In various papers it is suggested that up to now 20-30% of drivers regularly use drugs which can act and interact on skills related to driving (Kibrick & Smart, 1970; Glauz & Blackburn, 1975; Linnoila, 1976; Lukkari et al. 1976).

For this reason, extensive research in the field of drugs and driving has become an absolute necessity. One of the most commonly detected drugs among drivers arrested for driving under the influence of alcohol, is diazepam. (Twente et al., 1975; Alha, et al., 1977).

This drug is a tranquillizer of the benzodiazepine type and is widely used for the symptomatic relief of anxiety, insomnia, psychiatric disturbances, seizures and as preoperative medication (Greenblatt & Shader, 1974).

Epidemiological studies on benzodiazepines (especially on diazepam) and driving are scarce. Bø et al. (1975) demonstrated that diazepam was present in the plasma of 11% of injured traffic accident victims who were hospitalized; significantly less diazepam (2%) was found in a reference group. Recently, Skegg et al. (1979) reported the results of a prospective study of 43,117 people in which prescriptions issued by general practitioners over two years were linked with records of hospital admissions and deaths. For 57 people injured or killed while driving cars, motor cycles, or bicycles, the medicines that had been dispensed in the three months before were compared with those dispensed for 1425 matched controls. There was a highly significant association between the use of minor
tranquillizers, e.g. diazepam, and the risk of a serious road accident. Experimental studies on the effects of diazepam on driving related skills yielded conflicting results. The effects of benzodiazepines on psychomotor performance were critically reviewed by Wittenborn (1979), while Kleinknecht & Donaldson (1975) reviewed the effects of diazepam on cognitive and psychomotor performance. Clayton (1976) has demonstrated the lack of unanimity between the various laboratory studies of benzodiazepine derivatives. The differences are due to the interaction between acute and chronic dose regimens of the drug, patient and healthy volunteers and experimental methods used.

Attempts to extrapolate from the laboratory situation to the conditions in the real world are of no use in situations as mentioned above where a great validity problem exist. Many of the laboratory studies of the benzodiazepine derivatives have used an acute dose administration which does not represent the clinical situation where patients are consistently placed in chronic dose regimens. Furthermore, laboratory-experiments are hardly correlated to the complex skill of real driving. Subjects motivation to react in a certain way under simulated conditions mostly differs from their motivation under actual condition in real driving. Many authors (Biehl, Klonoff, 1974) agree that measuring influences on driving performance in a real driving situation is the best representation of what happens on the road. Although this approach has its own methodological problems it gives valid information about the performance of patients in real driving. A comparison with the performance in a laboratory experiment offers the opportunity to evaluate the laboratory tasks.

The primary goal of this study was to compare the performance of ambulant patients receiving diazepam medication on a laboratory task with their real driving performance. The performance on these tasks should be related to levels of diazepam and N-desmethyldiazepam in body fluids. To accomplish this goal a laboratory experiment and a driving test were selected (De Gier, et al., 1980a).

The design of the laboratory task was based on the observation by Nelemans (1968) that benzodiazepines could cause accidents on a quiet highway whereas the patients had just performed adequately in dense city traffic. It was felt that a long lasting experiment should be used to investigate an influence of diazepam on psychomotor and cogni-
tive skills. The design should be appropriate to vary the level of attention in order to simulate the attention demanding properties of the traffic conditions as mentioned above.

Some experiments are known in which auditory vigilance performance was studied after diazepam use by volunteer subjects (Hart, et al., 1976; Jones, et al., 1978). According to these authors it also seemed likely that protracted, repetitive and boring tasks would be more sensitive than short, entertaining or engrossing ones. However due to the conflicting results and the fact that a visual vigilance test is much more related to the driving task, it was decided to use an experiment based on visual vigilance. An experiment consisting of a high and a low attention demanding task was chosen.

Measurement of real driving performance was based on the method described by De Gier (1979).

In order to study level response relationships it was decided to take blood and saliva samples during the experiments. These samples were taken to determine the diazepam and N-desmethyldiazepam (the major metabolite) concentration.

The use of saliva samples in drugs and driving research was indicated by De Gier (1977). Saliva analysis may be an important approach to the dose-response problem since the concentration of most drugs in saliva corresponds to the free or unbound, pharmacological active, plasma drug concentration. Furthermore, the use of a non-invasive sampling technique prevents traumatic influences on performance.

Materials and Methods

Subjects
Twenty two male subjects participated in the study and were under normal medical control of the same physician. Nine subjects (mean age 45.6 years, standard deviation (SD) 9.6 years) were out-patients receiving diazepam medicamentation. Their number of years of driving experience was 23.3 years with a SD of 8.2 years. Thirteen subjects served as controls (mean age 40.6 years, SD 8.4 years) and their mean number of years of driving experience was 18.2 years with a SD 6.4 years.

Diazepam was given orally on prescription in a dose regimen of three
Driving Test  
During the test driving skills were rated by a trained observer in the front seat of the car using a 110 items check list. The items reflect behavioral components and cognitive skills involved in driving. Items were rated on a three point scale as "satisfactory", "moderate" and "insufficient". A final driving ability score was assessed by calculating the total number of items scored as "insufficient" in a selection of 22 so called "important" items. The importance of an item was indexed regarding the fundamental contribution of this item to traffic safety. A full description of the driving test conditions is given elsewhere (De Gier, 1979).

In order to decide which score on the 22 items would constitute a driver to be a hazard on the road the driving scores of 20 beginning drivers were used to accomplish this criterion. These drivers performed our driving test within seven days after they failed the practical examination for a license of the Dutch licensing authorities.

Observers  
The observers were employed by the Traffic Department of the Royal Dutch Tourist Association (A.N.W.B.). Normally they act as examinators for the advanced drivers test of the Association. The four observers were not aware of the subjects drug use, if any. They were ad random asked to score a certain subject, without knowing if the subject was a control subject or a patient.

To investigate the interjudge reliability, an experiment with five drivers was performed. These drivers were qualified drivers and not included in the experiment. First these drivers also performed a whole test to become familiarized with the car and driving area. Thereafter they participated in sessions with each of the observers.

Laboratory tests  
It was suggested that patients should not be able to perform a task demanding a low level of attention, assuming that this low level was indeed as such recognized by the patients. This impaired performance would also be shown when such a task is immediately preceded by a
A high-attention demanding task; a task the patients should perform like control subjects. This phenomenon could cause the "benzodiazepine type of accident" (Nelemans, 1968).

In order to simulate this phenomenon in a laboratory situation it was decided to carry out an experiment consisting of two tasks. The first task (task A) is expected to be a high-attention demanding task, while the second task (task B) is expected to be a low-attention demanding task. Both tasks are performed for one hour each, without an interval.

The stimulus material for both tasks is displayed on a BARCO color CRT (type CM 50 RGB), driven by a RAMTEK graphic controller (type 6200) and a DEC PDP 11/40 computer. The subjects performed the tasks in a soundproof room.

In the high-attention demanding task A the subject had to detect randomly introduced irregularities in the sequence in which a moving dot in a 10 x 10 stationary pattern of small open circles lights up (figure 1). Normally the dot moved from left to right, and top to bottom. An irregularity (b) in the sequence was said to occur if the moving dot skipped one dot of the stationary pattern. A normal sequence is said to occur if the moving dot did not skip a dot (a), or skipped two dots (c) of the stationary pattern.

After each displacement of the moving dot, one out of four possible classes of subject responses was obtained, and stored by the computer:

- **Hit:** irregularity detected by the subject
- **Miss:** irregularity not detected by the subject

![Figure 1. Dot pattern of task A.](image1)

![Figure 2. Moving square of task B.](image2)
false alarm: no irregularity, but irregularity reported by the subject

correct rejection: no irregularity and no irregularity reported by the subject

No feedback about actual performance was given.

The low-attention demanding task B was a task during which the subject had to keep the joystick controlled cursor within the boundaries of a slowly moving (0.5 cm/sec) CRT displayed square (dimension 70/30 mm) (figure 2).

The square moved in the horizontal direction only. When one of the sides of the square reached a display field boundary, the direction of movement of the square was changed. Changes of direction of movement were also randomly introduced. The subject's performance on this task was recorded by the computer, and defined as the number of seconds that the centre of the cursor was outside the boundaries of the square. Again no feedback about actual performance was given. A full description of both tasks is given by De Gier, et al. (1980a).

Blood and saliva samples

Blood samples were taken 2 h after the start of the driving test and the laboratory test. Saliva samples were taken just before the start of each test and just before the blood sampling. Saliva samples were collected after stimulation of the secretion by chewing on a slice of Teflon.

Diazepam and N-desmethyldiazepam were determined by electron-capture gas chromatography according to the method described by De Gier & 't Hart (1979). This procedure for plasma determinations has also been used successfully to quantitate concentrations of diazepam and N-desmethyldiazepam in saliva (De Gier et al. 1980b).

Experimental Design

Ignoring the difference between patients and control subjects, two subjects were assigned to each day of experimentation.

The two subjects were tested individually on this day; one subject started the driving test, the other subject the laboratory test. After a lunch they switched for the other part.

Breakfast and lunch were standardized (orange juice and sandwiches). Furthermore subjects were instructed not to drink alcoholic beverages.
within 24 h preceding their tests.
Subjects were allowed to smoke between the sessions and to have one cup of coffee at lunch. Furthermore they were instructed not to speak about their diazepam use, if any.

RESULTS
Driving test
The data used to investigate the interjudge reliability, are presented in table 1. According to these figures it was decided to discard the data of the patients and control subjects derived from the judgement of the fourth observer D.

Table 1. Scores as percentage "insufficient" items of 5 reference subjects (qualified drivers) obtained with each of the observers

<table>
<thead>
<tr>
<th>Reference subject</th>
<th>Observer A</th>
<th>Observer B</th>
<th>Observer C</th>
<th>Observer D</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>5.1</td>
<td>-1)</td>
<td>10.0</td>
</tr>
<tr>
<td>2</td>
<td>3.4</td>
<td>6.8</td>
<td>8.3</td>
<td>32.2 *</td>
</tr>
<tr>
<td>3</td>
<td>1.7</td>
<td>1.7</td>
<td>0</td>
<td>6.7 *</td>
</tr>
<tr>
<td>4</td>
<td>3.4</td>
<td>5.2</td>
<td>6.7</td>
<td>50.0 *</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>3.4</td>
<td>1.7</td>
<td>4.9 *</td>
</tr>
</tbody>
</table>

1) no data available
* p < 0.05, Friedman two-way analysis of variance by ranks

According to the scores of the 20 beginning drivers it was found that the scores on six particular items showed a good agreement. These items reflected "visual perception" and "anticipation of events". The scores on these items were highly significant correlated with the over-all scores on the 22 items (Spearman rank correlation coefficient of .96). None of the beginning drivers had less than four items scored as "insufficient" of the six particular items. This finding enabled us to divide our subjects into two classes: inadmissible and admissible drivers. The last class showed less than four items scored as "insufficient" of the six items reflecting "visual perception" and "anticipation of events".
Using the classes as mentioned, the scores of the subjects are represented by frequencies in a 2 x 2 contingency table. The exact probability of the occurrence of more inadmissible drivers in the patient group was .083 (Fisher exact probability test). In order to gain more insight into the significance of this difference it was decided to perform a power analysis. Cox (1969) described a method of analysis based on the linear logistic model. By using this method an exact power of .73 was indicated for the difference between the groups on a logistic scale.

The results of the driving test are given in table 2. The patients scored significantly (p < 0.05), Mann Whitney U test) more "insufficient" on the 22 important items. Significant effects for individual items were derived after analyses by use of the Fisher exact test (p < 0.10). These items were: "take account of obstruction of the view by passengers present", "to avoid cutting off corners", "observation while driving on a roundabout", "visual perception" and "anticipation of events".

Table 2. Number of items scored as "insufficient" (I) and "moderate" + "sufficient" (II) of 9 patients and 8 control subjects in a total of the 22 most important items

<table>
<thead>
<tr>
<th>Patient</th>
<th>Observer</th>
<th>I</th>
<th>II</th>
<th>Control subject</th>
<th>Observer</th>
<th>I</th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>C</td>
<td>0 (0)</td>
<td>22</td>
<td>1</td>
<td>C</td>
<td>0 (0)</td>
<td>22</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>1 (0)</td>
<td>21</td>
<td>2</td>
<td>B</td>
<td>6 (1)</td>
<td>16</td>
</tr>
<tr>
<td>3</td>
<td>B</td>
<td>11 (5)</td>
<td>11</td>
<td>3</td>
<td>B</td>
<td>1 (0)</td>
<td>21</td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>13 (5)</td>
<td>9</td>
<td>4</td>
<td>C</td>
<td>12 (5)</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>7 (3)</td>
<td>15</td>
<td>6</td>
<td>C</td>
<td>0 (0)</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>C</td>
<td>8 (4)</td>
<td>14</td>
<td>9</td>
<td>B</td>
<td>4 (2)</td>
<td>18</td>
</tr>
<tr>
<td>7</td>
<td>C</td>
<td>11 (6)</td>
<td>11</td>
<td>10</td>
<td>B</td>
<td>6 (3)</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>B</td>
<td>8 (3)</td>
<td>14</td>
<td>13</td>
<td>C</td>
<td>0 (0)</td>
<td>22</td>
</tr>
<tr>
<td>9</td>
<td>B</td>
<td>10 (4)</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1) Subjects 5, 7, 8, 11 and 12 were judged by observer D and thus rejected
2) () number of items scored as "insufficient" of the six items reflecting "visual perception" and "anticipation of events".
Laboratory test

The results of the laboratory test are given in figure 3. The patients showed significantly less (P < 0.01, Student t-test) "false alarms" in task A and significantly more (P < 0.05, Student t-test) "seconds out" in task B. Figure 4 shows the latter results where the "seconds out" are presented as periods of sustained seconds.

Figure 3. Results of the laboratory test

Figure 4. Results of task B presented as periods of sustained seconds

Correlation between Driving test and Laboratory test

The score on the 22 items of the driving test was significantly (P < 0.05) correlated with the score (as "seconds out") derived from task B of the laboratory test. The Spearman rank correlation coefficient (r_s) was .52. The results of the patients and control subjects on task B and expressed within the meaning of the two classes of drivers (inadmissible and admissible) are given in figure 5.

1017
Figure 5. Correlation between driving test and laboratory test.

As number of items scored as "insufficient" of the six items reflecting "visual perception" and "anticipation of events". (⭐ = score is zero)

*Saliva and plasma samples*

The results of the saliva and plasma concentration measurements of diazepam and N-desmethyl-diazepam are given in figures 6 and 7. Significant correlations were obtained between saliva and plasma levels of diazepam and N-desmethyl-diazepam. The saliva/total plasma concentration ratios for diazepam and N-desmethyl-diazepam were 0.013 ± 0.002 and 0.018 ± 0.004, respectively.

No significant correlations (all correlation coefficients were < .25) existed between driving performance and concentration of diazepam or N-desmethyl-diazepam in saliva and plasma. The same lack of correlation was found between task performance and drug level measurements in body fluids.
DISCUSSION

Due to the significant more rigorous level of judgement by observer D we decided to reject the data of the subjects judged by this observer as indicated in table 2. Afterall this rejection served a better matching of the observers because no patients were judged by observer D.

The results of the driving test indicated that patients receiving diazepam medication performed worse by higher "insufficient" scoring on the 22 most important items. As to the items where a worse performance was observed it may be concluded that these patients show an impairment in "visual perception" and "anticipation of events".

The results of the laboratory test (fig. 3) need some discussion. Because a lower false-alarm rate derived from the patients it looked like they showed a better performance in task A. The performance pattern of both tasks by the control subjects was a similar pattern as was found in two previous experiments with totally different groups of subjects (students) where the influence of alcohol on these tasks
was investigated (De Gier, 1980).

The decrement in mean number of false alarms in task A by the patients is probably due to the motivation of the patients to show their best because they were informed about the aims of our experiments. Nevertheless they performed worse in the simple low attention demanding task.

In our experiment the boring tasks (task B) proved to be more sensitive for an impairment of performance. Especially noticed in periods of sustained "seconds out" (fig.4) the patients scored prolonged outside the retangle. Further investigations need to be done to indicate the importance of this finding for real driving situations.

A correlation between performance in task B and driving performance existed. Expressed within the meaning of potential hazard on the road (figure 5) a certain score in task B seemed to discriminate between the patients and control subjects. After a score of 30 "seconds out" only patients (n = 5) were involved. Four of them showed a driving performance like beginning drivers that fail to pass an examination in order to obtain a license. The feasibility of the use of this correlation for screening of impaired drivers is now subject of further investigations.

It should be emphasized that the performance of patients receiving diazepam medication has been studied. It is not possible to blame diazepam alone for the impairment of driving behaviour and task performance. It is the combination "patient and drug" that showed an impairment of performance. Therefore this combination needs more attention regarding the drugs and driving problem.

The results of the saliva and plasma measurements of diazepam and N-desmethyldiazepam (figures 6 and 7) indicate the possibility of using saliva instead of plasma level measurement to establish level response studies in the future. The comparison of saliva and plasma diazepam measurements is recently discussed by De Gier et al. (1980b). In conclusion, the results in this study indicate the feasibility of the use of a real driving test. Its use is illustrated in a situation where differences in driving behaviour between two groups of subjects could be expected. The development of a laboratory experiment in order to accomplish information validated by results obtained in the real life situation is the next step to solve the drugs and driving problem.
problem. This approach is illustrated by use of patients receiving diazepam (as an example of the benzodiazepines) medication. Their performance on the road and in the laboratory was significantly impaired.

ACKNOWLEDGEMENTS
The authors are indebted to Ms. B.J. 't Hart, Department of Pharmacotherapy, for her technical assistance, and to H. Bergman and H. Zwaga, Psychological Laboratory, State University Utrecht, for their advice and constructive criticism.
This work was supported in part by grants from Hoffmann-La Roche, The Netherlands Association of Doctors Motorists (V.V.A.A.) and the Royal Dutch Tourist Association (A.N.W.B.).

REFERENCES
Biehl, B.: The effects of two tranquillizers on driving performance as measured in the normal driving task. Paper presented at the 18th Congress of Int. Ass. Applied Psychology, Montreal, 1974
Gier, J.J. de: A subjective measurement of the influence of ethyl alcohol in moderate levels on real driving performance. Blutalkohol


Gier, J.J. de: Contribution to the evaluation of the drugs and driving problem, Thesis, Utrecht, 1980


Twente, R.T., Theeuwen, A.B.E., Verwey, A.M.A.: Enige ervaringen met een gaschromatografische methode voor kwalitatief geneesmiddelen-