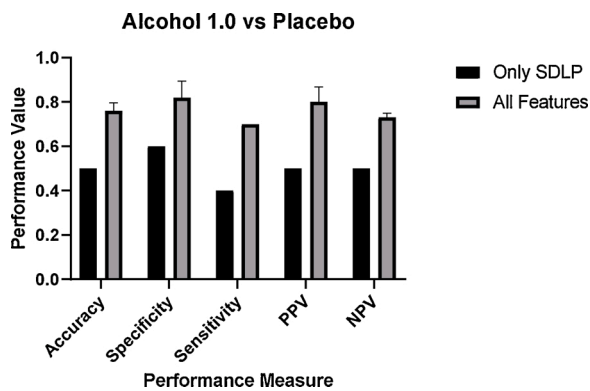

Fig. 4. Performances with standard deviation of the models using only the standard deviation of the lateral position (SDLP, black bars) and the models using all parameters on predicting alcohol intake (grey bars). PPV, positive predictive values. NPV, negative predictive values.

alprazolam or alcohol when trained on the data of 80 % of the subjects, and validated on the remaining 20 %. Accuracies obtained in SDLP only models were 65 % concerning alprazolam and 50 % as regards alcohol, indicating that the SDLP was more important for distinguishing alprazolam from placebo than alcohol. This is also shown in Figs. 2 and 5 on feature importance, which show that the feature importance is more equally distributed in the alcohol model than in the alprazolam model. Our findings suggest that alprazolam mainly affects various features affecting the lateral position of the car (SDLP, strip-index-diff, MLP), while alcohol also has a major influence on other driving parameters (such as steering speed, car speed and steer position). The relatively low accuracy percentages, as observed in the SDLP only models, are probably due to the high inter-subject variability. Since the effect of a medicine may vary substantially for each subject and the number of subjects in the datasets is relatively small, the change in SDLP may differ for subjects in a training set compared to subjects in a test set. In our dataset used for 'the alprazolam model', the prediction of alprazolam

use was difficult when based on the training set using SDLP only. Adding more features to train the model, the performance increased the predictive accuracy of alprazolam intake to over 80 %. This higher accuracy is associated with a clearer distinction between placebo and drug intake in the continuous / probability predictions.

Likewise, in the 'alcohol model', the predictive accuracy for alcohol ingestion increased to 76 %. These percentages are slightly higher than observed previously. (Chen and Chen, 2017) Chen et al. evaluated the effects of alcohol and additionally used physiological measurements in their model. In their study the authors could successfully distinguish drunk driving from normal driving with an accuracy of 70 %. Our results are likely to be more accurate, because we used a correction for baseline measurements of each subject. Part of the inter-subject variability can significantly be reduced by correcting the driving results for the baseline measurements of the same subject. This will make it substantially easier to evaluate the effect of a drug on driving behaviour. The increase in performance observed here is in line with the results of the study of Irwin et al. (Irwin, 2017) This study showed that alcohol affected not only the lateral control of the car, but also the longitudinal control of the car.

While for both the alprazolam and alcohol models the SDLP was a major determinant, the addition of other parameters such as the steering behaviour substantially increased the capacity to distinguish between drug usage and placebo. Extending on the current results in this manner it may be possible to develop systems that learn safe driving behaviour of an individual and detect abnormalities for that driver. This may be a substantial advantage particularly when assessing the effect of drugs or alcohol.

The models for alcohol and alprazolam showed differences in feature importance, indicating that the two interventions influence driving behaviour in a different way. This distinction between the effects of alcohol and alprazolam on driving behaviour could not be achieved as accurately with the use of SDLP only. Although multiple features have been analysed in several previous studies, these additional features are hardly used in the final assessment because of a lack of validation. The current study has shown that models combining all these driving features could solve this problem of validation. The probability predictions
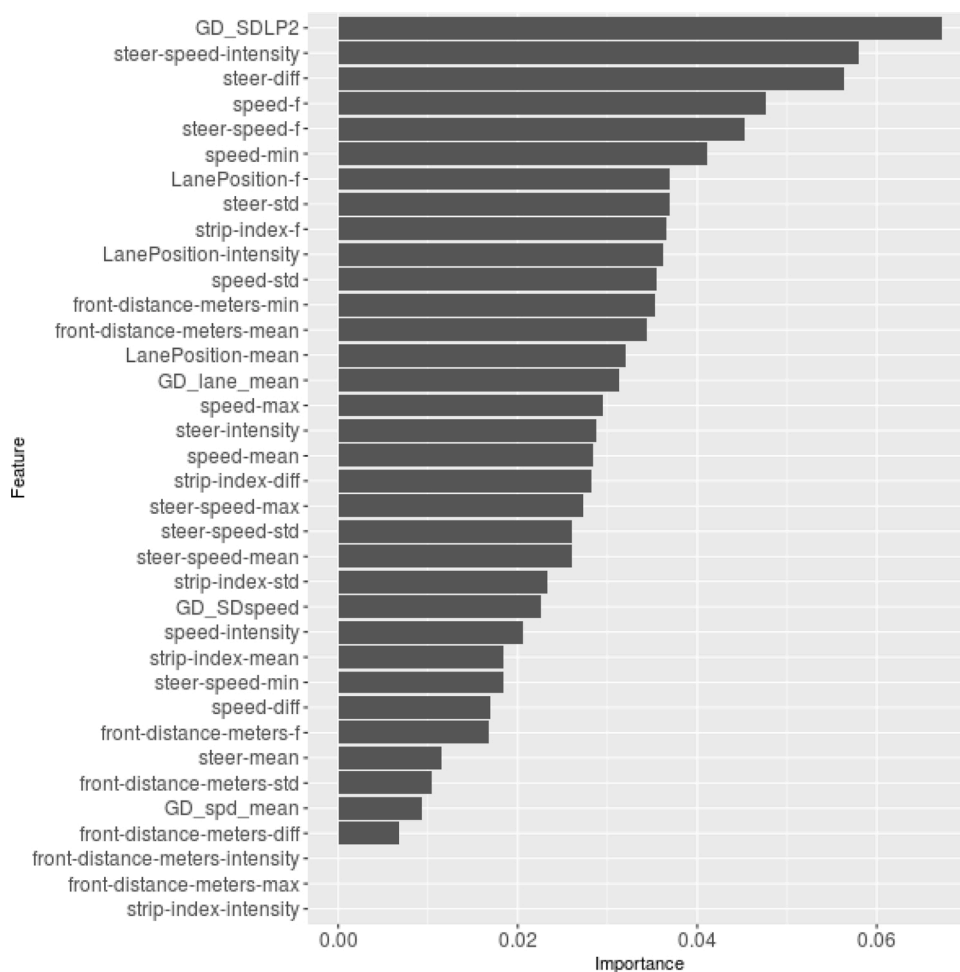
Fig. 5. Average feature importance of the models using all parameters predicting alcohol intake. GD_SDLP, standard deviation of the lateral position (SDLP). GD_lane_mean, mean lateral position (MLP). GD_SDspeed, standard deviation of speed (SDS). GD_spd_mean, mean speed (MS).

when using multiple driving features besides the SDLP show a clearer distinction between drug-induced and normal driving in comparison to using SDLP only. Even when a single feature does not show a significant change in driving behaviour, it may contribute to a significant outcome of a model that is combining multiple features. Therefore, the use of multiple features may allow a clear discrimination between the various effects of different interventions on car driving behaviour.

4.1. Limitations

In daily practice, a car accident caused by drug or substance is the true endpoint, but this is difficult to assess. It would seem a reasonable assumption that unsafe driving behaviour is a proxy for this endpoint. Preferably this proxy should be as predictive as possible.

Although the use of an ensemble machine model, such as the gradient boosting model used in this study, is more accurate and robust, this type of machine learning is accompanied by lack of interpretability (Wang et al., 2015). The importance of the features trained on can be extracted after training the model, but it is not directly clear how the features are being used by the model. The PPV and NPV are quite high, 83 % and 84 % for the alprazolam model, respectively, and 80 % and 73 % for the alcohol model, respectively, showing the reliability of the models. However, the model was tested on the measurements of 5 subjects in this study. Also, it has not yet been analysed how the model performs in other interventions.

A shortcoming of our study is that we only developed and optimised the models for alcohol and alprazolam and that we did not yet test them

for other similar medicines. The generalisability of the models (for instance for other benzodiazepine-like medicines or even other psychoactive medicines) requires more work and the models will be supplied (supplementary material) for other research groups for further validation.

These new prediction models can be used to create a unique 'fingerprint' (profile) with respect to both desired and undesired effects on driving. However, any inability regarding detection of deviant driving behaviour might be related to limitations of this test battery to detect driving behaviour abnormalities induced by novel compounds. This requires further studies. When these current uncertainties are resolved the proper use of created models in early drug development can provide important information that can be used to make a go/no-go decision regarding further development of new drugs. Similarly, they can be used to guide the decision-making process regarding the dosage range to be used in phase II studies, determining a therapeutic window, and even identifying the target study population (Groeneveld et al., 2016). This way novel psychopharmacological drugs could be tested for effect on on driving behaviour in the early phase of development. Using these models adequate probability scores can be given to test-drives, which provide an indication about the way and the extent to which these drugs are modifying driving behaviour.

5. Conclusion

In our study we demonstrated how machine learning may improve the assessment of drug-induced deviant driving behaviour. In particular,
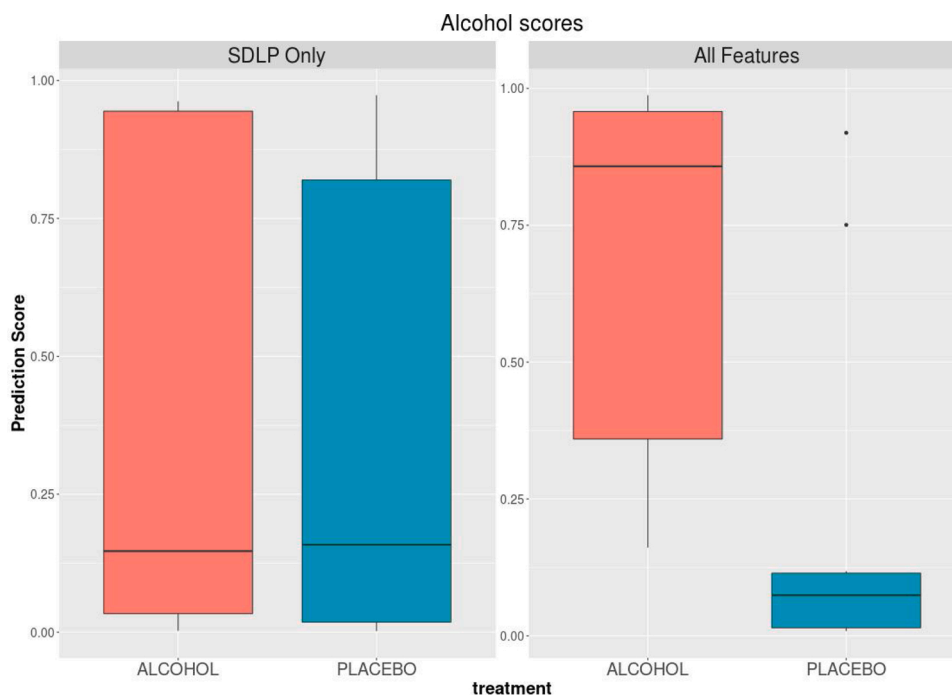
Alcohol scores

Fig. 6. Boxplots (indicating minimum, maximum, median, first and third quartiles) of the probability predictions of one of the repetitions of alcohol intake. Left: the model using standard deviation of the lateral position (SDLP) only. Right: the model using all features.

the inclusion of multiple driving features rather than SDLP alone as is currently the state-of-the art, improved the performance of a model characterizing the way driving behaviour was affected by alcohol or alprazolam. These models may facilitate quantitative description of deviant driving behaviour in the development and application of psychopharmacological medicines but require further evaluation in other groups of substances before they can be used to evaluate new (unknown) interventions.

CRediT authorship contribution statement

H.E.C. van der Wall: Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing. **R.J. Doll:** Methodology, Supervision, Writing - review & editing. **G.J.P. van Westen:** Methodology, Supervision, Writing - review & editing. **I. Koopmans:** Investigation, Methodology, Writing - review & editing. **R. G. Zuiker:** Investigation, Writing - review & editing. **J. Burggraaf:** Supervision, Writing - review & editing. **A.F. Cohen:** Methodology, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

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

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.aap.2020.105822.

References

Arnedt, J., Wilde, G.J.S., Munt, P.W., et al., 2000. Simulated driving performance following prolonged wakefulness and alcohol consumption: separate and combined contributions to impairment. J. Sleep Res. 9 (3), 233–241.

Bagley, S.C., White, H., Golomb, B.A., 2001. Logistic regression in the medical literature: standards for use and reporting, with particular attention to one medical domain. J. Clin. Epidemiol. 54 (10), 979–985.

BIOVIA, D.S., 2018. Pipeline Pilot (version 2018). Biovia.

Brunnauer, A., Laux, G., Zwick, S., 2009. Driving simulator performance and psychomotor functions of schizophrenic patients treated with antipsychotics. Eur. Arch. Psychiatry Clin. Neurosci. 259 (8), 483–489.

Brunnauer, A., Buschert, V., Segmiller, F., Zwick, S., Bufler, J., Schmauss, M., et al., 2016. Mobility behaviour and driving status of patients with mental disorders–an exploratory study. Int. J. Psychiatry Clin. Pract. 20 (1), 40–46.

Bunn, T., 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.

Chen, H., Chen, L., 2017. Support vector machine classification of drunk driving behaviour. Int. J. Environ. Res. Public Health 1 (14), 108.

Chen, T., Cao, Y., Zhang, Y., Liu, J., Bao, Y., Wang, C., et al., 2013. Random forest in clinical metabolomics for phenotypic discrimination and biomarker selection. Evid. Based Complement. Altern. Med. 2013.

Darby, P., Murray, M., Raeside, R., 2009. Applying Online Fleet Driver Assessment to Help Identify Target and Reduce Occupational Road Safety Risks Safety Science, 47, pp. 436–442 (3).

Deo, R., 2015. Machine learning in medicine. Circulation 132 (20), 1920–1930.

Du, X., Min, J., Shah, C.P., Bishnoi, R., Hogan, W.R., Lemas, D.J., 2020. Predicting in-hospital mortality of patients with febrile neutropenia using machine learning models. Int. J. Med. Inform., 104140

Grabe, H.-J., Wolf, T., Grätz, S., Laux, G., 1998. The influence of polypharmacological antidepressive treatment on central nervous information processing of depressed patients: implications for fitness to drive. Neuropsychobiology 37 (4), 200–204.

Groeneveld, G.J., Hay, J.L., Van Gerven, J.M., 2016. Measuring blood–brain barrier penetration using the NeuroCart, a CNS test battery. Drug Discov. Today Technol. 20, 27–34.

Guo, F., Fang, Y., 2013. Individual driver risk assessment using naturalistic driving data. Accid. Anal. Prev. 61, 3–9.

Hegde, J., Rokseth, B., 2020. Applications of machine learning methods for engineering risk assessment–A review. Saf. Sci. 122 (104492).

Huizinga, C., Zuiker, R.G., de Kam, M.L., et al., 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, Cea., 2017. Ffects of acute alcohol consumption on measures of simulated driving: a systematic review and meta-analysis. Accid. Anal. Prev. 102, 248–266.

Iwata, M., et al., 2018. Evaluation method regarding the effect of psychotropic drugs on driving performance: a literature review. Psychiatry Clin. Neurosci. 72 (10).

Jauhar, P., McClure, I., Hillary, C., Watson, A., 1993. Psychomotor performance of patients on maintenance lithium therapy. Hum. Psychopharmacol. Clin. Exp. 8 (2), 141–144.

Leufkens, T., Ramaekers, J.G., De Weerd, A., Riedel, W., Vermeeren, A., 2014a. On-the-road driving performance and driving-related skills in older untreated insomnia patients and chronic users of hypnotics. Psychopharmacology 231 (14), 2851–2865.

Leufkens, T., Ramaekers, J.G., De Weerd, A., Riedel, W., Vermeeren, A., 2014b. Residual effects of zopiclone 7.5 mg on highway driving performance in insomnia patients and healthy controls: a placebo controlled crossover study. Psychopharmacology 231 (14), 2785–2798.

Liguori, A., 2009. Simulator studies of drug-induced driving impairment. Drugs driving and traffic safety 75–82.

Mets, M., Kuipers, E., de Senerpont, D.L.M., et al., 2011. Effects of alcohol on highway driving in the STISIM driving simulator. Hum. Psychopharmacol. Clin. Exp. 26 (6), 434–439.

Moore, N., 1977. Medazepam and the driving ability of anxious patients. Psychopharmacology 52 (1), 103–106.

O'Hanlon, J.F., Vermeeren, A., Uiterwijk, M.M., van Veggel, L.M., Swijgman, H.F., 1995. Anxiolytics' effects on the actual driving performance of patients and healthy volunteers in a standardized test. Neuropsychobiology 31 (2), 81–88.

Obermeyer, Z., Ezekiel, J.E., 2016. Predicting the future big data machine learning and clinical medicine. N. Engl. J. Med. 375 (13).

Paltrinieri, N., Comfort, L., Reniers, G., 2019. Learning about risk: machine learning for risk assessment. Saf. Sci. 118, 475–486.

Robertson, R.D., Hing, M.M., Pashley, C.R., Brown, S.W., Vanlaar, W.G., 2017. Prevalence and trends of drugged driving in Canada. Accid. Anal. Prev. 99, 236–241.

Sagberg, F..Selpi, Bianchi Piccinini, G.F., Engström, J., 2015. A review of research on driving styles and road safety. Hum. Factors 57 (7), 1248–1275.

Shinar, D., 2017. Traffic Safety and Human Behavior. Emerald Group Publishing.

Soyka, M., Winter, C., Kagerer, S., Brunnauer, M., Laux, G., Möller, H.-J., 2005. Effects of haloperidol and risperidone on psychomotor performance relevant to driving ability in schizophrenic patients compared to healthy controls. J. Psychiatr. Res. 39 (1), 101–108.

Subasi, A., Gursoy, M.I., 2010. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst. Appl. 37 (12), 8659–8666.

Tango, F., Botta, M., 2013. Real-time detection system of driver distraction using machine learning. IEEE Trans. Intell. Transp. Syst. 14 (2), 894–905.

Van Laar, M.W., Volkerts, E.R., Van Willigenburg, A.P., 1992. Therapeutic effects and effects on actual driving performance of chronically administered buspirone and diazepam in anxious outpatients. J. Clin. Psychopharmacol.

Vanlaar, W., 2005. Drink driving in Belgium: results from the third and improved roadside survey. Accid. Anal. Prev. 37 (3), 391–397.

Vanlaar, W.G., Robertson, R.D., 2010. Road safety monitor 2010: drinking and driving in Canada. TIRF road safety monitor 2010 (10L), 1–6.

Verster, J., Roth, T., 2011. Standard operation procedures for conducting the ontheroad driving test and measurement of the standard deviation of lateral position SDLP. Int. J. Gen. Med. 359 (4).

Wang, J., Ryohei, F., Yosuke, M., 2015. Trading interpretability for accuracy: oblique treed sparse additive models. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 1245–1254.

Wylie, K.R., Thompson, D.J., Wildgust, H.J., 1993. Effects of depot neuroleptics on driving performance in chronic schizophrenic patients. J. Neurol. Neurosurg. Psychiatr. 56 (8), 910–913.

Zhang, Z., Zhao, Y., Canes, A., Steinberg, D., Lyashevska, O., 2019. Predictive analytics with gradient boosting in clinical medicine. Ann. Transl. Med. 7 (7).

Zoethout, R.W.M., 2012. Applications of Alcohol Clamping in Early Drug Development. Centre for Human Drug (CHDR), Faculty of Medicine, Leiden University Medical.