An Approach Towards Reducing Road Traffic Injuries and Improving Public Health Through Big Data Telematics: A Randomised Controlled Trial Protocol

Mehrdad Azmin, PhD; Ayyoob Jafari, PhD; Nazila Rezaei, MD; Kavi Bhalla, PhD; Dipan Bose, PhD; Saeid Shahraz, MD, PhD; Mina Dehghani, PhD Candidate; Parastoo Niloofar, MSc; Soraya Fathollahi, PhD Candidate; Javad Hedayati, MSc; Hamidreza Jamshidi, PhD; Farshad Farzadfar, MD, MPH, MHS, DSc

Abstract

Objective: Deaths due to road traffic accidents (RTAs) are a major public health concern around the world. Developing countries are over-represented in these statistics. Punitive measures are traditionally employed to lower RTA related behavioural risk factors. These are, however, resource intensive and require infrastructure development. This is a randomised controlled study to investigate the effect of non-punitive behavioural intervention through peer-comparison feedback based on driver behaviour data gathered by an in-vehicle telematics device.

Design, Setting, and Participants: A randomised controlled trial using repeated measures design conducted in Iran on the drivers of 112 public transport taxis in Tehran province and 1309 inter-city busses operating nationwide. Driving data is captured by an in-vehicle telematics device and sent to a centrally located data centre using a mobile network. The telematics device is installed in all vehicles. Participants are males aged above 20 who have had the device operating in their vehicles for at least 3 months prior to the start of the trial.

Intervention: The study had three stages: 1- Driver performance was monitored for a 4-week period after which they were randomised into intervention and control groups. 2- Their performance was monitored for a 9-week period. At the end of each week, drivers in the intervention group received a scorecard and a note informing them of their weekly behaviour and ranking within their peer group. Drivers in the control group received no feedback via short messaging service (SMS). 3- Drivers did not receive further feedback and their behaviour was monitored for another 4 weeks.

Primary and Secondary Outcome Measure: Primary outcome was changes in weekly driving score in intervention and control groups during stage 2 of intervention. Taxis and busses were analysed separately using generalised estimating equation analysis.

Funding and Ethical Approval: This project was funded by the National Institute for Medical Research Development (Grant No.940576) and approved by its ethics committee (Code: IR.NIMAD.REC.1394.016). This trial was registered at www.irct.ir as IRCT20180708040391N1.

Keywords: Behavioural intervention, Big data, Public health, Road traffic injury, Telematics

Cite this article as: R. An approach towards reducing road traffic injuries and improving public health through big data telematics: a randomised controlled trial protocol. Arch Iran Med. 2018;21(11):495–501.

Rationale and Objectives

Road traffic accidents (RTAs) are a leading cause of death and injury around the world. It has been estimated that in 2013, globally, 1.2 million people died and many more were injured from being involved in RTAs. This has led to the Sustainable Development Goal 3.6 (SDG 3.6), requiring the member states to halve the rates of deaths and injuries from RTAs by 2020. Approximately 90% of deaths due to RTAs occur in low to middle income countries even though these countries have 54% of the world’s vehicles. At 34–44 deaths per 100,000, Iran has one of the highest rates of road traffic mortality, almost double that of the...
global road mortality rate at 18 deaths per 100,000.  

The high rates of RTAs are attributable to issues relating to enforcement of safety regulations, driving environment, and user behaviour. Interventions by the traffic police in Iran through enforcement of various driving safety legislations have been shown to have lowered RTA related fatality and morbidity.  

Large-scale and continued implementation of such measures, however, face several challenges. Speeding, for example, is one of the most significant contributing risk factors in the occurrence and aftermath of RTAs. Increasing the speed of a vehicle from 50 to 80 km/h decreases the survival chance of an adult pedestrian struck by a car from 80% to 40%. The primary method for prevention of over-speeding is the use of speed cameras. Effective operation of speed cameras requires financial resources for the costs of equipment, maintenance, and the infrastructure necessary for its operation. Such preventative measures are, therefore, employed in limited scopes and do not provide the required coverage. It is thus necessary to explore other methods of reducing RTA risk factors.

The project presented here aims to address a number of issues regarding the identification and rectification of RTA risk factors by collection of real world driving data. It focuses primarily on reducing driver behavioural risks through low-cost intervention methods, which is the subject of this paper. It additionally investigates the possibility of detecting infrastructure issues such as road design and condition through analysis of the collected driving data. Finally, it seeks to deal with vehicle specific risk factors through collection and analysis of vehicle diagnostic data.

### Study Design and Methods

#### Definitions

**Behavioural Interventions**

Behavioural interventions have been developed to affect individuals’ actions and modify their traits. They are widely used to address population health issues such as smoking cessation, low physical activity, alcohol misuse and illicit drug use, unhealthy diet, sexual risk taking, and unsafe driving behaviour. There are three important steps in planning behavioural change. These are: correct identification of target behaviour and outcomes, suitable choice of methods and interventions, and appropriate implementation of the intervention. One of the most popular of these interventions is “Nudging”, which seeks to influence an individual to make “better” decisions through manipulation of their choice architecture in a low-cost and not heavily invasive manner.

**In-Vehicle Telematics**

Telematics refers to monitoring of a vehicle by combining its on-board computer and a GPS system. It is increasingly employed in the automobile insurance industry and fleet management systems (FMS) across Europe, Australia, and North America. The device collects vehicle operational features like speed, acceleration, mileage, and driving times which are then used by the insurance companies or fleet managers to assess the driving safety of customers and employees.

**Study Setting**

This was a randomised controlled trial to assess the effectiveness of behavioural interventions in the risk behaviour of the drivers of light and heavy commercial public transport vehicles (taxis and busses). The taxi group comprised drivers operating in three southern counties of Tehran, namely Robat Karim, Eslamshahr, and Qarchak. The bus group comprised drivers operating intercity busses across Iran. The participants were all males aged above 20. A team of engineers installed and tested the devices in both taxis and busses starting in February 2017. Overall, 355 taxis and 1673 busses were fitted with the device by August 2017.

**Telematics Device**

This project utilises a bespoke Telematics device fixed inside a vehicle fuse box or under its steering wheel. The device is GPS enabled, has a 3-axis accelerometer, and can interface with a vehicle’s “On-board Diagnostics” (OBD) port where available. A schema of device data flow is presented in Figure 1. The system collects approximately 3 million data points per day. A visualisation of data point density for vehicles.

---

**Figure 1. Schematic of the Device and Study Data Flow.**
travelling at over 20 km/h collected over a 24-hour period is provided in Figure 2.

Study Phases
The study was organised in three stages:
1. Baseline Measurement and Randomisation – Driving data was collected for a 4-week period to establish a baseline measure of risky behaviour for each driver. The drivers of each vehicle type were then randomised into intervention and control groups. The randomisation process was designed to ensure similar distribution of baseline risk amongst drivers in each group.
2. Intervention Phase – Individual drivers were followed for a period of 8 weeks, during which they received weekly individual feedback based on their behaviour.
3. Post-intervention Phase – At the end of this phase, the drivers received an SMS text message informing them that they will not receive further weekly updates on their driving performance. Their performance, however, was monitored for another 4 weeks to ascertain the endurance of the behaviour change.

Intervention Details
Overview
The behavioural intervention in this study was designed to reduce the risky driving behaviour through social norms intervention, in which individuals are presented with the decisions other people make in a given situation. This approach uses peer comparison to influence individuals. This method has been shown to have greater effect on individual behaviour than other factors like biological, personality, familial, religious and cultural influences. There are 2 types of norms, descriptive norms and injunctive norms. The former indicates how most others act in a given situation, and the latter implies what is accepted or rejected by society.

Peer comparison feedback consists of providing individuals with feedback about their own function compared with the performance of others. This approach has shown to be effective in lowering residential energy use and improving physician performance in colorectal cancer screening programs. In this study, we expect that reminding drivers of their own performance, by providing them a descriptive social norm and demonstrating the behaviour of the best performing peers in their group will encourage them to improve their driving behaviour.

Risk Factor Identification and Scoring Scheme
To provide drivers with feedback on their performance relative to their peers, we devised a model based on their driving behaviour with which a driver can be scored. The factors considered here were:

a) Harsh Braking (HB) may indicate poor decision making in speed management and/or failure to maintain a safe distance to the car in front. Harsh braking may also cause the trailing vehicle to brake harshly and/or swerve to other lanes and thus contributes to increased crash risk. Having considered various publications on the subject, the figure for HB threshold was set at -0.4 g (km.h⁻¹.s⁻¹) for both light and heavy vehicles following the threshold recommended by Verizon.

b) Car Handling refers to factors such as “Harsh Turning” (HT) and “Harsh Lateral Angle”. The threshold of the former was set to 0.7 g (24.7 km.h⁻¹.s⁻¹). The latter of the 2 has not been yet considered at this stage for the purposes of scoring. Another factor considered in this category is “Harsh Acceleration” which may result in loss of control and was set at 0.22 g (7.8 km.h⁻¹.s⁻¹) following recommendations.

c) Speed Management – Speed violation (SV) is perhaps the most important contributing factor in risk of crash as well as its consequences. The safe vehicle speed is currently assumed to be the speed limit of the road on which a vehicle is travelling. These violations are detected off-line by comparing the GPS speed sent...
by the telematics device and its location to the posted speed limit of the road the vehicle is travelling on. Road geography, type, and speed data was obtained through OpenStreetMap (OSM) API and was used under 'Open Data Commons Open Database License'. Speed violations are divided into three classes based on severity, details of which have been included in Table 1.

d) Driver Fatigue (DF) is defined as a gradual and cumulative process whereby the driver experiences loss of efficiency and experiences deterioration in their vigilance, alertness, and overall performance. It must be noted that 'fatigue' terminology in the literature is used interchangeably with 'sleepiness', 'tiredness', and 'drowsiness'. Sleepiness, in particular, refers to difficulty in remaining awake and is determined by the 'body clock' and 'sleep homeostasis'. In this study, 'fatigue' and 'sleepiness' are considered to be interlinked and interact with each other. This is recognised by inclusion of 2 metrics: 1- Continuous driving of more than 240 minutes without a 20-minute rest period (fatigue); 2- Driving after 12:00 AM (sleepiness).

The formula for overall driving score per day per driver is then given by,

\[ S = S_{\text{max}} \times W \times W \]

where \( H \) and \( W \) are, respectively, the vectors of incident frequency and their associated weights and the subscript \( i \) refers to an individual vehicle. \( S_{\text{max}} \) is the maximum driving score possible and is constant throughout the scoring scheme.

The frequency of incidents is likely higher in drivers who have higher mileage. A driver who records 5 faults within 5 km, for instance, should be considered worse than one who records 5 violations in 50 km. In order to obtain a more meaningful means of comparison between drivers’ behaviour, the number of each incident for each period is normalised relative to the distance travelled in that period. For each driving-day, the incident per distance vector is given by,

\[ H = \left[ \frac{HA}{D_i}, \frac{HB}{D_i}, \frac{HT}{D_i}, \frac{SV_{i,1}}{D_i}, \frac{SV_{i,2}}{D_i}, \frac{SV_{i,3}}{D_i}, \frac{DF}{D_i}, \frac{ND}{D_i}, \frac{T_i}{T_{\text{lim}}} \right] D_{ij} > D_{\text{lim}} \]

where \( D_{\text{lim}} \) denotes a minimum distance which a vehicle has to have travelled per scoring period for its record for that period to be included in the scoring process and \( W \) refers to daily driving time.

The weight vector \( W \) is determined using a Neuro-Fuzzy process using an adaptive network-based fuzzy inference system (ANFIS) (20). The Fuzzy rules take the form of IF-THEN statements such as “IF \( SV_i \geq 3 \) THEN Very Poor” or “IF \( SV_i = 1 \) AND (\( HA < 2 \) OR \( HB < 2 \)) THEN Moderate”. Daily behaviour data of the participants were collected for a month and each driver-day set was labelled by a group of traffic injury experts. The labels were given on a 5-level scale (very poor to very good) each of which was mapped to a continuous scale representing driving scores. The labelled data was then used as a training dataset to tune fuzzy inference system (FIS) parameters which then produce a driving score based on new driving-day data for each participant.

**Driver Feedback**

Having ranked the drivers according the scoring scheme previously presented, each driver in the intervention group received a weekly feedback SMS. The message will detail the number of risk events during that week, their driving score (0–100 with 100 being the safest), and how they ranked in the previously presented, each driver in the intervention group having ranked the drivers according the scoring scheme.

The primary outcome will be the weekly driving scores from Saturday to Friday, reflecting the working week in Iran. Secondary outcomes will be the average number of weekly incidences of over-speeding, harsh brakes, turnings, and accelerations normalised by driving mileage.

**Enrolment and Randomisation**

The drivers considered for inclusion in this study have had the device installed on their vehicle for at least 3 months and were active for the 4-week period of baseline measurement. The 3-month period was chosen to minimise dropouts after the start of the intervention and for the remaining drivers to become accustomed to the having the Telematics device on their vehicle and for any behaviour change because of it to be normalised amongst the subjects. It additionally lowers the rate of dropouts because of early device failures due to

**Table 1. Overview of Driving Behaviour Factors Affecting Safety**

<table>
<thead>
<tr>
<th>Threshold or Range</th>
<th>Measurement Method</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;20 km/h over the speed limit</td>
<td>GPS speed</td>
<td>Speed Violation Class 3 (Severe)</td>
</tr>
<tr>
<td>10 to 20 km/h over the speed limit</td>
<td>GPS speed</td>
<td>Speed Violation Class 2 (Moderate)</td>
</tr>
<tr>
<td>Up to 10 km/h over the speed limit</td>
<td>GPS speed</td>
<td>Speed Violation Class 1 (Minor)</td>
</tr>
<tr>
<td>&gt;0.7 g</td>
<td>Accelerometer Y axis</td>
<td>Harsh Turning</td>
</tr>
<tr>
<td>&lt;=0.4 g</td>
<td>Accelerometer X axis</td>
<td>Harsh Braking</td>
</tr>
<tr>
<td>&gt;0.22 g</td>
<td>Accelerometer X axis</td>
<td>Harsh Acceleration</td>
</tr>
<tr>
<td>&gt;240 minutes of continuous driving*</td>
<td>Driving time</td>
<td>Driver Fatigue</td>
</tr>
<tr>
<td>Driving after 12:00 AM</td>
<td>Time of day</td>
<td>Night Time Driving</td>
</tr>
</tbody>
</table>

*Defined as continuous operation of the vehicle without taking a 20-minute break.
manufacturing defects. In order to minimise selection bias, treatment allocation was carried out automatically using an especially written script in MATLAB 2016b and inputted into the database by a technician independent from the study. The treatment allocation table used by the automatic SMS system for sending out feedback messages was then locked until the end of the study.

Out of the 355 taxis fitted with the device, 112 drivers met this criterion at the start of the intervention. Taxi drivers were randomised into treatment and control groups using stratified block randomisation method. The three strata were based on district of operation, each having the same number of intervention and control participants. The stratification was performed to account for different driving environments in which the drivers operated. The numbers of bus drivers meeting the inclusion criteria were 1309 and were randomised by simple randomisation. Randomisation quality was assessed by graphical checks and the results are presented in Figure 4. Overall, we detected no significant difference in mean baseline driving score between the control and intervention groups for either taxis or buses.

**Sample Size**
The statistical analysis is based on repeated measure design in which the driving score of each participant is measured at 10 points during the intervention phase. We included all participants who met the inclusion criteria prior to starting of the intervention phase. We calculated the smallest detectable difference between the intervention and control groups using the “Tests for Two Means in a Repeated Measures Design” option in the PASS 11 software (NCSS, LLC, Utah). A sample size of 1309 bus drivers, taking into account a 10% dropout rate, would allow the detection of difference of 1.0 driving point with power of 0.8 and type I error of 0.05. A sample size of 112 taxi drivers would allow the detection of 10.3 driving points with power of 0.8 and type I error of 0.05 considering a 10% dropout rate.

**Statistical Analysis**
Data relating to driver scores will be analysed using generalized estimating equations (GEE) for analysing longitudinal data with multiple measurements. The intervention effect will be analysed separately for bus and taxi groups. The unit of the analysis will be individual participants and the repeated measures will be their weekly driving scores at midnight of the seventh day of the week. The model will account for baseline driver scores and Taxis’ districts of operation. Drivers who drop out prior to the...
start of the intervention are excluded from the analysis. All drivers whose driving data is registered at least for 1 day during the intervention phase will be included in the analysis. Appropriate adjustments will be made in the analysis if attritions cause significant imbalance between control and intervention groups. The treatment effect on the primary outcome and its 95% confidence interval will be reported. The data related to secondary outcomes will be analysed by graphical checks and comparing the mean and standard deviation of each risk factor normalised to the distance travelled.

**Patient and Public Involvement**

The study participants were not involved in the design of this study.

**Conclusions**

In conclusion, the trend of road fatalities from RTAs in Iran and globally is decreasing. This, however, is not true for non-fatal injuries from RTAs. Implementation of traditional countermeasures, up to the coverage level necessary to reduce the rates of RTAs to that required in SDG goals, face challenges relating to governance, human resources, and capital investment in equipment and infrastructure. It is, therefore, necessary to evaluate other methods for promotion of safe driving and reduction of trauma resulting from RTAs. Behavioural interventions present a promising method to promote and improve safe driving behaviour. This study investigates the effect of peer comparison feedback, a non-monetary incentive, on drivers' behaviour using the data collected and sent over the mobile network through an in-vehicle telematics device. Success of this project may have several policy implications regarding road safety measures. The intervention may be scaled up to cover all public transport vehicles. The vehicle diagnostic information gathered by the device can also aid in lowering accident risk due to direct and indirect consequences of vehicle failure on the road. For example, our system allows for real-time data collected by the OBD port and is able to send a text message to alert the driver or the vehicle owner. As mentioned briefly in the introduction, driver behaviour is only one of the factors influencing the likelihood of RTAs. It must be noted that the driver is also expected to compensate for the inadequacies of the driving environment. For example, a poorly designed junction which blocks a driver's view to the oncoming traffic or abrupt changes in the number of highway lanes may increase the likelihood of RTAs. Collection and analysis of telematics data may aid in identification of such hazardous points by investigating the clustering of various risky behaviours on the road. Finally, the research infrastructure developed as a result of conducting this study may, in the future, allow more complex behavioural research such as the development of ‘Agent-based’ driver models.

**Authors’ Contribution**

FF, AJ, and NR designed the concept; FF, NR, MA, and MD designed the study; AJ and NR acquired the data; AJ and MA designed the scoring method; MA and PN processed the data; MA and PN analysed the data and performed the randomisation; MA and MD drafted the manuscript; FF, AJ, NR, KB, DB, JH, and SS revised the manuscript; FF approved the final manuscript for publication.

**Funding**

This study was funded by the National Institute of Medical Research Development (NIMAD) through the grant No.940576.

**Data Sharing**

No additional data is available. Behavioural data will be available upon request.

**Conflict of Interest Disclosures**

The authors have no conflicts of interest.

**Ethical Statement**

The participants were required to sign an informed consent form prior to installation of the device on their vehicle. Participants consented to installation of the device, taking part in a randomised controlled trial, and for their driving data to be used in other exploratory research projects. They were also informed that their driving data used for this study, real-time or aggregate remains confidential and is not disclosed to employers. This project has been ethically approved by the National Institute of Medical Research Development ethics code: IR.NIMAD.REC.1394.016. The trial is registered in at http://www. irtc. ir/ as IRCT20180708040391N1.

**Acknowledgments**

The authors would like to thank the National Institute of Medical Research Development (NIMAD) for funding this project by grant No.940576.

**References**

RCT Protocol of Big Data Telematics in Road Traffic Injuries


